International Conference on Advanced Manufacturing Engineering and Technologies

Conference Chair
Prof. Cornel Mihai Nicolescu (KTH Royal Institute of Technology, Stockholm)

Co-Chairs
Prof. Viorel Paunoiu (University of Galati, Romania)
Prof. Miroslav Piska (Brno University, Czech Republic)
Prof. Mauro Onori (KTH Royal Institute of Technology, Stockholm)

Scientific Committee
Prof. Bengt Lindberg (KTH Royal Institute of Technology, Sweden)
Prof. Lars Mattsson (KTH Royal Institute of Technology, Sweden)
Prof. Torsten Kjellberg (KTH Royal Institute of Technology, Sweden)
Prof. Lihui Wang (KTH Royal Institute of Technology, Sweden)
Prof. Marco Santochi (University of Pisa, Italy)
Prof. Reijo Tuokko (Tampere University of Technology, Finland)
Prof. Philippe Lutz (University of Franche-Comté, France)
Prof. Fabrizio Quadri (University of Rome Tor Vergata, Italy)
Prof. Paul Shore (Crainfield University, UK)
Prof. Laszlo Monostori (Budapest University of Technology and Economics, Hungary)
Prof. Lars Nyborg (Chalmers University of Technology, Sweden)
Prof. S.Jack Hu (University of Michigan, USA)
Prof. José Barata (New University of Lisbon, Portugal)
Prof. Patrick Martin (ENSAM, France)
Prof. Takayuki Hama (Kyoto University, Japan)
Prof. Trevor Dean (The University of Birmingham, UK)
Prof. Gino Dini (University of Pisa, Italy)
Prof. Luis Norberto Lopez De La Calle (Technical School of Engineering of Bilbao, Spain)
Prof. Satyandra K. Gupta (University of Maryland, USA)
Prof. Krzysztof Jemielniak (Warsaw University of Technology, Poland)
Prof. Johan Stahre (Chalmers University of Technology, Sweden)
Prof. Jerzy Jedrzejewski (Wroclaw University of Technology, Poland)
Prof. Paulo E. Miyagi (University of São Paulo, Brasil)
Prof. Niels Bay (Technical University of Denmark, Denmark)
Prof. Adinel Gavrus (National Institute of Applied Sciences, France)
Prof. Antonio Gonçalves Coelho (New University of Lisbon, Portugal)
Prof. Kamal Youcef-Toumi (Massachusetts Institute of Technology, USA)
Prof. Wit Grzesik (Opole University of Technology, Poland)
Prof. Patricio Franco (Technical University of Cartagena, Spain)
Prof. Terje Lien (Norwegian University of Science and Technology, Norway)
Prof. Vytautas Ostasevicius (Kaunas University of Technology, Lithuania)
Prof. P.G. Maropoulos (University of Bath, UK)
Prof. Francisco Restivo (University of Porto, Portugal)
Prof. George Chryssolouris (University of Patras, Greece)
Prof. Shiv G. Kapoor (University of Illinois at Urbana-Champaign, USA)
Prof. Dimitris Mourtzis (University of Patras, Greece)
Prof. João Paulo Davim (University of Aveiro, Portugal)
Prof. Hong Hocheng (National Tsing Hua University, Taiwan)
Prof. Loredana Santo (University of Rome, Tor Vergata, Italy)
Prof. Luis Gomes (New University of Lisbon, Portugal)

Organizing Committee at Royal Institute of Technology
Dr. Andreas Archenti (Editor)
Dr. Antonio Maffei (Editor)
Tech Lic. Hakan Akillioglu
Dr. Danfang Chen
Dr. Lorenzo Daghini
Mr. Joao Ferreira
Mr. Costantinos Frangoudis
Tech Lic. Pedro Neves
Mr. Johan Pettersson
Ass. Prof. Amir Rashid
Mr. Tomas Österlind
Conference Partners

The organizing committee would like to express their deep gratitude to all the partners for their active support and contribution, without which this event would not have been possible.
Preface to themes treated in this volume

The theme of the conference is “Advanced Manufacturing Engineering and Technologies”, and the aim of this initiative is providing a forum for researchers and practitioners working on the diverse issues of such a broad topic. In particular, authors, both from academia and industry have been invited to submit papers for all aspects of theories, methodologies, applications, and case studies related to their work in this context.

This volume collects the papers treating the following themes: Forming, Assembly and Automation Technology, Material Science, Additive manufacturing, Welding, Operation Management and Inspection and Quality Assurance.

Production Engineering is often defined as the “decathlon” of engineering sciences. To be able to produce in a sustainable manner and still answer to the needs of the market, industrial enterprises need a deeper understanding about manufacturing strategies. Material properties, improved analysis and design techniques as well as operation management and automation technologies contribute to the creation of a holistic view for the integration of the critical processes and components.

New materials are required to produce future resource efficient and complex products, for instance low emission vehicles. The optimization of the production process in its whole requires deep knowledge of the material properties, which, in turn, affects the way a component is produced, either through traditional methodologies, such as forming or welding, or through emergent techniques such as additive manufacturing. Assembly and automation technologies have also to adapt to these dynamically changing demands and, last but not least, more demanding quality requirements put a burden on quality assurance methodologies.

This volume is an attempt to cover these themes and give an overview of the latest developments in these fields.

The Editors
Dr. Andreas Archenti & Dr. Antonio Maffei
# Table of contents

**Theme 4: Forming**

- Holistic approach to pulse magnetic forming of magnesium alloy AZ31 at low forming Temperatures ............................................... 11
- Mechanical Properties Characterization of Tin-Lead Open-Cell Foams using Upsetting Experimental Tests and Finite Elements Modelling .......................................................... 25
- Numerical modeling of magnetic induction and heating in injection molding tools .......................................................................................................................... 35
- New Concepts for Offline Dimensional Control in Sheet Metal Forming ................................................................................................................... 47
- Numerical optimization of die geometry in open die forging ................................................................. 59
- Design of Experiment for Roll-to-roll Hot embossing Process ......................................................... 69
- Feeding and Positioning of Linked Parts in Micro Production Chains ............................................. 75
- Influences on the FEM-results of different parameters of the rotating straightening process ................................................................. 85
- Finite Element Analysis on the Friction Effects in the Gear Rolling Process ......................................... 93

**Theme 5: Assembly and Automation Technology**

- Electronic Component Cleaning in Remanufacturing ............................................................................. 105
- Vision-assisted and 3D model-based remote assembly ..................................................................... 115
- Minimising energy consumption for robot arm movement .................................................................. 125
- Detection of Battery Screwdriver’s Optimal Working Regimes .................................................. 135
- Supply chain collaboration on the competitiveness of Basque Country manufacturing companies .............................................................................................................. 141

**Theme 6: Material Science, Additive manufacturing and Welding**

- A methodology for parameter setting in the laser cladding by process simulation ..................................... 153
- A novel strategy for the incorporation of optical sensors in Fused Deposition Modeling parts ................................................................. 163
- Cooling Rate Prediction in case of Pipelines Longitudinal Welds performed by Submerged Multi-Arc Welding ................................................................. 171
- A new approach to modelling friction stir welding using the CEL method ......................................... 179
- Modelling the mechanical deformation of nickel foils for nanoimprint lithography on double-curved surfaces ......................................................................................... 191

**Theme 7: Operation Management**

- A Genetic Algorithm For Simultaneous Scheduling Problem In Flexible Flow Shop Environments With Unrelated Parallel Machines, Setup Time And Multiple Criteria ........................................................................................................................................................................... 203
- The ICTs in the Extended Enterprises ........................................................................................................ 213
- Reference Model of Manufacturing Resources ......................................................................................... 221
- Towards a Downtime Cost Function to Optimise Machine Tool Calibration Schedules ................................................................. 231
- A case study of long-haulage transport in China for a sustainable transport and manufacturing business ........................................................................................................................................................................... 241

**Theme 8: Inspection and Quality Assurance**

- Barkhausen Noise Analysis of Surface Integrity in Grinding of Large Bearing Rings ................................................. 253
- New Technologies for Individual Joint Implants ......................................................................................... 263
- A Roadmap to Model Based Manufacturing ............................................................................................ 273
- A Study of the Surface Integrity after Machining by means of Non-Destructive Testing Methods ........................................................................................................................................................................... 283
Impact Acoustic Testing as NDT Method for Classification of Compacted Graphite Iron ....................................... 293

**Theme 9: Industry oriented session** ........................................... 303
Tolerance chain design and analysis of in-process workpiece ......................................................... 305
Nonparametric identification of stiffness and damping in nonlinear machining systems ......................................................... 317
Using concept modelling enables improvement system development ......................................................... 329
The IDEAS Plug & Produce System ........................................................................................................... 339
Laser assisted milling of a nickel based alloy using different insert geometries ......................................................... 347
Theoretical and Experimental Studies of Internal Turning Using Damped Boring Bars ......................................................... 357
Steps Towards Holistic Modelling of HSC Machining Centre Errors ......................................................... 369

**Theme 10: New research endeavors** ........................................... 379
Advanced cutting materials for hard turning ......................................................... 381
Wear and renewal of cutting tools properties ......................................................... 391
The reverse engineering to optimize the assembly of a conical spur gear by CAD ......................................................... 397
International Conference on Advanced Manufacturing Engineering and Technologies

Theme 4

Forming

Holistic approach to pulse magnetic forming of magnesium alloy AZ31 at low forming temperatures
Eckart Uhlmann, Lukas Prasol
Institute for Machine Tools and Manufacturing Technology, TU Berlin, Germany

Mechanical Properties Characterization of Tin-Lead Open-Cell Foams using Upsetting Experimental Tests and Finite Elements Modelling
Abd-Elmouneim BELHADJ1, Adinel GAVRUS2, Mohammed AZZAZ1, Fabrice BERNARD2
1Laboratory of Science and Engineering of Materials (LSGM), University of Science and Technology Houari Boumediene (USTHB), Algeria
2 National Institute of Applied Science (INSA Rennes), UE, Laboratory of Civil Engineering and Mechanical Engineering (LGCGM), Rennes, France

Numerical modeling of magnetic induction and heating in injection molding tools
Patrick Guerri Ferrari Jesper H. Hattel
Technical University of Denmark, Department of Mechanical Engineering, Section of Manufacturing Engineering, Denmark

New Concepts for Offline Dimensional Control in Sheet Metal Forming
Viorel Paunoiu, Alexandru Epureanu
Dunarea de Jos University of Galati, Department of Manufacturing Engineering, Romania

Numerical optimization of die geometry in open die forging
Peter Christiansen1, Jesper H. Hattel1, Niels Bay1, Paulo A.F. Martins2
1Department of Mechanical Engineering, Technical University of Denmark, DTU, Denmark
2Instituto Superior Tecnico, Technical University of Lisbon, Portugal

Design of Experiment for Roll-to-roll Hot embossing Process
Dongwon Yun, Youngsu Son, Heechang Park, Byungin Kim, Sangyong Ham, Seyoung Kim
Korea Institute of Machinery & Materials (KIMM)

Feeding and Positioning of Linked Parts in Micro Production Chains
Philipp Wilhelmi, Eric Mouni, Eike Foremny, Bernd Kuhfuss, Christian Schenck
University of Bremen, bime - Bremen Institute for Mechanical Engineering, Germany

Influences on the FEM-results of different parameters of the rotating straightening process
Graziella Kreiseler
Fraunhofer Institute for Manufacturing Engineering and Automation, Germany

Finite Element Analysis on the Friction Effects in the Gear Rolling Process
Alireza Khodaei, Arne Melander
KTH Royal Institute of Technology, Sweden
Holistic approach to pulse magnetic forming of magnesium alloy AZ31 at low forming temperatures

Eckart Uhlmann, Lukas Prasol

Institute for Machine Tools and Manufacturing Technology, TU Berlin, Germany
Corresponding author: prasol@iwf.tu-berlin.de

ABSTRACT
Magnesium alloys are suitable for lightweight constructions due to their low density of \( \rho = 1.8 \text{ g/cm}^3 \). Due to the lattice structure of magnesium (hcp) a relatively brittle material behaviour results. Therefore it is necessary to form magnesium alloy AZ31 at elevated temperature of 220°C. Pulse magnetic forming is an innovative forming technology to improve the forming behaviour of magnesium alloy AZ31. Due to high strain rates and induction process, both process-related, there is a local change of thermodynamic conditions in the sheet metal. In this paper the systematic study of the behaviour of the magnesium alloy AZ31 with pulse magnetic forming at low forming temperature is presented. First, experimental investigations with a suitable experimental setup were progressed. Analogously a FE model, including all important physical domains, was developed. The obtained results show a good agreement between experimental and numerical investigations.

KEYWORDS: pulse magnetic forming, magnesium, low forming temperature, FE simulation
1. INTRODUCTION

Different national and international guidelines, e.g. CO$_2$ reduction in automotive industry or weight reduction in aviation industry, as well as the E-mobility sector require new lightweight components and associated with these new production technologies. In comparison to steel components lightweight components like magnesium alloys, aluminium alloys or CFRP exhibit lower density and therefore lower mass.

The use of different materials, a so-called “multi-material design”, was investigated in the EU project “Super Light Car”. The mass of the front end structure of a vehicle was reduced for a total of 35 % of 100 kilograms, thus the emission of CO$_2$ could be reduced to 8.4 gram per 100 kilometer. This value corresponds to a fuel economy from 0.3 to 0.5 liters [1]. Especially in the area of E-mobility the total weight of the vehicle structure exhibits a very important role. The average energy density of a lithium-ion battery is 0.10 kWh/kg, that of a gasoline engine which was fuelled with premium gasoline 12.0 kWh/kg. To achieve comparable distances with an electric motor an accumulator with a huge mass is required. The use of lightweight components at different areas in the vehicle body is therefore essential to reduce the whole weight of the vehicle. An efficient and environmentally sustainable processing of lightweight components is a production-technical challenge.

Production processes have to meet the requirements of these tendencies and must be suitable for mass production. Forming processes with optimal utilization of material and high productivity offer potential for excellent accuracy. Forming process of magnesium alloys is accomplished at high temperatures of 220°C currently. A general challenge in forming of magnesium alloys will be in the realization of forming processes at lower temperatures.

In the last years the pulse magnetic forming processes have gained an increasing attention from both manufacturing companies and research facilities [2, 3]. A considerate advance has been made in the field of simulation in the recent years providing a wide field of software as well as the necessary hardware. In this paper a holistic approach to pulse magnetic forming of magnesium alloy AZ31 at lower forming temperature is presented.

2. PROCESS PRINCIPLES

The pulse magnetic forming process is based on the physical effect of induction [4]. The energy which is necessary for the forming process of the sheet metal is stored in a bank of capacitors by charging them to a high voltage U. By discharging the capacitors over a high-current switch, the arising large currents $I_{\text{ind}}(t)$ generate an intense magnetic field $H(t)$ outside the tool coil with the magnetic flux density $B = \mu H$ (see Fig. 1.). This magnetic field $H(t)$, whose effective duration depends on the process time $t_p$, induces eddy currents $I_{\text{eddy}}(t)$ in the workpiece which are running in the opposite direction compared to the primary currents $I_{\text{ind}}(t)$ in the tool coil (see Fig. 1).
Due to high frequencies $f_p$ the skin effect\(^1\) is causing that the induced eddy currents $I_{\text{eddy}}(t)$ are running at the surface – exactly in depth $t_{\text{skin}}$ which is determined by the so-called ‘skin depth’ – of the workpiece. Consequently the resulting Lorentz forces $F_L(t)$ which depend on the primary magnetic field $H(t)$ are acting for a short time of 50 µs to 100 µs on the surface of the workpiece as a magnetic pressure $p_{\text{mag}}$. As now workpiece and tool coil repel each other, the workpiece will be deformed. Here, the material dependent yield stress $k_f$ is exceeded and as a consequence plastic deformation of the workpiece takes place within a few microseconds.

Pulse magnetic production processes are assigned to high speed forming processes. The forming process is realized without any mechanical contact between workpiece and tool coil by the stored energy. In principle there are three different process variants – compression, expansion and flat forming. Depending on the forming process different tool coils are used (see Fig. 2).

---

\(^1\) Skin effect is the tendency of an alternating current to become distributed within a electrical conductor such that the current density is largest near the surface of the conductor, and decreases with greater depths in the conductor.
3. STATE OF THE ART

Among metallic construction materials magnesium is the one with the lowest density. Due to this magnesium alloys exhibit a low mass. The magnesium alloy AZ31, main alloy aluminium (about 3%) and zinc (about 1%), exhibits a weight saving of 30% in comparison to aluminium and a saving of 75% in comparison to steel alloys. The specific strength of AZ31 reaches significantly higher values in relation to aluminium or steel alloys, see Fig. 3.

![Fig. 3. Illustration of specific strength of magnesium alloy AZ31 in comparison to different material [5].](image)

Magnesium and its alloys, e.g. AZ31, have hexagonal lattice structure (hcp). Materials with this lattice structure posses at room temperature (20°C) following sliding and twinning systems: basal and prismatic slip, pyramidal slip of first and second order as well as twinning (shear twinning or compression twinning). In dependence of different critical shear stresses $\tau_{\text{CRSS}}$ of these systems [6] the so-called ‘von Mises criterion’ – which describes a homogenous deformation of polycrystal and requires at least five linear independent sliding or twinning systems [7] – is fulfilled by superposition of sliding (basal, prismatic or pyramidal) and twinning. As a result of this magnesium alloy AZ31 has a relatively brittle material behaviour at room temperature and fails at low strains [8, 9, 10, 11].

Changing the thermodynamic boundary conditions, e.g. by supplying energy, cause the activation of further slip planes with different sliding directions. As a consequence more than five independent sliding systems are available. By these much higher strains $\varepsilon$ can be reached without material failure.

DROEDER showed in tensile tests with magnesium alloy AZ31 that by increasing the workpiece temperature at 235°C and strain rate $\dot{\varepsilon} = 0.002 \text{ s}^{-1}$ the logarithmic deformation could be doubled [5]. Strain rates $\dot{\varepsilon}$ illustrate a further thermodynamic boundary condition in forming processes. The principal effect of different strain rates $\dot{\varepsilon}$ on the formability $\varphi$ of metallic materials is shown in Fig. 4.
At high strain rates $\dot{\varphi}$ a decrease of yield stress $k_f$ which is associated with an increase of formability $\varphi$ occurs. Due to the short process time $t_p$ the generated heat $Q$ cannot be dissipated quickly enough from the local forming area. Heat $Q$ occurs as a result of internal friction during the forming process. \cite{12} had shown in experiments with magnesium alloy AZ80 that an increase in the degree of compression fracture occurs by increasing strain rate $\dot{\varphi}$, regardless of the workpiece temperature. This effect is observed at strain rates from $\dot{\varphi} = 2,000 \, \text{s}^{-1}$. This means that the quasi-adiabatic character of the process is lasting on the compression process.

During the pulse magnetic forming of metallic materials strain rates $\dot{\varphi}$ occur up to $\dot{\varphi} = 25,000 \, \text{s}^{-1}$. The maximum magnetic pressure $p_{\text{mag}}$ at the surface of the workpiece reaches up to $p_{\text{mag}} = 1,000 \, \text{MPa}$. The process dependent strain rates $\dot{\varphi}$ reach significantly higher values than the previously study strain rates $\dot{\varphi}$. Hence it can be expected that a significant increase in the formability of magnesium alloy AZ31 will occur at room temperature \cite{13, 14}.

### 4. EXPERIMENTAL PART

For the experimental study of pulse magnetic forming of magnesium alloy AZ31 at room temperature the following axis-symmetric experimental setup was designed, see Fig. 5 a). The variably designed experimental setup makes it possible to investigate the influence of different parameters which are explained in the following.

#### 4.1. Influencing Factors

As to be investigated parameters the charging energy $E$, the die diameter $D$, the radius $R$ of the drawing edge and the friction $\mu$ between die and workpiece are identified.

- An increase of the charging energy $E$ causes that during the same process time $t_p$ more energy $E$ is introduced into the workpiece due to excessive current thereby an increase of deformability $\varphi$ occurs.
- An increase of the die diameter $D$ causes higher effective magnetic pressure $p_{\text{mag}}$ acting on the workpiece. Due to the coil geometry (axis-symmetric, see Fig. 5.) in the area of the coil centre a significant decrease of the magnetic pressure $p_{\text{mag}}$ occurs. Hereby a lower magnetic pressure $p_{\text{mag}}$ is acting on the workpiece while using smaller die diameter $D$.
- A decrease of the drawing radius $R$ causes that the material flow along the drawing radius is favoured. For smaller drawing radii $R$ a failure of the shaped material occurs.
- The friction between $\mu$ the workpiece and blank holder can be minimized by using lubricants. Thereby the flow of the material along the blank holder is facilitated.
4.2 Experimental Results

The formability of magnesium alloy AZ31 was investigated at room temperature with the experimental setup which is shown in Fig. 5 a). The forming process was realized by pulse magnetic forming with the pulse generator FA-60-1440-SW Magnepuls, see Fig. 5 b). The maximum realizable forming height $h$ as well as the maximum strain $\varepsilon$ of the sheet metal are determined by charging energy $E$, die diameter $D$, drawing radius $R$ and the friction ratio sheet metal/die with constant sheet metal thickness $d = 1.5$ mm (see Fig. 6). Height $h$, strain $\varepsilon$ as well as the sheet metal thinning were measured with an optical measuring system (GOM ARGUS).

![Fig.5. Schematic drawing of the experimental setup for pulse magnetic forming of magnesium alloy AZ31 a) as well as the pulse generator FA-60-1440 SW Magnepuls b).](image-url)

With increasing charging energy $E$, that means with increasing magnetic flux density $B(t)$, a higher magnetic pressure $p_{mag}$ is acting on the workpiece. Hence an increase of realizable forming height $h$ in $z$ direction occurs in dependence of die diameter $D$ and drawing radius $R$.

The use of lubricant (industrial grease) leads to a minimization of friction $\mu$ between sheet and die. Hereby the material flow along the die ($x$ direction) is favoured and therefore no material failure occurs at comparable forming heights $h$.

An increase of the die diameter of $D = 50$ mm to $D = 80$ mm results in an enormous gain of realizable forming height $h$ while using lubricant. This increase is motivated by the fact that...
the use of such a tool coil leads to higher magnetic pressure $p_{\text{mag}}$ when using a die with larger diameter $D$. Therefore higher Lorentz forces $F_L$ acting during the process on the workpiece.

A larger drawing edge radius $R$ favours also the material flow into the die. Hence better forming results are achieved. If the pulse magnetic forming process is realized without lubricant, even at low charging energies $E$ cracks occur in the forming area, e.g. along the drawing edge. Fig. 7 shows the measured strains $\varepsilon$ and also the thinning of the sheet metal in the region of maximum forming height $h$.

The measurements illustrate that with increasing charging energy $E$ significant increase of strain $\varepsilon$ in the forming area takes place. Furthermore it is determined that only a slight thinning of the sheet metal within the forming process occurs. The determined strains $\varepsilon$ assume that the increase of the workpiece surface is dominated by the material flow in the deformation area but not by thinning of the material. In this way a high speed deep drawing process of magnesium alloy AZ31 is realized.

The obtained results show that a pulse magnetic forming process of magnesium alloy AZ31 at low forming temperatures significantly improves the formability of the sheet metals. Furthermore the obtained results demonstrate that during the forming process a material flow into the die takes place, and thus a high speed forming deep drawing process is realized.

5. FE-SIMULATION

A detailed description of the whole pulse magnetic forming process requires the consideration of all physical domains within the simulation. This requires an implementation of the electromagnetic, the thermal and the structural domain as well as the discharging circuit. For the FE simulation, the commercial software ANSYS as well as ANSYS APDL are used. A strong coupling (coupled simulation) of physical domains is carried out, see Fig. 8.
For a detailed illustration of the whole pulse magnetic forming process it is necessary to map the discharging process of capacitors with an equivalent circuit, see Fig. 9. As a result of the equivalent circuit the results of the discharging process are given to electromagnetic simulation. The results of the electromagnetic simulation (Lorentz force \( F_L \)) are given to the thermal simulation. In the final step the results of the electromagnetic and the thermal simulation are given to structural simulation. As a yield function a combination of the HILL model and the PERZYNA model was selected [15, 16]. This approach enables the combination of the anisotropy which occurs in HCP materials (AZ31) and also the strain rate-dependency caused by the process in ANSYS Implicit. The strength differential (SD) which occurs in AZ31 appears at higher strains weak [17]. So it is possible to neglect this effect and use HILL approach to map the anisotropy.

Based on the discharging current (see Fig. 9) the required parameters \( L_i \) and \( R_i \) were calculated using the following equations:

\[
I(t) = \hat{I} \cdot \sin(\omega t) \cdot \exp(-\delta t)
\]  

Equation (1) is the analytic function of the discharging current \( I(t) \) with

\[
\omega = \sqrt{\omega_0^2 - \delta^2}, \quad \omega_0 = \frac{1}{\sqrt{LC}}, \quad \delta = \frac{R}{2L}.
\]

Here, \( \omega \) demarks the angular frequency, \( \delta \) the damping constant and \( \omega_0 \) the resonant frequency. The only entity still missing in the equivalent circuit diagram is the inductivity of the tool coil.
It can be determined by measurement of the discharging current in the same way the inner parameters have been determined, taking into account, that
\[ L = L_i + L_f, \]
\[ R = R_i + R_f. \]  

Strong coupling method is based on the fact that within each time step all domains are solved. This procedure is iterative. In addition there is an iterative re-meshing of the distorted mesh. Based on the experimental setup, see Fig. 5, the following 2D experimental setup is simulated as shown in Fig. 10. All shown areas are relevant, such as the axis-symmetric tool coil (material: cooper), sheet metal (material: magnesium alloy AZ31), die with a defined drawing radius (material: steel) and the surrounding airspace. For simulation a die with defined inside diameter \( D = 80 \) mm and a defined die drawing radius \( R = 10 \) mm was selected.

![Fig10. Illustration of simulated experimental setup.](image)

### 5.1. Simulation Results

**Electromagnetic simulation**

The current flow \( I_{ind}(t) \) in the tool coil is determined as a boundary condition for electromagnetic simulation. The electromagnetic simulation is used to determine Lorentz forces \( F_L \) acting during the process on the workpiece, which results as plastic deformation of the workpiece. Due to the ‘skin effect’ which occurs at high frequencies \( f_p \) and induced eddy current flow \( I_{eddy}(t) \) at the workpiece surface, Lorentz force \( F_L \) act as a volume force on the workpiece. The following Fig.11 shows the whole acting Lorentz force \( F_L \) on the workpiece as well as the chronological characteristics of the whole induced eddy current \( I_{eddy}(t) \) in the workpiece.
Thermal simulation

Due to the mutual induction an eddy current $I_{\text{eddy}}(t)$ is induced in the workpiece. Because of the ohmic resistance $R_2$ of the sheet metal a heat generation takes place. As a result temperature rises in the sheet metal. The heat distribution along the surface causes a temperature gradient across the sheet metal thickness. The highest temperatures occur at the surface where the eddy current $I_{\text{eddy}}(t)$ is running. The following Fig.12 shows the chronological characteristics of JOULE dissipation in the workpiece.
Fig. 12. Chronological characteristics of JOULE dissipation in the whole workpiece.

*Structural simulation*

The Lorentz force $F_L$ is given as a boundary condition to structural simulation. The required material data for the yield function $k_f$ by HILL and PERZYNA were determined in tensile tests under different boundary conditions varying strain rates $\dot{\varphi}$ ($10^{-3}\,\text{s}^{-1}, 10^{0}\,\text{s}^{-1}, 10^{3}\,\text{s}^{-1}$) temperature (20°C, 150°C, 250°C) and texture (rolling direction (RD), 45° regarding to rolling direction, transversal direction (TD)).

In the next step the adaption of the HILL and PERZYNA parameters to experimental data was carried out. The optimization of numerical data in relation to experimental data was done by least square method (LSM). A correlation of 98% between numerical and experimental data (stress-strain behaviour) was reached. Optimization of stress-strain behaviour (RD, 45°RD, TD) was carried out for the following approaches for HILL model: $10^{-3}\,\text{s}^{-1}$ and 20°C; $10^{-3}\,\text{s}^{-1}$ and 150°C; $10^{-3}\,\text{s}^{-1}$ and 250°C. Afterwards the verification of HILL parameter was done (see Fig. 13). In the next step the optimization of stress-stress behaviour (RD) was carried out for PERZYNA approach: $10^{0}\,\text{s}^{-1}$ and 20°C; $10^{0}\,\text{s}^{-1}$ and 250°C; $10^{3}\,\text{s}^{-1}$ and 20°C; $10^{3}\,\text{s}^{-1}$ and 250°C. Afterwards the verification of PERZYNA parameter was done (see Fig. 13).
Fig. 13. Verification of HILL parameter (upper image); verification of PERZYNA parameter (lower image).

Fig. 14 shows the simulated maximum height h in comparison to experimental data. The numerically calculated height h was determined by the displacement of a defined node (point A) on the sheet metal surface (see Fig. 10).

Fig. 14. Comparison of numerical and experimental results (maximum forming height h).

Fig. 14. illustrates that the simulated height h achieves a good correlation with experimental results. The biggest difference (height h) between simulation and experimental data is about 12%. Furthermore, the simulation results show that the influence of heating (JOULE heating)
occurs due to induced currents in the workpiece. Hence a decrease of yield stress $k_f$ and an increase of formability $\varphi$ during the process occur.

6. CONCLUSION

The results in this article show the influence of defined parameters to pulse magnetic forming of magnesium alloy AZ31 at room temperature. The experiments were carried out with an axis-symmetric tool coil. This experimental setup allows a systematic study of defined parameters: charging energy, die diameter, drawing edge radius and friction between sheet metal and blank holder. An increase of charging energy causes higher currents to flow through the tool coil with the result that a larger magnetic pressure occurs and thereby higher deformations are reached. By reducing the friction factor between workpiece and blank holder the forming process is facilitated. In this way a continued flow of the material into the die is favoured. The pulse magnetic forming process enables a high-speed deep-drawing process of magnesium alloy AZ31.

A detailed description of the whole pulse magnetic forming process requires the consideration of all physical domains within the simulation. For this purpose a FE simulation with ANSYS Implicit was carried out. As a yield function a combination of the HILL model and the PERZYNA model was selected. This yield function considers the anisotropy of magnesium alloy AZ31 which occurs in HCP materials as well as the strain-rate dependency caused by the process. The simulation results show a good agreement between simulation and experimental data (maximum difference about 12%). Especially the results of thermal simulation show an increase of temperature in the sheet metal due to ohmic heating. Hereby the formability of magnesium alloy AZ31 increases significantly.

In the next step the influence of strain rate compared to the influence of JOULE heating will be investigated experimentally and numerically. Hereby a precise delimitation of pulse magnetic forming compared to conventional high-speed forming processes (such as explosive forming) is possible.
References:


Mechanical Properties Characterization of Tin-Lead Open-Cell Foams using Upsetting Experimental Tests and Finite Elements Modelling

Abd-Elmouneïm BELHADJ¹, Adinel GAVRUS², Mohammed AZZAZ¹, Fabrice BERNARD²

¹ Laboratory of Science and Engineering of Materials (LSGM), University of Science and Technology Houari Boumediene (USTHB), BP 32, El-Alia, Bab-Ezzouar 1611, Algeria
² National Institute of Applied Science (INSA Rennes), UE, Laboratory of Civil Engineering and Mechanical Engineering (LGCGM), 35708 Rennes, France
adinel.gavus@insa-rennes.fr

ABSTRACT

The last few years, many scientific researches are interested to elaboration of metallic foams with open-cell structure, to analyse their mechanical or physical properties and to apply these new materials in several manufacturing process or industrial applications. This paper presents an analysis of mechanical properties of tin-lead open-cell foams using experimental upsetting tests. New analytical models are proposed to describe the variation of the plateau stress ratio with the relative density and a finite element modelling is used to a better understanding of the densification process during the compression process. Isotropic hardening crushable foam’s model has been chosen to describe the compressible material behaviour.

KEYWORDS: Metallic foams, Open-cell structure, Mechanical characterization, Plateau stress vs. relative density relationships, Finite Element modelling

1. INTRODUCTION

During the last twenty years the goal of minimization of the engineering structure’s mass simultaneously with the increased of its mechanical performances has been required the design of advanced materials. Consequently high interests of industrial and academic researches have been focused on the elaboration of a class of metallic foams represented by porous metals with a high porosity varying from 50 to 97% and relatively small density. Then new manufacturing technologies have been recently developed with a wide range of applications for the industrial conception of impact energy absorbers (Fig.1a), heat exchangers (Fig.1b), dental implants (Fig.1c), silencers (Fig.1d), components for drying processes, air bearings, oil-lubricated bearings or some novel engineering structures.
Fig.1. Examples of industrial applications using metallic foams: a) shock absorbers, b) cooler devices, c) dental implants of porous titanium foam, d) sound dampers or heat exchangers

Over the last three decades, different innovative technologies have been proposed for producing metallic foams: the foaming of a liquid metal either by injecting a gas [1] or by the decomposition of gas releasing particles [2], by replication techniques of a polymer [3], known that an open cell foam precursor, by investment casting using as model a polymer foam [4] or from a porous material infiltrated by a liquid metal. In order to obtain a high porosity, metallic foam with open cell has been produced using an infiltration process, forcing a molten metal to flow in a porous network [5-7]. This paper will be focused on this class of materials and presents the mechanical properties of Tin-Lead open cell foams starting from liquid metal produced by replication using a NaCl perform.

In a first way, a better understand of the principal mechanical strength properties and of the compressibility behaviour requires experimental compression tests, performed for samples with various initial relative densities. It is then possible to evaluate the response of Sn-Pb foam in terms of the stress-strain curves and its dependence on the chosen relative densities $\rho_r$. The obtained experimental data permits to conclude that the shape of the experimental curves fit the classical description of foam materials. Moreover the results show that the compressive stress-strain curves of Tin-Lead foams exhibit three distinct variation domains: a linear elastic one, where the stress rises linearly with increasing strain, a plastic “plateau” region with slight stress fluctuation over a wide range of the plastic strain and a densification region where the stress increases rapidly. It is also observed that the plastic collapse stress of Tin-Lead foams rises with increasing of the $\rho_r$.

Concerning the stress “plateau” $\sigma_p$, a rigorous analysis is proposed to describe the influence of different $\rho_r$ values. Identification by a non-linear regression is used to compare the classical Gibson–Ashby model [8], a simple Avrami model [10] and a generalized Avrami formulation [11]. This work gives important improvements of previous studies [12] by proposing new physically based relationships defining the $\sigma_p/\sigma_0$ variation (where $\sigma_0$ is the plastic yield stress of the dense state of the material) as a function of $\rho_r$. The principal aim is to obtain a better prediction of the curve in the whole range of variation of the relative density: from the smallest values to the higher ones using an asymptotic variation close to the value of the dense material. Exploration of a numerical simulation to obtain the variation with the plastic strain during the compression process will be presented in the last part.

2. EXPERIMENTAL ANALYSIS

In order to a better understanding of the mechanical and physical properties of porous metallic materials with different relative density values, this study is focused on tin–lead open-cell foams obtained by an infiltration technique, forcing a molten metal to flow in a porous NaCl salt network, analysing essentially the mechanical behaviour under quasi-static compression loads.
2.1. Specimens elaboration

The Sn–Pb specimen foams were prepared from an alloy with a content of 50% tin and 50% lead heated at a temperature of 180°C in order to benefit from the fluidity of the binary Sn38–Pb62 eutectic structure (38 % of Sn in weight and respectively 62 % of Pb). The used elaboration method requires the manufacturing of several salt performs having different grain sizes \( g^* \). Then the molten metal infiltrates the salt mould during the foundry process followed by a cooling at the room temperature and dissolution of the salt in water. After the solidification process, open-cell metallic foams with different porosity are obtained with a structure corresponding to the negative of the salt perform using then a “negative” replication process. Figure 2 present four cylindrical specimens with a diameter of 30 mm and a small height equal to 6 mm, obtained from salts with different grain sizes in order to vary the porosity and the relative density \( \rho_r \). More details concerning the elaboration process and confirmation of their reproducibility are presented in [12].

![Fig.2. Different specimens of Sn–Pb open-cell foams obtained by liquid metal infiltration from various salt grain sizes \( g^* \) [12].](image)

2.2. Uniaxial upsetting tests and mechanical characterisation

On the mechanical point of view it is very important to describe correctly the compressibility behaviour. It is then necessary to quantify the mechanical strength characteristics of the foam materials, taking into account the influence of the initial density value. This last variable depends strongly on the main manufacturing parameters defined by the initial grains sizes controlled via an adequate choice of the salt performs morphology. Consequently, to analyse the macroscopic mechanical behaviour of the Sn-Pb foams, upsetting experimental tests were carried out using an INSTRON machine at a constant deformation speed until the sample material reaches a “densification” state. Figure 3 illustrates the obtained intrinsic compression curves for five samples with various sizes of pores and different relative density values, where the corresponding stresses are computed dividing the recorded experimental loads by the initial specimen area \( \sigma = F / S_0 \) and the deformations from the rapport between the measured height displacements and the initial specimen height \( \epsilon = \Delta h / h_0 \). As an example, the analysis of the engineering stress-strain variation for the specimen D (Fig.4) shows that the compression curve have generally three stages: a first one which corresponds to an elastic behaviour (characterized by an apparent Young modulus \( E^* \)), controlled by the elastic flexion of both the edges and the walls of the cells, a second stage corresponding to the compression plateau, where progressive collapse of the cells can be observed and a third one, when start the contacts between the struts and the material densification mode. In order to estimate the stress at the beginning of the plateau i.e. \( \sigma_p \), this one corresponds to the coordinate of the point obtained by intersection of the tangents of the initial linear part and of the plateau one. The obtained results together with the principal
physical property concerning the various density values for the five metallic foams are presented in Table 1.

![Fig.3](image-url) The curves of the engineering stress-deformation variations obtained during the uniaxial compression tests of the five foams Sn–Pb samples [12].

![Fig.4](image-url) The curve of the engineering stress-strain variation obtained during the uniaxial compression of sample D until beginning of the “densification” process.

Table 1. General physical and mechanical properties of 50%Sn-50%Pb foam specimens (the corresponding massive alloy is characterized at room temperature by a density of \( \rho =0.00886 \text{ g/mm}^3 \) and a yield flow true stress \( \sigma_0 \) around of 41 MPa).

<table>
<thead>
<tr>
<th></th>
<th>Density ( \rho^* ) [g/mm(^3)]</th>
<th>Relative Density ( \rho_r = \rho^*/\rho )</th>
<th>Plateau Stress ( \sigma_p ) [MPa]</th>
<th>Plateau Stress Ratio ( \sigma_p/\sigma_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.002786</td>
<td>0.3144</td>
<td>2.88</td>
<td>0.0702</td>
</tr>
<tr>
<td>B</td>
<td>0.003265</td>
<td>0.3685</td>
<td>3.84</td>
<td>0.0937</td>
</tr>
<tr>
<td>C</td>
<td>0.003711</td>
<td>0.4189</td>
<td>5.00</td>
<td>0.1220</td>
</tr>
<tr>
<td>D</td>
<td>0.002967</td>
<td>0.3349</td>
<td>3.16</td>
<td>0.0771</td>
</tr>
<tr>
<td>E</td>
<td>0.002916</td>
<td>0.3291</td>
<td>3.00</td>
<td>0.0732</td>
</tr>
</tbody>
</table>

3. Analysis of plateau stress ratio vs. relative density variation

For these open-cell metallic foams is important to find the realistic relationships which describe the variation of plateau stress ratio with different values of the relative density. The scientific literature proposes various empirical models.

3.1. Empirical models

Previous research works of Gibson and Ashby [8] concerning the variation of the plateau stress ratio \( \sigma_p/\sigma_0 \) with \( \rho_r \) show that a large class of foams seems to be characterized by a relationship close to a power law:

\[
(\sigma_p/\sigma_0) = C(\rho^*/\rho)^a
\]

defined by a constant \( C \) around of 0.3 and an index \( a \) close to 1.5.
In previous works Féret and other authors [9] have analysed porous compressible materials used especially in civil engineering applications, showing that the constant $C$ is rather close to the unity, while the coefficient $a$ vary between 1.5 and 2.5. Concerning the Sn-Pb foams analyzed in this paper, starting from the Table 1 and the previous power law (1), a non-linear regression analysis leads to the different results presented in Table 2 and Fig. 5.

Table 2. Identification results obtained from non-linear regression analysis of power law

<table>
<thead>
<tr>
<th>Empirical Models</th>
<th>$C$</th>
<th>$a$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashby</td>
<td>0.6858</td>
<td>1.9913</td>
<td>0.9935</td>
</tr>
<tr>
<td>Féret</td>
<td>0.9913</td>
<td>2.3409</td>
<td>0.9991</td>
</tr>
<tr>
<td>Model I</td>
<td>0.9999</td>
<td>2.3663</td>
<td>0.9950</td>
</tr>
</tbody>
</table>

![Fig.5. Non-linear regression curves defining the plateau stress ratio variation vs. relative density: a) Ashby Model, b) Féret Model, c) General Power Model.](image)

In the case of the classical Ashby model only the measured five experimental points are used for the non-linear regression analysis, while for a Féret model has been added the point
corresponding to the dense state of the material \((\rho_r = 1, \rho^* = \rho\) and \(\sigma_p/\sigma_0 = 1\)). It is then possible to obtain a \(C\) coefficient close to 1 and an index parameter \(a\) around of 2.34 placed in the interval \([1.5, 2.5]\). If the origin point is also taken into account (General power model or Model I using then a complete experimental data base), the results are similar with the Féret ones, excepting the index value \(a\) identified around of 2.36. This analysis shows the need to be covered the influence of all the relative density values, between that defined by very low values \((\rho_r \rightarrow 0)\) until that corresponding to a dense solid alloy \((\rho_r = 1)\) i.e. \(\rho_r \in [0,1]\). The Model I seems to be more adequate and in concordance with the model proposed by Féret. It remains to analyze the extrapolation towards values close to unity, where the kinetics of the densification process should be slower.

### 3.2. Proposed physically based models

Outside the above empirical analysis, it is important to add that during an upsetting load, a compressible foam material undergoes a continuous densification process and the relative density gradually increases to 1. Physically it is also necessary that in the vicinity of the unity value the variation of the mechanical properties with the relative density to be asymptotic toward those of the compact state. Inspired from other experimental phenomena analyzed by Avrami which concerns evolution or kinetics of thermo-mechanical, metallurgical or microstructure variables \([10, 11]\), the following relationship can be used:

\[
\xi(x) = \alpha \left\{ 1 - \exp \left[ -k x^n \right] \right\}
\]  

(2)

where \(x\) represents the time or an other basic state variable. If \(x \ll 1\) and \(n > 1\), \(\xi(x) \approx C x^n\) and if \(x \to \infty\) then \(\xi(x) \to \alpha\) (value of the saturation bearing). Or here \(x\) must be represented by the relative density \(\rho_r\) (i.e. \(0 \leq x \leq 1\)) and it is necessary to have \(\xi(0) = 0, \xi(1) = 1\).

Consequently the plateau stress ratio can be described by a simple Avrami form:

\[
\frac{\sigma_p}{\sigma_0} = \frac{1 - \exp \left[ -k (\rho^*/\rho)^n \right]}{1 - \exp(-k)}
\]

(3)

where \(k\) is a material constant and \(n\) is the index defining the “velocity” of the \(\rho_r\) influence.

For a better control of the evolution curve shape, a modified Avrami equation it is proposed in the form:

\[
\frac{\sigma_p}{\sigma_0} = \frac{\left\{ 1 - \exp \left[ -k' (\rho^*/\rho)^{n'} \right] \right\}^{m'/n'}}{\left\{ 1 - \exp(-k') \right\}^{m'/n'}}
\]

(4)

defined by the material constants \(k', n'\) and \(m'\).

### Table 2. Parameters values obtained from non-linear regression analysis

<table>
<thead>
<tr>
<th>Proposed Models</th>
<th>(k)</th>
<th>(k')</th>
<th>(n)</th>
<th>(n')</th>
<th>(m')</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Avrami (II)</td>
<td>4.10(^{-6})</td>
<td>-</td>
<td>2.37</td>
<td>-</td>
<td>-</td>
<td>0.995</td>
</tr>
<tr>
<td>Modified Simple Avrami (II-1)</td>
<td>1.9832</td>
<td>-</td>
<td>3.16</td>
<td>-</td>
<td>-</td>
<td>0.986</td>
</tr>
<tr>
<td>Generalized Avrami (III-a)</td>
<td>-</td>
<td>0.0762</td>
<td>-</td>
<td>25.198</td>
<td>2.37</td>
<td>0.995</td>
</tr>
<tr>
<td>Generalized Avrami (III-b)</td>
<td>-</td>
<td>0.0105</td>
<td>-</td>
<td>0.5868</td>
<td>2.37</td>
<td>0.995</td>
</tr>
<tr>
<td>Modified Generalized Avrami (III-1)</td>
<td>-</td>
<td>1.5749</td>
<td>-</td>
<td>25.558</td>
<td>2.43</td>
<td>0.994</td>
</tr>
</tbody>
</table>
The Table 2 synthesizes the numerical results obtained from non-linear regressions using the complete experimental data base and Fig. 6 pictures the corresponding curves which permit to have comparisons of the plateau stress ratio evolution and of its morphology shape.

![Fig. 6. Non-linear regression curves defining the plateau stress ratio variation vs. relative density: a) Model II, b) Model III.](image)

For the proposed modified Avrami’s models (II-1 and III-1), in order to have a tangent of the asymptotical variation toward the unity for a value of $\rho_r$ equal to 1, it is necessary that $n = \frac{\exp(k)-1}{k}$ and $m' = \frac{\exp(k')-1}{k'}$. It is true that this tangent must to be close to 0, but this condition is very difficult to be obtained and generally expect that $k$ or $k'$ strive toward infinity. A tangent value equal to unity can be a good compromise which permits to overcome the dense material density value with a relatively small error around of 5%-10%.

### 3.3. Discussions

Using the simple Avrami equation (Model II and Eq. (3)) the curve shape is similar with the power one (Model I) fitting correctly the experimental points (Fig. 6a). Requiring that the tangent corresponding to the maximal value of the relative density ($\rho_r = 1$) to be equal to 1, the modified Avrami model (Model II-1) leads to a sigmoid variation curve which permits to predict saturation around the bearing $\sigma_p = \sigma_0$ (stress ratio equal to 1). It is true that in absence of experimental points for $\rho_r \geq 0.5$, the validation of the extrapolation values obtained from these models is based only in the physical considerations due to the material behaviour in the vicinity of the dense state.

If the generalized Avrami equation is used (Model III and Eq. (4)) a best fitting of the experimental data is obtained (Fig. 6b) together with a more realistic prediction for the modified generalized form (Model III-1). Regarding the values of index parameters the solution of $n'$ around 25 (Model III-a and Model III-1) and of $m'$ around of 2.4 seem to be more realistic. Furthermore in this last case the $m'$ parameter replaces the role of the $n$ index describing the simple Avrami equation. It is also possible to mention the similar magnitude of these two parameters with the index used by the power law.
4. **FINITE ELEMENT MODELLING**

A finite element model based on commercial code Abaqus is used to simulate the compression test of cylindrical foam specimens. A homogenization technique using isotropic conditions is employed to describe the material properties and the distribution of all the state variables: density, stress, pressure and either elastic or plastic deformation.

**4.1. Mesh definition, kinematic conditions and material model**

Using an axis symmetric formulation (linear elements of CAX4R type), the geometry and the mesh are pictured in Fig. 7a. The elements have a size of 0.75 mm in the both radial and axial directions. A constant velocity of 0.055 mm/s is applied in the axial direction in order to obtain a compression loading. According to the specimen (D) having the mechanical properties presented in Fig. 3 and Table 1, the experimental compression test gives the load-displacement curve. The true stress and the true plastic strain must be computed by:

\[
\sigma_0^c = \frac{F}{S} \text{ and } \bar{\epsilon} = -\ln(1 - \epsilon) \text{ with } \epsilon = \Delta h / h_0
\]  

During the experiment have been observed that the real area of the specimen is equal to the initial one \(S = S_0\), consequently the material is purely compressible. The material hardening behaviour can be then represented by the true stress-plastic strain curve pictured in Fig. 7b.

\[
\sigma^2 + \sigma^2 \beta^2 \sigma_0^c \quad \text{i.e.} \quad \frac{\sigma^2}{\sqrt{3} \sigma_0^c} + \frac{p^2}{\frac{1}{\sqrt{3}} \sigma_0^c} = 1
\]

where \(\sigma\) is the Von-Mises stress, \(p\) is the hydrostatic pressure and \(\sigma_0^c\) is the true stress obtained from an uniaxial compression deformation.

---

**Fig. 7.** a) Geometry and mesh of the half of cylindrical disc specimen D, b) Isotropic material hardening expressed in terms of the true stress - plastic strain variation

A large class of metallic foams has a Poisson coefficient near to 0. In this case the ratio between the experimental compression stress value and the pressure obtained in a uniform hydrostatic test is equal to \(\sqrt{3}\). If an isotropic hardening crushable foams model is used to describe the material rheological behaviour, the plastic compressible criterion is defined by:

\[
\bar{\sigma}^2 + \alpha^2 p^2 = \beta^2 \sigma_0^c \quad \text{i.e.} \quad \left(\frac{\bar{\sigma}}{\sqrt{3} \sigma_0^c}\right)^2 + \left(\frac{p}{\frac{1}{\sqrt{3}} \sigma_0^c}\right)^2 = 1
\]
The cumulated plastic strain can be defined by
\[ \varepsilon_p = \int_{t_0}^{t} \dot{\varepsilon} d\tau \] where the equivalent plastic strain rate is obtained from the total dissipated plastic power expressed in terms of the stress tensor and of the plastic strain rate one i.e.
\[ \dot{\varepsilon} = \left[ \sigma \right] \left[ \dot{\varepsilon} \right] / \sigma_0^c. \] It is easy to verify that in the case of a simple compression test, \( \bar{\sigma} = \sigma_0^c, \) \( p = -\sigma_0^c / 3, \) \( \dot{\varepsilon}^c = \dot{h} / h, \) consequently
\[ \dot{\varepsilon} = \dot{\varepsilon}^c = \dot{h} / h \text{ and } \dot{\varepsilon} = \ln(h_0 / h), \] results which confirms the relationships (5) and (6). The above material model requires the use of a Dynamic Explicit version and the hardening curve is introduced via the pairs of experimental points. To avoid the problem based on the transition between the compressible state and the incompressible one (characterising the dense state of the material), in this study the extrapolation of the experimental curve after a plastic strain of 70% is made by the use of a rigid plastic part defined by a bearing stress around of 18 MPa which correspond to the observation of the beginning of the “densification” process.

4.2. Numerical Simulation Results

The Fig. 8 pictures the time evolution of the density distribution during the uniaxial specimen compression. It can be see that after 3 mm of the cylindrical disc crushing the density is about twice as large that the initial one (beginning of the densification process), and around of 3.6 mm the material has a density close to the compact material (\( \rho = 0.00886 \text{ g/mm}^3 \)).

Fig. 8. Time evolution of the numerical density distribution during the compression (\( t = 0 \text{ s, } t = 28 \text{ s - } \Delta h = 1.5 \text{ mm, } t = 57 \text{ s - } \Delta h = 3.1 \text{ mm and respectively } t = 65 \text{ s - } \Delta h = 3.6 \text{ mm} \))

After the time step equal to 57 s the simulation uses the extrapolation of the compressible behaviour that it is not in accord with the real evolution of the material which becomes to be incompressible, having a behaviour close to a compact solid state. It is then necessary to develop more adequate models which permit to describe correctly the transition between a metallic foam state and a perfect compact one.
4.3. Discussions

Until a height displacement of 3 mm the predicted numerical load-stroke curve is in a very good agreement with the experimental one (Fig. 9.a). At the same time, in the middle area of the sample, the evolution of the density with the true plastic strain (Fig. 9b), shows that this one seems to be correct, excepting the points which following the beginning of the densification (70.6% of the true plastic strain or 51.6% of the engineering deformation), where the numerical results can be treated only as an indication of the transition toward the compact state.

![Graphs showing experimental and FE values](image)

**Fig. 9.** a) The comparison between the experimental load curve and the computed one, b) The variation of the numerical density value with the true plastic strain.

5. CONCLUSIONS

This study has been made to analyse on the mechanical point of view the behaviour of St-Pb open-cell foams, especially in terms of new relationships proposed to describe the variation of the plateau stress ratio with the relative density. A finite element model based on a crushable compressible behaviour has been used to a better understand of the densification process during a simple upsetting test. However the obtained results confirm the need in the future to develop more rigorous models able to describe during a forming process the transition between a compressible behaviour and an incompressible one.

REFERENCES


34
Numerical modeling of magnetic induction and heating in injection molding tools

Patrick Guerrier, Jesper H. Hattel

Technical University of Denmark, Department of Mechanical Engineering, Section of Manufacturing Engineering, Denmark
email pagu@mek.dtu.dk

ABSTRACT

Injection molding of parts with special requirements or features such as micro- or nanostructures on the surface, a good surface finish, or long and thin features results in the need of a specialized technique to ensure proper filling and acceptable cycle time. The aim of this study is to increase the temperatures as close as possible to the cavity surface, by means of an integrated induction heating system in the injection molding tool, to improve the fluidity of the polymer melt hereby ensuring that the polymer melt will continue to flow until the mold cavity is completely filled. The presented work uses numerical modeling of the induction heating in the mold to investigate how the temperature in the mold will be distributed and how it is affected by different material properties.

KEYWORDS: Induction heating, injection molding, finite element, coupled.

1. INTRODUCTION

In order to injection mold special parts having micro- or nanostructures on the surface, being long and having thin features, or requiring a good surface finish, elevated mold temperatures will help ensuring complete filling of the cavity [1-3]. However, this can also lead to an increased cycle time due to a longer cooling time, if the heating is not only done locally where needed. The aim is therefore to only increase the temperature as close as possible to the surface by means of an induction heating system to improve the fluidity of the polymer melt hereby ensuring that the polymer melt will continue to flow until the mold cavity is completely filled. The temperature required will depend on the transition temperature of the polymer [1].

Recent studies [4-10] have been investigating induction heating of the mold tool, but they are mainly focused on using an induction coil to heat up the surface of the mold before the injection of the polymer. One drawback with this method is that a lot of the heat will already be dissipated out into the mold and away from the surface. This will in turn not be much warmer than without induction, as the whole mold will be heated up instead, again leading to a longer cooling time.
The basic idea behind this new concept is that a coil should be placed beneath the mold cavity surface, encased, on the backside, in a magnetic material, typically ferrite. The mold cavity in front should be made of a non-magnetic material electroplated with a layer of magnetic material e.g. nickel, to have the magnetic field running as close as possible to the surface of the cavity side.

The presented work uses numerical modeling of the induction heating in the mold to investigate the temperature distribution in the mold and especially at its surface. There are two main mechanisms taking place in induction heating, namely electromagnetism and thermal conduction. The electromagnetism is described by Maxwell’s equations which need to be solved. A rapid changing magnetic field induces Eddy currents, which gives rise to a resistive heating effect, known as Joule heating. The thermal conduction is controlled by the heat conduction equation wherein the Joule heating is acting as a source term. The Eddy currents tend to run at the surface of a magnetic conducting material due to the well-known effect of the skin depth [11]. It is desired to find a combination of materials with different electrical properties to get the heating as close as possible to the cavity surface. In this study different material properties are investigated to find such a combination.

2. MECHANISMS

The electromagnetic part of induction heating is controlled by Maxwell’s equations, which here are presented in terms of free charges and currents [12]:

\[ \nabla \cdot \mathbf{D} = \rho_f \]  \hspace{1cm} (1)  

\[ \nabla \cdot \mathbf{B} = 0 \]  \hspace{1cm} (2)  

\[ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \]  \hspace{1cm} (3)  

\[ \nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} \]  \hspace{1cm} (4)

Where \( \mathbf{D} \) is the electric flux density, \( \mathbf{B} \) is the magnetic flux density, \( \mathbf{E} \) is the electric field, \( \mathbf{H} \) is the magnetic field, \( \mathbf{J}_f \) is the free current density, \( \rho_f \) is the free charge density, \( t \) is the time. Relating \( \mathbf{D} \) and \( \mathbf{H} \) in terms of \( \mathbf{E} \) and \( \mathbf{B} \) depends on the material, and for linear media it can be related through the following constitutive relations:

\[ \mathbf{D} = \varepsilon \mathbf{E} = \varepsilon_0 \varepsilon_r \mathbf{E} \]  \hspace{1cm} (5)  

\[ \mathbf{H} = \frac{1}{\mu} \mathbf{B} = \frac{1}{\mu_0 \mu_r} \mathbf{B} \]  \hspace{1cm} (6)

Where \( \varepsilon_0 \) is the vacuum permittivity, \( \varepsilon_r \) is the relative permittivity, \( \mu_0 \) is the vacuum permeability, and \( \mu_r \) is the relative permeability. Furthermore the current density and electric field can be related by the well-known Ohm’s law:

\[ \mathbf{J} = \sigma \mathbf{E} \]  \hspace{1cm} (7)
Where $\sigma$ is the electrical conductivity. The displacement current in equation (4) will be ignored and the assumption that a time-harmonic varying source current density will result in a sinusoidally varying magnetic field will be made. Combining Maxwell’s equations (1)-(4) with the constitutive equations (5)-(7), the following complex diffusion equation can be derived [11]:

$$\frac{1}{\mu} \nabla^2 \vec{A} - i\omega \sigma \vec{A} = -\vec{J}_s$$  \hspace{1cm} (8)

Where $\vec{A}$ is the magnetic vector potential related to the magnetic flux by $\vec{B} = \nabla \times \vec{A}$, $\vec{J}_s$ is the source current density in the coil, $\omega = 2\pi f$ is the angular frequency (and $f$ the frequency), and the overbar is denoting the peak value or the amplitude. Assuming an axisymmetric cylindrical system, the magnetic vector potential has an azimuthal component only and equation (8) reduces to:

$$\frac{1}{\mu} \left( \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \vec{A}_\theta}{\partial r} \right) + \frac{\partial^2 \vec{A}_\theta}{\partial z^2} \right) - i\omega \sigma \vec{A}_\theta = -\vec{J}_{s,\theta}$$  \hspace{1cm} (9)

The magnetic flux can be found from the solution of the magnetic vector potential by taking the curl of $\vec{A}_\theta$, hence for the axisymmetric case it can be calculated as:

$$\vec{B}_r = -\frac{\partial \vec{A}_\theta}{\partial z}; \quad \vec{B}_z = \frac{\partial \vec{A}_\theta}{\partial r} + \frac{\vec{A}_\theta}{r}$$  \hspace{1cm} (10)

The magnetic field can then be found using equation (6). The induced Eddy currents in the conductors can be found from:

$$\vec{J}_e = -i\omega \sigma \vec{A}_\theta$$  \hspace{1cm} (11)

From which the Joule heating can be found via:

$$\dot{Q} = \frac{I}{2\delta} |\vec{J}_e|^2$$  \hspace{1cm} (12)

Which is the volumetric heat source induced by the Eddy currents. A well-known and useful expression is the current skin depth or penetration depth, which can also be derived from Maxwell’s equations. This quantity is defined as the distance for which the amplitude of a plane wave decreases by a factor of $e^{-1} = 0.368$, and will prove useful later:

$$\delta = \frac{I}{\sqrt{\pi \mu \sigma}}$$  \hspace{1cm} (13)

The other mechanism taking place in induction heating is the heat conduction, which can be described by the transient heat conduction equation:
\[ \rho c_p \frac{\partial T}{\partial t} = k \nabla^2 T + \dot{Q} \]  \tag{14}

Where \( T \) is the temperature, \( \rho \) is the density, \( c_p \) is the specific heat capacity, \( k \) is the thermal conductivity, and \( \dot{Q} \) is the heat source from equation (12).

### 3. FINITE ELEMENT FORMULATION

The finite element method has been used to solve equation (9) and (14). This implies first finding the heat source term from the electromagnetic solution and then using it in the transient heat conduction equation to get the resulting temperature distribution. The implementation of the finite element formulation is done in-house and is self-developed using MATLAB. The implementation is tested against the commercial software COMSOL and QuickField, to validate the code. For the actual experimental setup presented here, it was chosen to use the self-developed implementation due to its code flexibility and fast solution time.

Using the Galerkin method to approximate the solution on a set of discrete points we get

\[
\int_{\Omega} [N]^T \left[ \frac{1}{\mu} \left( \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial A_0}{\partial r} \right) \right) + \frac{\partial^2 A_0}{\partial z^2} \right] - i \omega \alpha A_0 + J_{s,0} \right] \, d\Omega = 0 \]  \tag{15}

where \([N]\) is a row vector containing the shape functions, and \( \Omega \) is the domain of interest. The total matrix system becomes:

\[
[[K] + i[C]] [A_0] = \{f\} \]  \tag{16}

with the following element matrix equations:

\[
[K]_e = \int_{\Omega} [B]^T [B] \, d\Omega; \quad [C]_e = \int_{\Omega} \omega [N]^T [N] \, d\Omega; \quad \{f\}_e = \int_{\Omega} [J_s] [N]^T \, d\Omega \]  \tag{17}

where \([B]\) is the derivative of the shape function also called gradient matrix. Linear triangular elements have been employed and a closed form solution of the integration has been done with the resulting elemental matrices being similar to those described in [11].

The heat conduction equation is discretized in a similar way. Employing equation (14) in cylindrical coordinates and using the Galerkin method we obtain

\[
\int_{\Omega} [N]^T \left[ k \left( \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial T}{\partial r} \right) \right) + \frac{\partial^2 T}{\partial z^2} + \dot{Q} - \rho c_p \frac{\partial T}{\partial t} \right] \, d\Omega = 0 \]  \tag{18}

Inserting the spatial approximation of the temperature will then result in the following form of matrix equation:
\[
[C] \left( \frac{\partial T}{\partial t} \right) + [K][T] = \{f\}
\] (19)

With similar element matrices to those in equation (17) but substituting \( \frac{1}{\mu} \) with \( k \), \( \omega \sigma \) with \( \rho c_p \), and \( \bar{J} \) with \( \dot{Q} \). The time derivative is discretized using an implicit Euler approximation (backward finite difference) to yield the following form:

\[
([C] + \Delta t[K])[T]^{n+1} = [C][T]^n + \Delta t\{f\}
\] (20)

Where \( \Delta t \) is the time step. The implicit formulation has the advantage of being unconditionally stable. Both equation systems seen in equation (16) and (20) are built as sparse matrices and both are solved using MATLAB’s build-in solver, which can solve unsymmetrical sparse linear systems on the form \([A]x = b\).

\section*{4. GEOMETRICAL SETUP}

An induction heating setup has been employed with the self-developed finite element code and in COMSOL. Only the results from the self-developed code will be shown as they are identical to the COMSOL implementation. This is done to investigate the temperature distribution in the mold utilizing a unique combination of materials to get the Eddy currents as close as possible to the mold cavity. All components have been integrated into the mold itself, in order to avoid external inductors to heat the mold surface. Also, the unit can be used with a conventional injection molding machine, with the only extra equipment being the power supply.

The mold seen in Fig. 1 consists of a mold plate, which from the backside contains a core with a coil inserted. The coil consists of windings of insulated copper wires. The core (or field concentrator) is used to carry the magnetic flux and to concentrate it where it is needed, here in front of it towards the surface of the mold plate.

![Fig.1. Left: Exploded view of the main parts in the mold with the build-in inductor. Right: Cross section of the mold with the cavity side up, where the molten polymer will flow.](image_url)
5. NUMERICAL MODEL SETUP

In the finite element model it is assumed that the geometry is axisymmetric, and consists of a cross section as seen in both Fig. 1 to the right and in Fig. 2. Besides the mold plate, core, and coil other features in the geometry includes an added layer of a magnetic material on the top of the mold plate. The magnetic layer’s purpose is to carry the magnetic field as close as possible to the cavity surface. The coil is encased in a non-conducting material like a polymer, and inside each winding is a cooling channel, which prevents the coil from being heated up due to Joule heating.

![Fig.2. Detailed drawing of the sections in the axisymmetric numerical model](image)

The boundaries in the models are of Dirichlet type at the axis of symmetry for the electromagnetic field, as the field vanishes toward the symmetry axis. The same type of boundary is also used for the three other outer boundaries, as the field also vanishes towards those, due to the magnetic material carrying the magnetic flux.

The current density is applied in the coils, and it is assumed that each winding carries 200 A at a frequency of 10 kHz. This gives a current density of approximately $7.6 \times 10^7 \text{ A/m}^2$.

The boundaries for the heat conduction equation are adiabatic (neutral Neumann condition) at all the outer boundaries. For all material interfaces the heat transfer coefficient is set to infinity. The cooling channel domains are set to a Dirichlet type condition with a specified temperature of the cooling channel liquid. This is chosen to be the initial temperature of the mold which is room temperature at 25°C. The time step used in the simulation was 0.1 seconds. The electrical and thermal material data used for the finite element calculation have been summarized in Table 1.

A two-step coupling between the electromagnetic and thermal solution is used. This has been chosen because the material data are assumed to be constant, and not depend on temperature. This approximation can give good results at low temperatures (<200°C) [11] and is very efficient to employ, as the complex steady state diffusion equation (9) only needs to be solve once to find the heat source term, which can then be switched on and off as desired in the transient thermal analysis.
Due to the skin depth the current will run in a thin layer in the magnetic materials. This gives a rather strict requirement for the mesh in the magnetic layer in order to get sufficiently accurate results. A good practice [11] is to have at least three skin depths of fine mesh. At three skin depths the current is reduced to $e^{-3} = 0.05$ of its original value, meaning that 95% of the currents will run in the fine mesh. For the mold plate the skin depth can be calculated to be $\delta = 4.18$ mm and for the magnetic layer to be $\delta = 0.1336$ mm, so a maximum element size of 0.315 mm should satisfy the requirement in the mold and 0.0550 mm in the nickel layer at the boundary facing the mold plate should satisfy the requirement in the magnetic layer. The mesh, see Fig. 3, has been created in COMSOL and subsequently imported into the in-house developed FE-code.

6. RESULTS AND DISCUSSION

The solution time of the MATLAB implementation is around 11 seconds, compared to the COMSOL validation simulation (linear elements and 1st order accuracy in time) which takes around 15 seconds. From the solution of equation (9) the magnetic flux is found, and the flux lines are depicted in Fig. 4. It can be seen that the magnetic field is concentrated in front of the core, and the flux lines move relatively easy through the mold material and into the magnetic layer which carries the flux in a very thin layer due to the skin depth.
Fig. 4. Contours of equal magnetic vector potential magnitudes or the magnetic flux lines induced by the coil.

Figure 5 shows the Joule heat arising from the Eddy currents, which are found directly from the electromagnetic solution using equation (11) and (12). In the figure to the left it is possible to see the distribution in the z-direction. A high peak is observed in the layer, but also a lot of heat is generated in the mold plate itself.

The reason for the substantial amount of heat generated in the mold, even though it is a non-magnetic material, can be related back to the skin depth, equation (13), as both the electrical conductivity and the permeability contribute to how deep the currents run in a material. This means that lowering the electrical conductivity further in the mold material will allow the magnetic flux to easier penetrate the mold, and be concentrated in the magnetic layer close to the cavity. A ten times lower electrical conductivity in the mold will give a skin depth of $\delta = 13.22$ mm compared to 4.18 mm, which is an indication that the magnetic flux can more easily penetrate the mold and into the magnetic layer.

Different variations of the electrical conductivity have been tested for the mold material and for the magnetic layer. In Fig. 6 the Joule heating in the z-direction can be seen using different combinations of electrical conductivity. The red curve to the left is the same curve as in Fig. 5 to indicate the difference in magnitude.

Fig. 5. Joule heat distribution in different positions: Left: in the z-direction above the inductor (here, a high peak is observed in the magnetic layer). Right: in the radial direction at different distances from the mold cavity surface.
Fig. 6. Distribution of the Joule heat in the z-direction above the inductor for different electrical conductivities, $\sigma$, in the mold (denoted with m) and the layer (denoted with l). Downward pointing arrow means a ten times lower conductivity, and upward pointing arrow means a ten times higher. Having a low electrical conductivity in the mold is favorable.

A clear improvement can be seen for a (here ten times) lower electrical conductivity in the mold material. Raising the electrical conductivity of the magnetic layer does give a higher Joule heating in the layer. It still generates a lot of heat in the mold plate which is not favorable, and gives an even stricter requirement for the mesh. What is the most interesting is how lowering the electrical conductivity of the mold plate will promote the Joule heating in the magnetic layer, resulting in an effect which is greatly desirable for this setup. The temperature distribution in the mold can been seen in Fig. 7 after four seconds of heating for both standard and low value of the electrical conductivity of the mold plate.

Fig. 7. Left: (Unfavorable) Temperature distribution after four seconds of induction heating using the electrical conductivities specified in Table 1. Right: (Favorable) Temperature distribution after four seconds of induction heating using a ten times lower electrical conductivity in the mold as compared to Table 1.

Even though the maximum temperature after four seconds is higher for the simulation with standard electrical conductivity in the mold plate, it is located close to the inductor and not at the surface where it is desired. This means a higher amount of energy would have to be removed after the injection phase, making it more energy inefficient. The transient behavior of the temperature at two measurement points $P_1$ (surface) and $P_2$ (center: Axis-symmetric line at the surface) shown in Fig. 2 and the maximum temperature at any given point are shown in Fig. 8 for four seconds of induction heating and four seconds of cooling.
Fig. 8. Comparison of the temperature at the surface above the inductor ($P_1$), center of the mold ($P_2$), and the maximum temperature, during four seconds of heating and four seconds of cooling, without (left) (unfavorable) and with (right) (favorable) a ten time reduction in the electrical conductivity of the mold.

It is evident that the temperature at the surface above the inductor is very close to the maximum temperature during the heating phase, and still cools down together with the center point, when having a low electrical conductivity in the mold. It should be noted that the simulations were done with the mold being at room temperature initially, so the cooling phase of both simulations in Fig. 8 are quite similar, but it is expected that having a lower electrical conductivity in the mold will reduce the cooling phase when a steady cycle is reached.

7. CONCLUSION

A finite element model of an induction heating system to be used in an injection mold has been presented. The main mechanisms and their associated equations, namely the complex steady state diffusion equation and the transient heat conduction equation, have been presented. The finite element discretization using the Galerkin method is also presented. The model is subsequently used in a study of an injection molding tool with a built-in inductor, where the Joule heating and temperature distribution were investigated. An interesting finding was, that using a magnetic layer on the cavity surface alone is not enough to promote a maximum temperature field in the surface. Due to the skin depth, a non-magnetic material with a low electrical conductivity will be needed in order to promote the heating of the tool as close as possible to the cavity surface.

REFERENCES


New Concepts for Offline Dimensional Control in Sheet Metal Forming

Viorel Paunoiu, Alexandru Epureanu

Dunarea de Jos University of Galati,
Department of Manufacturing Engineering,
viorel.paunoiu@ugal.ro

ABSTRACT
Springback is one of the factors that affects the final shapes of a sheet metal part. In this paper, we present a novel method for springback-based offline dimensional control which is based on an integrated-optimal approach. The key idea is to build a reduced-order model of the entire blank-process-tooling system and, by using it, to find the optimal values of the system parameters for which the part accuracy and precision are the highest. The reduced order model is built on the base of Finite Element simulation results and Neural Networks modeling while the optimum parameters values, for which the part accuracy and precision become maximum, are finding using appropriate search technique. A case study, namely the sheet multipoint forming with interpolators, come to validate the new proposed method.

KEYWORDS: sheet metal forming, precision, accuracy, optimization, multipoint forming

1. INTRODUCTION
Springback is one of the factors that affects the final shapes of a sheet metal part. The springback is defined as the geometric difference between the loaded and unloaded configurations of the thin sheet deformed part. The springback percentage in entire dimensional error is much higher than the percentage of other sources of errors. Reduction of the error component due to springback means drastically reduction of the entire dimensional error. In other words, the springback control could be equivalent with the control of entire dimensional error. So we can accept the term springback-based dimensional control. On the other hand, the sheet metal forming is often a high-speed process, done with high frequency mechanical presses. Thus, the real time control of springback is extremely difficult to achieve. Therefore the offline control of springback, i.e. the control in preparatory stages, such as tooling development and press work-cycle programming, appears to be a more realistic solution.

By combining the two mentioned aspects of the springback we can define and use the term offline springback-based dimensional control.
The offline springback-based dimensional control, in classical approach, means to acting in two directions: for the springback compensation or for the springback prediction.

In terms of springback compensation, trial-and-error, displacement adjustment and spring-forward methods are used. The trial-and-error methods are based on engineering experience. Displacement adjustment is a geometrical method and consists in moving the surface nodes defining the die surface in the direction opposite to the springback error. The spring-forward method is based on the stress tensor that cause springback and computing the constraint forces to maintain equilibrium following forming [1, 2].

In terms of springback prediction, some analytical or numerical methods could be used. The analytical methods are difficult to apply due to the lack of understanding of the stresses distribution throughout the sheet. The methods are limited to simple geometries and simple deformations such as bending and flanging. The numerical methods are based on the Finite Element (FE) analyses of the sheet metal forming processes. The springback prediction using FE is also a sensitive process, which is not only influenced by springback computation itself, but depends on the accuracy of previous forming simulation process [2, 9].

The current methods for offline springback-based dimensional control, have the following main conceptual shortcomings: they refer to the part dimensional accuracy only; the part dimensional precision is always neglected; for solving the problem of springback, even many solutions are used, neither comparative analysis nor optimal search were applied; at present the increased potential of the integrated approaches for dimensional control remains unused.

Starting from these we are proposing an overall approach for the springback-based offline dimensional control, namely integrated-optimal approach and a new control method, namely springback optimization on the base of an integrated-optimal approach.

2. CHARACTERIZATION OF THE SPRINGBACK OPTIMIZATION METHOD

Fig. 1 presents a hypothetical example. In this example, there is only one controlled variable, i.e. the L dimension. Reference is presented in Fig. 1.a, where \( L_{\text{accepted}}^{\min} = 50.1 \text{mm} \) and \( L_{\text{accepte}}^{\max} = 50.7 \text{mm} \) are the limits and \( \Delta L_{\text{accepte}} = 0.6 \text{mm} \) is the range of tolerance domain of L. The value \( L_{\text{middle}} = 50.4 \text{mm} \) corresponds to the middle of this domain and is considered as desired value of dimension.

In the following, we will characterize the springback by the parameter \( SB \), defined as springback-caused error and computed as difference between the real value \( L_{\text{result}} \) of dimension L, which is affected by springback phenomenon, and its desired value \( L_{\text{middle}} \). In addition, to define the reference, let us consider the case in which the batch of products are processed without any offline control, the target dimension, \( L_{\text{target}} \), is equal with \( L_{\text{middle}} \), and the values of all other variables are in the middle of their domains of variation. Also, let us consider that the real value of the dimension L is \( L_{\text{result}} = 50.9 \text{mm} \). In this case the value of springback parameter is \( SB_{w.c.} = 0.5 \text{mm} \). As a consequence, at different samples of the batch, the perturbed variables fluctuate, then the springback varies as well, with \( \Delta SB_{w.c.} = 0.8 \text{mm} \), between \( L_{\text{result}}^{\min} = 50.5 \text{mm} \) and \( L_{\text{result}}^{\max} = 51.3 \text{mm} \). If the tolerance limits of L are exceeded, as is the case shown in Fig. 1.b, then the springback control is required.

In the springback optimization method, Fig. 1.c, the optimal offline control should be addressed as below:
- First step is to set the geometry of contact surfaces between part and tooling, by giving to each geometric parameter of these surfaces the value that corresponds to the situation in which the values of parameters that describe the blank, process, and part are in the middle of their domain of variation.

- In the second step, two groups of manipulated variables are selected from the variables that refer to the blank, tooling, process, and part. The selected variables should be independent of each other. The two groups have different targets. One is dedicated to minimize the springback parameters and another to minimize the variations of these parameters during the batch manufacturing. In example from Fig. 1.c, where the springback parameter is unique, the first group of variables is dedicated for minimizing the parameter \(SB\) value (its minimum value is denoted \(SB_{opt}\)), while the second, for minimizing the variation \(\Delta SB\) of this parameter (its minimum variation is denoted \(\Delta SB_{opt}\)).

- In the third step is checked if the domain of \(\Delta SB_{opt}\) is located within the domain of \(\Delta L_{accepted}\). If this condition is not met, then springback optimization approach is not effective. In example from Fig. 1.c, this approach is successful because \(SB_{opt}=0\), while \(\Delta SB_{opt}=0.4\)mm is smaller than \(\Delta L_{accepted}=0.6\)mm.

### 3. THE VARIABLES IN THE NEW PROPOSED APPROACH

The blank-process-tooling system, in our approach, is characterized by two notions. First is the notion of system variable, which has a mathematical meaning and refers to the models that describe the system behaviour. The second is the notion of system parameter, which has a physical meaning and refers to the system design, programming, and operation.

The system variables are selecting from the system parameters, but not only, and are defining according to the structure of their mathematical model.

In terms of variables values changing, we define four categories, namely fixed, perturbed, manipulated, and controlled variables. Fixed variables do not change during both process set-up and material processing stages. Perturbed variables change randomly in a neighbourhood of their nominal values, due to the inherent perturbations, in a perturbation domain. Manipulated variables are changing in a controlled manner by the designer for the
different outputs offline control, in a manipulation domain. Controlled variables are controlled for belonging to the tolerance domain.

The system parameters are selected from the system features or states and are defined according to the structure of their physical model.

In terms of parameters values defining, for every parameter of the blank-process-tooling system, five values are defined, namely the real, desired, target, programmed, and nominal value. Real value results after the system operation is completed or is estimating by physical measurement. Desired value corresponds to the middle of the tolerance domain. Target value is strategically imposed for the output parameters. Programmed value is selected from the input parameters to obtain the target values. Nominal value is only a reference, used for defining the tolerance domain.

The desired and real values refer to the output parameters only. The difference between them is called error. Dimension accuracy is estimated as the absolute value of those dimension errors, which appears when the perturbed and manipulated variables take their programmed values. Moreover, the overall part accuracy is defined as

\[ \sqrt{\text{Accuracy}_1^2 + \text{Accuracy}_2^2 + \ldots} \]

where Accuracy1, Accuracy2, ... are the accuracy values for all tolerated dimensions. Dimension precision is estimated as the difference between the maximum and minimum values of the dimension error, when the manipulated variables are fixed to their optimal values, while the perturbed variables explore their domains of variation.

4. THE INTEGRATED-OPTIMAL APPROACH

The proposed integrated-optimal approach has three main major modules: simulation, reduced-order modeling, and optimization (Fig. 2).

4.1. Simulation module

Firstly, the structure of the system full dimensional model is established. The start point consists in building the physical model of blank-process-tooling system along with their input/output parameters. Based on this, a full dimensional mathematical model of the system is built, using the FE technique. The result of the simulation is a basic dataset, describing the system behaviour in a number of cases. More details about this stage are presented [9].

4.2. Reduced-order modeling module

On the base of previous obtained dataset, in this module we will identify as well as we will reduced-order model the causal relations between each basic output variable, and the perturbed and manipulated variables. The reduced-order modeling addresses the problem of obtaining an approximate model that is computationally tractable and capture the essential physics, but neglect details that are not critical for the problem at hand. For this, the best neural network model technique is applied. This technique consists in testing, for each basic output variable, all possible neural network structures and in selecting the ones which best fit the basic dataset resulted from simulation. Each possible structure is defined by the input variables and also by the internal structural parameters of the network, like the number of the hidden layers and the number of neurons on each layer. The NN Model software incorporates this technique and it is used in this paper, to obtain the reduced-order model of blank-process-tooling system.
4.3. Optimization module

In this module, the reduced-order model will be used for assessment of the control approach. We will define the output synthetic variables and the performance indicators.

As output synthetic variables we can define the sensitivity of each part tolerated dimension to perturbed variables modification, numerically defined as the ratio between dimension change and the correlated change of perturbed variable, along with the modifications of those part dimensions which are not tolerated and were used as manipulated variables, as well as the variation domain of the part tolerated dimensions, which was generated by the perturbed variables changing.

As performance indicators, we are choosing the dimension accuracy and precision, obtained by a search technique. More details about this stage are presented [5, 10].

The structure of the optimization module for a specific case study is presented below.

5. THE SPRINGBACK OPTIMIZATION METHOD - A CASE STUDY

The multipoint forming is based on the concept of discrete approximation of a die continuous surface. It consists of a number of closely spaced multiple rigid surface tool elements, known as pins, each of which is a surface element of an expected contour. The heights of the pins can be adjusted to approximate the desired surface shapes either manually or using a computer control [3, 4, 7, 8].

![Diagram of the optimization process](image.png)
To apply for the proposed approach a simulation model was developed using Dynaform finite element software. The model was constructed considering the obtaining a single curvature plate. Fig. 3 shows the simulation model, in different deformations stages. The model includes two arrays of rigid active elements and two interpolators (upper and down rubber) between these and the blank. No blank holder is used so the rubber blocks are free to expand. The figure shows the complexity of the deformation process which is developing during the forming considering both the deformation of the rubber interpolator and the blank. This was the reason for which we choose this process which implicates a lot of parameters who are not controllable and could influence the springback values of the part.

5.1. SIMULATION STAGE

The target values dimensions are: length $B_i = 120$ mm, depth $H_i = 21.345$ mm, (radius $R_i = 95$ mm). The blank length is 129.883 mm and width 130 mm.

The perturbed variables are: Thickness of blank material, $t$, varies between 0.8 and 1.2 mm; Material characteristics, $K$, varies between 616 and 680 MPa; Friction coefficient, $\mu$, varies between 0.1 and 0.15.

The manipulated variables are: Pins stroke or punch stroke, $s$, will vary between 17 and 29 mm, according to interpolator thickness; Rubber thickness, $t_R$, varies between 2 and 10 mm; Rubber elastic modulus, $E_R$, varies between 14 and 44 MPa.

The fixed variables are: The part is a sheet with a single curvature with a width of $B_{tp} = 120$ mm, a maximum depth $H_{tp} = 21.345$ mm (radius of 95 mm); Material hardening exponent $n$ is 0.22. The $R$ anisotropy coefficients values are set to: $R_{00} = 1.87$; $R_{45} = 1.27$; $R_{90} = 2.17$; Punch speed is 100 mm/second; Tool is modeled as rigid surfaces. The tool is characterized by two arrays of pins, with 100 pins for each array. The pins are disposed face to face, both on $x$ and $y$-direction; Rubber interpolator is a material type Elvax 460. The interpolator is modeled as Mooney Rivlin Rubber material. The fixed input material data are: density, $\rho = 0.946$ g/cm$^3$; Poisson ratio, $\nu = 0.499$. Solid elements are used for the rubber interpolator mesh.

Controlled variables will be defined below.

5.2. REDUCED-ORDER MODELING

The reduced order model is based on neural network modeling. The training is made using the data obtained after FEM simulation. The input data, variables V1-V8 from table 1, are the values of the manipulated and perturbed variables. The outputs data, variables V9-V10, are the values of dimensions $B$ and $H$ after the springback simulation.

The best neural network model technique was applied for the causal relation identification and modeling. Using this technique resulted that only the perturbed/manipulated
variables showed in Tables 2 and 5 are in causal relation with \( B \) and \( H \). In this way, for each of these variables, it was structured and trained a neural network. These two neural networks represent the reduced-order model of the blank-process-tooling system which will be used, according to the considered springback-based dimensional control approach.

Table 1. Training data for neural network modelling

<table>
<thead>
<tr>
<th>INPUT</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t ), [mm]</td>
<td>( \mu ), [MPa]</td>
</tr>
<tr>
<td>( tR ), [mm]</td>
<td>( s ), [mm]</td>
</tr>
<tr>
<td>( ER ), [MPa]</td>
<td>( Bi ), [mm]</td>
</tr>
<tr>
<td>( Hi ), [mm]</td>
<td>( B ), [mm]</td>
</tr>
<tr>
<td>( H ), [mm]</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.125</td>
</tr>
<tr>
<td>…</td>
<td>0.125</td>
</tr>
<tr>
<td>…</td>
<td>1.2</td>
</tr>
<tr>
<td>228</td>
<td>1</td>
</tr>
</tbody>
</table>

5.3. THE SPRINGBACK OPTIMIZATION METHOD

Fig. 4 presents the method flowchart. The start point is the reduced-order model of the blank-process-tooling system, while the end is the optimal values of system parameters for which the accuracy and precision of the part tolerated dimensions have their best values. The method steps are the following: precision improving, accuracy improving, and precision evaluation.

5.3.1. Precision improving

The perturbed variables, \( t \), \( \mu \) and \( K \) are fixed to their nominal values (1; 0.125; 648 respectively). For precision improving, are considered the variations of the manipulated variables. The variation domains of manipulated variables are discretized, table 2 -partial data.

To each triplet of the divided manipulated variables, a perturbed variable matrix is allocating. An interrogation matrix was built, according to Fig. 4, step 1. The system reduced-order model represented by the trained network will be interrogating using this matrix. The result of interrogation will be the dimensions values of \( B \) and \( H \), as the basic output variables.

The springback sensitivities values are given by:

\[
S_{AB_g} = \frac{\Delta B( g_i ) - \Delta B( g_{i-1} )}{g_i - g_{i-1}} \Delta g ; \quad S_{AH_g} = \frac{\Delta H( g_i ) - \Delta H( g_{i-1} )}{g_i - g_{i-1}} \Delta g
\]

\[
S_{AB_\mu} = \frac{\Delta B( \mu_i ) - \Delta B( \mu_{i-1} )}{\mu_i - \mu_{i-1}} \Delta \mu ; \quad S_{AH_\mu} = \frac{\Delta H( \mu_i ) - \Delta H( \mu_{i-1} )}{\mu_i - \mu_{i-1}} \Delta \mu
\]

\[
S_{AB_K} = \frac{\Delta B( K_i ) - \Delta B( K_{i-1} )}{K_i - K_{i-1}} \Delta K ; \quad S_{AH_K} = \frac{\Delta H( K_i ) - \Delta H( K_{i-1} )}{K_i - K_{i-1}} \Delta K
\]

The \( S_{AB} \) and \( S_{AH} \) global sensitivities are given by:
To obtain the minimum springback errors due to $S_{AB}$ and $S_{AH}$, we will introduce the sensitivity error index:

$$TS = \sqrt{S_{AB}^2 + S_{AH}^2}$$

(6)
Table 2. Values of manipulated variables used in interrogation matrix, partial data – step 1

<table>
<thead>
<tr>
<th>Manipulated variables - Interpolator and process parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubber thickness, $t_R$, [mm]</td>
</tr>
<tr>
<td>--------------------------------</td>
</tr>
<tr>
<td>1 10</td>
</tr>
<tr>
<td>2 10</td>
</tr>
<tr>
<td>… …</td>
</tr>
<tr>
<td>10 8.5</td>
</tr>
<tr>
<td>11 8.5</td>
</tr>
</tbody>
</table>

Table 3. Accuracy level in springback optimization method, partial data – step 1

<table>
<thead>
<tr>
<th>Synthetic variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SA_Bg$</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>1 0.046</td>
</tr>
<tr>
<td>2 0.023</td>
</tr>
<tr>
<td>… …</td>
</tr>
<tr>
<td>10 0.389</td>
</tr>
<tr>
<td>11 0.218</td>
</tr>
</tbody>
</table>

The smallest values of the sensitivities are found using exhaustive search technique. According to $TS$ index, using the exhaustive search, table 3 presents the top smallest obtained interrogation values, in ascending order. By comparison, we are considering that the best combination of the manipulated variables is composing from the triplet (8.5; 26; 44).

5.3.2. Accuracy improving

In this step, we will model the connection between the variations of the tool geometry parameters and the springback values, considering the fixed values of interpolator and process parameters and perturbed variables. The variation domains of tool dimensions will be discretized. The perturbed variables, $t$, $\mu$ and $K$ are fixed to their nominal values (1; 0.125; 648) and interpolator and process parameters variables are fixed to the values (8.5; 26; 44). The tool dimensions, $B_i$ and $H_i$, are presented in table 4. An interrogation matrix is building, Fig. 4, step 2. The system reduced-order model represented by the trained network will be interrogating using this matrix. The result of interrogation will be the dimensions values of $B_2$ and $H_2$, as the basic output variables. The dimension errors, as output variables, are given by:

$$EpsB = B_2 - Btp; \ EpsH = H_2 - Htp$$  \hspace{1cm} (7)

where: $B_2$ and $H_2$ are the dimensions after springback; $Btp$ and $Htp$ – desired part dimensions.

Overall part error index is given by:

$$TP = \sqrt{EpsB^2 + EpsH^2}$$  \hspace{1cm} (8)

The optimal values of the manipulated variables are finding using appropriate search technique. According to $TP$ index, using the exhaustive search, table 4 presents the smallest resulted interrogation values, in ascending order.
Table 4. Accuracy level in springback optimization method, partial data – step 2

<table>
<thead>
<tr>
<th>Manipulated variables - Tool dimensions</th>
<th>Basic output variables</th>
<th>Synthetic variables: dimensions error and overall part error for the top 12 best sets of manipulated variables values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Bi$</td>
<td>$Hi$</td>
<td>$B2$</td>
</tr>
<tr>
<td>1</td>
<td>115.2</td>
<td>25.655</td>
</tr>
<tr>
<td>2</td>
<td>115.25</td>
<td>25.615</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10</td>
<td>118.8</td>
<td>22.580</td>
</tr>
<tr>
<td>11</td>
<td>118.3</td>
<td>22.993</td>
</tr>
</tbody>
</table>

It results that the optimal values of the manipulated tool geometry along with the accuracy improved level are the combination (115.2; 25.655) mm.

5.3.3. Precision evaluation

In this stage the divided search space refers to the perturbed variables and theirs domains of perturbation. For this, the system reduced-order model represented by the trained network will be interrogated with the above triplet of optimal values. The tooling geometry values will remain constant at their target part dimensions (115.2; 25.655) and the perturbed variables domains will be discretizing. An interrogation matrix is building, Fig. 4, step 3.

Table 5 presents the first input variables values, used in network interrogation.

Table 5. Values of perturbed variables for precision estimation, partial data – step 3

<table>
<thead>
<tr>
<th>Perturbed variables</th>
<th>Material thickness, $g$, [mm]</th>
<th>Friction coefficient, $\mu$</th>
<th>Material characteristic $K$, [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.1</td>
<td>647</td>
</tr>
<tr>
<td>2</td>
<td>0.9</td>
<td>0.1</td>
<td>647.5</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10</td>
<td>0.9</td>
<td>0.125</td>
<td>647</td>
</tr>
<tr>
<td>11</td>
<td>0.9</td>
<td>0.125</td>
<td>647.5</td>
</tr>
</tbody>
</table>

Table 6. Precision level in springback optimization method, partial data – step 3

<table>
<thead>
<tr>
<th>Output variables</th>
<th>Dimensions deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B3$</td>
<td>$H3$</td>
</tr>
<tr>
<td>1</td>
<td>120.535</td>
</tr>
<tr>
<td>2</td>
<td>120.541</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10</td>
<td>120.590</td>
</tr>
<tr>
<td>11</td>
<td>120.593</td>
</tr>
</tbody>
</table>

The result of the interrogation will be the dimensions values of $B3$ and $H3$, as the basic output variables (Table 6). The dimensions deviations, $\Delta B$ and $\Delta H$, as the output variables, are obtained:

$$\Delta B = B3 - Btp; \quad \Delta H = H3 - Htp$$ (9)
where: $B3$ and $H3$ are the dimensions after springback; $B_{tp}$ and $H_{tp}$ – dimensions desired values.

In the two columns of the synthetic variables $\Delta B$ and $\Delta H$ from table 6, we will search for the maximum values respective the minimum values of them. The dimensions precision, $\delta B$ and $\delta H$, as output variables, are given by:

$$
\delta B = \Delta B_{\text{max}} - \Delta B_{\text{min}}; \quad \delta H = \Delta H_{\text{max}} - \Delta H_{\text{min}}
$$

(10)

where: $\Delta B_{\text{max}}$, $\Delta B_{\text{min}}$ and $\Delta H_{\text{max}}$, $\Delta H_{\text{min}}$ – extreme dimensions deviation.

The dimension precisions, $\delta B$ and $\delta H$, are presented in table 7.

Table 7. Dimension precisions, $\delta B$ and $\delta H$ in springback optimization method

<table>
<thead>
<tr>
<th>Deviation limits - dimension $B$:</th>
<th>Deviation limits - dimension $H$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta B_{\text{max}} = 0.66459$</td>
<td>$\Delta H_{\text{max}} = 0.65727$</td>
</tr>
<tr>
<td>$\Delta B_{\text{min}} = -0.10166$</td>
<td>$\Delta H_{\text{min}} = -0.18929$</td>
</tr>
</tbody>
</table>

Precision - dimension $B$: $\delta B = 0.76625 \text{ mm}$

Precision - dimension $H$: $\delta H = 0.84656 \text{ mm}$

6. COMPARATIVE RESULTS

The values of performance indicators, resulted by springback optimization method implementing are presented in table 8. Table 8 presents, also, by comparison, the values of the same indicators for the other two methods of springback control, namely optimized springback reduction and optimized springback compensation, developed by the authors and are based on the proposed integrated-optimal approach.

Table 8. Comparative results obtained by applying the three methods based on the integrated-optimal approach

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$EpsB$</td>
<td>$EpsH$</td>
</tr>
<tr>
<td>Optimized springback reduction, [10]</td>
<td>0.778</td>
<td>0.081</td>
</tr>
<tr>
<td>Optimized springback compensation, [6]</td>
<td>0.378</td>
<td>0.409</td>
</tr>
<tr>
<td>Springback optimization</td>
<td>0.231</td>
<td>0.297</td>
</tr>
</tbody>
</table>

Fig. 5. The springback optimization method
Figure 5 is a representation of the data from table 8, for the springback optimization method, $B$ dimension. The maximum error is $\varepsilon_{B_{\text{max}}}^{\text{opt}} = 0.664$. By comparison with optimized springback reduction and optimized springback compensation methods, see table 8, using the springback optimization method, the obtained accuracy and precision increase was up to 80% and 54%, respectively. It results the best choice which is the springback optimization.

7. CONCLUSIONS

In the paper, we propose a method for offline dimensional control in sheet metal forming, based on integrated-optimal approach and applied to sheet multipoint forming with interpolators. Here, the integrated-optimal approach developed by the authors and their main stages, namely simulation, reduced-order modeling, and optimization, were revised only.

The integrated-optimal approach includes a new control method, namely springback optimization, which is here presented. A reduced order model of the blank-process-tooling system at whole is built and, using it, were found the optimal values of the system parameters for which the part accuracy and precision are the highest. The reduced order model is built on the base of FE simulation results. For optimization, the part accuracy and precision are evaluated, using the reduced-order model and considering many system parameters values, which belong to their tolerated domains of variation. The result is a significantly increasing of the accuracy and precision.

The application of the integrated-optimal approach shows the importance of dimension precision and the fact that this is an important output parameter which, till now, was no considered. The effect of random variation of the perturbed variables is major which demonstrate that the proposed approach is useful.

The springback optimization method leads to a better part accuracy and precision so it can be used when the precision imposed level is higher.

REFERENCES


Numerical optimization of die geometry in open die forging

Peter Christiansen¹, Jesper H. Hattel¹, Niels Bay¹, Paulo A.F. Martins²

¹Department of Mechanical Engineering, Technical University of Denmark, DTU – Building 425, DK-2800 Kgs. Lyngby, Denmark,
²Instituto Superior Tecnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
petc@mek.dtu.dk

ABSTRACT

This paper deals with numerical optimization of open die forging of large metallic ingots made by casting implying risk of defects, e.g. central pores. Different material hardening properties and die geometries are combined in order to investigate, which geometry gives rise to maximum closure of a centreline hole in a single compression operation. Friction is also taken into account.

The numerical analysis indicates that a lower die angle of approximately 140° results in the largest centreline hole closure for a wide range of material hardening. The value of optimum die angle is not influenced by friction, which was found only to change the degree of centreline porosity closure in case of small lower die angle.

KEYWORDS: Ingot forging, Centreline porosity closure, Numerical simulation, Die design

1. INTRODUCTION

The open die forging process is applied for production of large metallic parts such as shafts for ship’s propellers, power plant turbines or wind turbines. It is important to ensure the mechanical soundness of the shafts, since failure of such components is expensive and unacceptable.

Manufacturing large shafts consists of a number of operations: First a large steel block (known as an ingot) is cast. After solidification and cooling, the cast ingot is reheated to approximately 1200°C and subsequently hot forged. The forged ingot is then left to cool down after which it is machined by turning to its final shape. The machined workpiece is then hardened and grinded.

The open die forging process is schematically illustrated in Fig. 1. The ingot is placed in-between a flat upper die and a V-shaped lower die, after which the upper die is moved downwards to compress the ingot. After compression, the ingot is lifted from the lower die and rotated around its centre axis (usually 45°) and the compression operation is repeated.
Repetition is carried out a number of times to finish a cross-section of the ingot, after which the ingot is moved forward and a new cross-section is forged.

According to ASM Handbook [1] the main reason for performing the open die hot forging is to minimize defects originating from the casting process. Such defects are for instance segregations, inclusions, porosities due to gas entrapment or poor feeding, or coarse grain structure due to long cooling time. The centreline of the ingot is where these defects are most pronounced. This is because the centreline of the ingot is the last place to solidify. It is possible to diminish and sometimes even remove these defects by hot forging: recrystallization may be triggered by the plastic deformation, inclusions can be crushed to smaller particles and porosities may be closed during forging. The latter is the scope of the present investigation.

Pioneering investigation of the open die forging process was conducted by Nasmyth [1]. Based on practical experience he found that a V-shaped lower die was superior to a flat lower die regarding closure of centreline porosities and to ensure soundness of the final product.

A theoretical analysis of the open die forging process using V-shaped and flat dies was presented by Johnson [2], who applied slipline analysis for die compression of circular cylindrical ingots. Despite the need to assume ideal-plastic materials and plane strain deformation, slipline analysis improved the understanding of why utilization of two plane dies resulted in hydrostatic tension, hence opening of porosities, at the centre of the workpiece during forging.

Upper bound analysis of compression of cylindrical billets with square cross section containing porosities was presented in Keife & Ståhlberg [3] and corresponding model wax experiments in Ståhlberg et al. [4]. Although providing a better understanding of why porosities close during plastic deformation, the analysis is based on plane strain assumption and rigid-plastic materials. The porosities are furthermore equally distributed across the specimen analysed, thereby only approximately resembling the porosity distribution in a cast ingot, where porosity density increases towards the centre of the ingot.

Dudra & Im [5] carried out plane strain numerical simulation of open die forging of model ingots with a centreline hole to mimic a porosity using a pair of plane and a pair of V-shaped 135° dies. They confirmed the aforementioned work of Nasmyth and Johnson that V-shaped dies are superior to plane dies regarding closure of centreline porosities.

A numerical study using the Gurson-Tvergaard-Needleman porous plasticity model was presented in Christiansen et al. [6]. A real size ingot with a porous centre region was compressed in a number of forging steps using a flat and a 90° V-shaped lower die. It was
found that the flat lower die resulted in hydrostatic tension at the centre of the ingot while the V-shaped die was able to suppress the hydrostatic tension hereby preventing porosity increase at the centre of the ingot. However only these two lower die angles were investigated, hence there could be other die angles more suited for centreline porosity closure.

An experimental study by Christiansen et al. [7], where downscaled lead model ingots with drilled centreline holes to mimic centreline porosity were compressed using different lower die angles, indicated that different degrees of centreline porosity closure were achieved while applying the same length of press stroke. The study indicated an optimum lower die angle, where centreline porosity closure would be maximized. The study also showed good agreement between the numerical simulation and the physical modelling.

The scope of the present paper is to estimate, which lower die angle may be most suited for centreline porosity closure. The paper furthermore aims to investigate, whether this optimum lower die angle is affected by variations in strain or strain rate hardening. The change in mechanical behaviour, i.e. the stress response, may be due to ingots of different materials, thermally induced changes in hardening properties due heating to different initial temperature or cooling during deformation.

2. NUMERICAL MODELLING

2.1. Finite element flow formulation

Since the open die forging of large ingots is normally performed by means of the relatively slow moving punch of a hydraulic press, accelerations are assumed to be insignificant and dynamic effects may be neglected. This allows the numerical modelling of the process to be carried out with the quasi-static finite element flow formulation, which is based on the following weak variational formulation:

$$\int_V \bar{\sigma} \delta \dot{\varepsilon}_{pl}^{\text{v}} dV + K \int_V \dot{\varepsilon}_{pl}^{\text{v}} \delta \varepsilon_{pl}^{\text{v}} dV = \int_S \tau_i \delta u_i dS$$

(1)

where $V$ is the control volume limited by surfaces $S_u$ and $S_t$, where velocities $u_i$ and tractions $\tau_i$ are prescribed, $\bar{\sigma}$ is the effective stress, $\dot{\varepsilon}_{pl}^{\text{v}}$ is the effective plastic strain rate, $K$ is a large constant penalizing volume shrinkage and hereby enforcing plastic incompressibility and $\dot{\varepsilon}_{pl}^{\text{v}} = \dot{\varepsilon}_{pl}^{\text{v}}$ is the volumetric plastic strain rate. Friction is modelled using the constant friction model $\tau = m_f k$ where $0 \leq m_f \leq 1$ is the friction factor and $k$ is the shear flow stress. Both full and reduced integration schemes are utilized when performing the numerical integration. Details on computer implementation of the flow formulation may be found in Nielsen et al. [8].

2.2. Simulation layout

The ingot investigated is considered to be 2000mm in diameter and having a centreline hole 100mm in diameter (5% of total diameter) in order to mimic a porous centre region. The ingot is placed between a flat upper die and a V-shaped lower die and compressed once 200mm, i.e. 10% of the ingot diameter. Lower die angles ranging from $60^\circ$ to $180^\circ$ (flat lower die) are employed. In order to reduce computational time only 2D models taking advantage of centreline symmetry are utilized. Discretization is performed by means of approximately 1700 quadrilateral elements. Plane stress loading conditions are assumed. This is a reasonable
assumption if the ratio of die width \( w \) to ingot diameter \( d \) is reasonable small. In an industrial case of open die forging, an ingot with a diameter of approximately 2000mm was forged using a die with a width of 1000mm giving a ratio \( w/d = \frac{1}{2} \), hence plane stress is a relevant assumption for such a forging case. An example of mesh and simulation layout can be seen in Fig. 2.

![Fig. 2. Ingot before (a) and after (b) being forged by a 120° lower die.](image)

Two different friction scenarios are investigated: frictionless and constant friction with a friction factor \( m_f = 0.5 \).

The simulation is performed using 200 increments (corresponding to approximately 0.05% reduction in diameter per step) and the convergence criterion is \( \Delta u/u < 0.01 \), where \( \Delta u \) is the change in the Euclidian norm of the velocities from previous to current iteration and \( u \) is the Euclidian norm of the velocities from previous iterations.

### 2.3. Ingot material

Since the open die forging process is performed with steel ingots preheated to approximately 1200°C, the flow stress of the steel is lowered considerably as compared to room temperature and may be taken to increase not only with increasing strain but also with increasing strain rate. A combination of Hollomon [9] and Norton [10] hardening is used to describe the hardening behaviour:

\[
\sigma_o = C \left( \frac{\dot{\varepsilon}^{pl}}{\varepsilon^{pl}} \right)^n \left( \frac{\dot{\varepsilon}}{\varepsilon} \right)^m
\] (2)

where \( \sigma_o \) is the flow stress of the material, \( C \) is the strength coefficient, \( n \) is the strain-hardening exponent and \( m \) is the strain-rate sensitivity exponent. As listed in Table 1 a number of different combinations are investigated to determine how varying hardening may affect the optimum lower die angle.
Table 1. Different combinations of material parameters utilized in the investigation.

<table>
<thead>
<tr>
<th>C [MPa]</th>
<th>n</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>100</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>100</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>100</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

The exact value of C is not important in the present study because it only affects the flow stress level and not the variation in flow stress within the ingot in contradiction to the values of n and m, which are of interest, since they affect the material flow. An approximate value of C = 100 MPa is taken from Spittel & Spittel [11].

3. RESULTS AND DISCUSSION

3.1. Plastic flow under different material behaviour

Fig. 3 shows results obtained in simulating forging of an ideal rigid-perfectly plastic material ingot using different lower die angles. The effective plastic strain is used as output to indicate the plastic flow.

Fig. 3. Effective plastic strain when forging an ingot with constant flow stress, $\sigma_0 = 100\,\text{MPa}$ using lower die angles of $60^\circ$ (a), of $100^\circ$ (b), of $140^\circ$ (c) and $180^\circ$ (d).

A clear variation in deformation pattern with varying lower die angle can be seen in Fig. 3, $140^\circ$ resulting in the largest plastic straining around the centreline hole.
An example of forging a strain-hardening material can be seen in Fig. 4. Comparing Fig. 3 with Fig. 4 it is noticed that strain hardening influences the flow pattern. Whereas non-hardening material results in deformation confined to small zones of the ingot, the strain hardening distributes the deformation to a larger part. This lowers the maximum strain (compare values corresponding to the same lower die angles).

In case of strain-rate hardening material Fig. 5 similarly shows more widespread deformation than predicted by the ideal-plastic material. The plastic strain is, however, not so evenly distributed as for a strain-hardening material with $n = 0.4$.

$$\sigma_0 = 100MPa \left( \varepsilon^{pl} \right)^{0.4}$$

using lower die angles of $60^\circ$ (a), of $100^\circ$ (b), of $140^\circ$ (c) and $180^\circ$ (d).
Fig. 5. Effective plastic strain when forging an ingot having a flow stress \( \sigma_0 = 100\text{MPa} \left( \varepsilon_{pl} \right)^{0.4} \) using lower die angles of 60° (a), of 100° (b), of 140° (c) and 180° (d).

### 3.2. Optimum lower die angles vs. material behaviour and friction

In order to determine the optimum lower die angle as a function of material hardening behaviour and friction quantitatively, curves displaying the area ratio between initial and final centreline hole cross-section area \( A_{\text{ratio}} = \frac{A_{\text{final}}}{A_{\text{initial}}} \) are used to express the amount of centreline porosity closure. The resulting finite element predictions are plotted in Fig. 6 - Fig. 8. It is worth noticing the influence of the inserted friction factor \( m_f \) in the overall modelling.

![Fig. 6. \( A_{\text{ratio}} \) as function of strain hardening and friction.](image-url)
As seen in Fig. 6, the largest degree of closure for the tested strain hardening range is obtained when utilizing a lower die angle of approximately 140°. Friction seems to increase the closure for small lower die angles, whereas the influence is insignificant when applying larger lower die angles.

![Graph showing area ratio as function of strain-rate hardening and friction.]

**Fig. 7.** $A_{ratio}$ as function of strain-rate hardening and friction.

From Fig. 7 it can also be concluded that an optimum lower die angle regarding centreline hole closure is approximately 130°-140° for the entire range of simulated strain-rate hardening. Again friction seems to increase the closure, when applying small lower die angles, whereas it has insignificantly influence in case of larger lower die angles.

![Graph showing area ratio as function of combined strain and strain-rate hardening and friction.]

**Fig. 8.** $A_{ratio}$ as function of combined strain and strain-rate hardening and friction.

In case of material with combined strain and strain-rate hardening as seen in Fig. 8, maximum closure occurs, once again, when applying a lower die angle of approximately 140°. Friction only has a significant influence in case of small lower die angles.
3.3. **Forging load**

Previous sections lead to the conclusion that material behaviour of the ingot plays a marked influence on the level of centreline hole closure. In general it was found that an increase in either strain or strain-rate hardening gave rise to an increase in centreline hole closure. If material strain hardening is not present due to elevated workpiece temperature a rise in forging speed, which increases the strain-rate, may be feasible to promote closure of the centreline hole. The disadvantage of increasing strain hardening or strain-rate hardening is a larger forging load.

It is also noticed that friction increases the centreline hole closure when applying smaller lower die angle, while it does not significantly influence the closure in case of larger die angles. This is due to increased forging load, which is especially sensitive to friction in case of small die angles. If the lower die is flat (180°) the load is not increased with increasing friction, see Fig. 9.

![Fig. 9. Comparison of forging loads for different lower die angles and friction factors (m).](image)

4. **CONCLUSION**

A numerical investigation of the open die forging process has been conducted with special emphasis on the influence of the mechanical material behaviour on closure of centreline defects e.g. pores. These defects, which originate from the casting process, are modelled by means of a centreline hole. Different strain and strain-rate hardening have been investigated and results indicate that optimum performance regarding centreline hole closure is achieved with a lower die angle of approximately 140°. Friction is only influencing the amount of centreline hole closure, not the angle at which the optimum performance occurs. Strain or strain-rate hardening gives rise to a larger degree of closure than an ideal-plastic material. Since plane stress deformation has been assumed, the study is in essence limited to open die forging processes of large specimens compared to the width of the dies. This will often be the case forging of large ingots.
ACKNOWLEDGEMENTS

The work is supported by the Strategic Research Center “REWIND - Knowledge based engineering for improved reliability of critical wind turbine components”, Danish Research Council for Strategic Research, grant no. 10-093966

Paulo Martins would like to acknowledge the financial support provided by the Velux Foundation during his sabbatical license at the Danish Technical University.

REFERENCES

Design of Experiment for Roll-to-roll Hot embossing Process

Dongwon Yun, Youngsu Son, Heechang Park, Byungin Kim
Sangyong Ham and Seyoung Kim

Korea Institute of Machinery & Materials (KIMM),
dwyun@kimm.re.kr

ABSTRACT

In this paper, experiments for the roll-to-roll (R2R) hot embossing system are performed and analyzed. The developed roll-to-roll hot embossing system is for manufacturing fine patterns on the flexible film. The generated embossed patterns have different depth and width for each polymer and various process conditions. The pattern quality is varied according to the temperature, pressure, film speed and many other conditions. So, it is not easy to find the proper or optimized process condition for hot embossing process.

To find desired hot emboss process conditions, the design of experiment is conducted which helps users to decide which condition should be applied to the process for good products. Especially, Response Surface Method (RSM) is utilized to find the optimum process condition and it can be concluded that the optimum process point can be easily identified in the hot embossing process.

KEYWORDS: Hot embossing, Design of experiment, Roll-to-roll

1. INTRODUCTION

Recently, as electronic and display parts become smaller and thinner, requirement for small and fine pattern technology increases. The demand for cost saving, high performance and high precision also increased. The manufacturing technology for semi-conductor, display device, bio device and optical device needs higher functionality and more degree of integration. To meet these needs, micro or sub-micron pattern should be made, but conventional process like LIGA or MEMS has weakness in productivity and that has limits in broadening the surface. To increase the productivity of micro pattern generation, hot embossing process becomes popular in the field of micro patterning field recently. A hot embossing lithography was proposed for replication of nanostructures[1] and high aspect ratio structures for polymer microstructures was fabricated using hot embossing process[2]. A microfeatured fluidic platform was made by hot embossing technology[3].

This process usually uses pressure and heat to generate many fine patterns on polymer film. But, the plate type hot embossing produces a bottle neck in increasing the production films because the surface tension increases as the surface gets broader. So, the maximum film size is about 400mm [4]. So, plate type hot embossing has some limits to large substrate. The
continuous fabrication system for $\mu$m scale electronic device on flexible substrate should be
developed. So the roll-type hot embossing system has been studied for large substrate
production. Parameters that effect micro channel quality have been investigated [5]. And an
induction heating roll was applied for even temperature distribution and better pattern quality
[6].

Roll-to-roll (R2R) hot embossing uses rotating rolls to fabricate fine patterns on flexible
polymer films, but plate type hot embossing process uses a flat plate as a stamp. The rolls
enable the roll-to-roll hot embossing process make continuous production. The film should be
heated before the mold on the roll surface carves the film. After the patterns are made, the film
should be cooled. So, the pressure and heat are the most importance process variable in the
roll-to-roll hot embossing process.

In this paper, a roll-to-roll hot embossing system will be introduced made by our
research team. After that, several roll-to-roll hot embossing experiments are conducted and
optimum process condition will be searched using the Design of Experiment (DOE).
Especially, Response Surface Method (RSM) will be adopted to find optimum process.

2. ROLL-TO-ROLL HOTEMLBOSSSING SYSTEM

A roll-to-roll hot embossing system was developed. Fig. 1 shows the photograph of the
developed system. This system consists of a main heater roller, a back-up roller, pressurizing
cylinders, a winder and an unwider. An induction heating roll was designed and fabricated and
this roll was used as a main heating roller, so there is no external heater. The patterns are
carved on the main heating roller and the hydraulic pressurizing cylinder pushes the back-up
roller to apply pressure on the film between two rollers.

![Fig. 1. Developed R2R hot embossing system](image)

Line shaped patterns are made on the roller surface and the patterns are carved
horizontally and vertically as shown in Fig.2. Three kinds of line width are used as 20$\mu$m, 30
$\mu$m and 40 $\mu$m.
When the polymer substrate goes between the pattern roll and the backup roll, a hydraulic cylinder above the pattern roll gives pressure onto pattern roll. By pressurizing and heating the substrate, micro sized pattern on the mold can be transferred to the substrate.

(a) Vertical line.  (b) Horizontal line.  
Fig. 2. Generated patterns.

3. ROLL-TO-ROLL HOT EMBOSsing EXPERIMENT

In this section, roll-to-roll hot embossing experiments will be conducted using the developed system. The film material is polyethylene terephthalate (PET) and its thickness is 250 µm. Among the patterns, 20 µm width mould is used for study. To measure the depth of the generated patterns, a confocal microscope (Lasertec Optelics C130) was used. As stated above, the process variables like roll temperature, film speed and pressure are important in the roll-to-roll hot embossing process. But, it is not easy to find the optimum process condition because the combination of variables is complex and the number of cases is large. So, we used the design of experiments (DOE) to find the optimum process variables. During the DOE methods, we adopted the Response Surface Method to search for optimum value. A commercial code, Minitab, was used for RSM analysis. To do this, we set the factors for RSM like Table 1. Roll temperature, film speed and roll speed are used as factors because they are important in deciding the pattern quality.

Table 1. Factor for RSM

<table>
<thead>
<tr>
<th>Factor</th>
<th>Range</th>
<th>unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roll temperature</td>
<td>Low</td>
<td>Mid</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>90</td>
</tr>
<tr>
<td>Force</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>Roll speed</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Through the RSM, 20 experimental runs were designed and experiments were performed for each process condition. The design table and measured depth for each case are displayed at Table 2. Through regression analysis like Fig. 3, we deleted T(temperature)× S(Roll speed) terms and find the response surface like
where $T$, $P$, and $S$ are roll temperature, press force and roll speed respectively.

To view the influence of each factor on the patterning depth, the response surfaces are plotted in Fig. 4. As shown in this figure, we can know the influence of combination of each two variables. Fig. 4(a) shows that the depth gets deeper as force becomes smaller and roll speed gets slower and Fig. 4(b) show that higher temperature and slower speed are helpful for better pattern quality. It can be also known that lower force and higher temperature are desirable through Fig. 4(c).

### Table 2. Design table.

<table>
<thead>
<tr>
<th>Run</th>
<th>point</th>
<th>block</th>
<th>Temp [°C]</th>
<th>Press [ton]</th>
<th>Roll speed [m/m]</th>
<th>Height [µm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>140.454</td>
<td>3</td>
<td>1</td>
<td>0.45</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>120</td>
<td>3.5</td>
<td>1.5</td>
<td>0.237</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1</td>
<td>0.092</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1.84</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>120</td>
<td>2.5</td>
<td>0.5</td>
<td>0.58</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>120</td>
<td>3.5</td>
<td>0.5</td>
<td>0.12</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1</td>
<td>0.16</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>3.5</td>
<td>1.5</td>
<td>0.094</td>
</tr>
<tr>
<td>10</td>
<td>-1</td>
<td>1</td>
<td>90</td>
<td>3.84</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>2.5</td>
<td>0.5</td>
<td>0.09</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1</td>
<td>0.041</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>2.5</td>
<td>1.5</td>
<td>0.086</td>
</tr>
<tr>
<td>14</td>
<td>-1</td>
<td>1</td>
<td>90</td>
<td>2.1591</td>
<td>1</td>
<td>0.124</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1</td>
<td>0.099</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1</td>
<td>60</td>
<td>3.5</td>
<td>0.5</td>
<td>0.133</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>1</td>
<td>120</td>
<td>2.5</td>
<td>1.5</td>
<td>0.452</td>
</tr>
<tr>
<td>18</td>
<td>-1</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>0.1591</td>
<td>0.288</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>1</td>
<td>90</td>
<td>3</td>
<td>1</td>
<td>0.148</td>
</tr>
<tr>
<td>20</td>
<td>-1</td>
<td>1</td>
<td>39.546</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Using Eq. (1), optimization was conducted. Through optimization, we find the optimum process condition like

- Temperature: 140.5 °C, roll speed: 0.16 m/s, force: 2.16 ton

This condition can be expected from the results of Fig. 4. That is, the temperature should be higher and roll speed and force should be minimized for this process. The optimization results of Minitab told that the maximum depth at the obtained optimum condition is 0.407 µm. To check the optimized results, some experiments at the optimum condition were performed.
Experiments have been conducted three times at the same condition. Table 3 shows the measured depth of patterns that is generated at the optimum condition. As shown in this table, the measured value is similar to the analytical result.

Fig. 3. Residual plots

(a) Force versus Roll speed.  
(b) Roll speed versus temperature.

(c) Temperature versus Force.

Fig. 4. Contour plot.
Table 3. Experimental results at the optimum condition.

<table>
<thead>
<tr>
<th></th>
<th>Exp 1</th>
<th>Exp 2</th>
<th>Exp 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth</td>
<td>0.378 um</td>
<td>0.357 um</td>
<td>0.420 um</td>
</tr>
</tbody>
</table>

4. CONCLUSION

In this research, we have introduced the roll-to-roll hot embossing system that was made by our research team. This system used an induction heater installed inside the pattern roller. It is important to find optimum process condition for roll-to-roll hot embossing system as there are many process variables. To find the optimum condition, the RSM was used. Using DOE method, we could find the meaningful experimental condition and we have conducted experiments following the obtained conditions. After completing the experiments, the response surface could be obtained by using the measured depth. Using the response surface, we have also performed optimization and we could find the optimum conditions consequently. Finally, we have conducted experiments at the given condition and identified the performance of our system. Future study will include the system modification to increase the patterning depth as the current performance is not satisfactory.

REFERENCES

Feeding and Positioning of Linked Parts in Micro Production Chains

Philipp Wilhelmi, Eric Moumi, Eike Foremny, Bernd Kuhfuss, Christian Schenck

University of Bremen, bime - Bremen Institute for Mechanical Engineering
wilhelmi@bime.de

ABSTRACT

For micro parts production the introduction of linked parts is a method to simplify handling operations. In subsequent operation steps the effort to provide the position and the orientation of the part can notably be reduced, which leads to higher productivity.

Nevertheless a fast and accurate measuring system is required to reference the individual part to the linked part and the feeding system. In the presented study a part of a processing chain is investigated. It consists of a conveyor, part detection system and a forming machine. The process requirements in the micro range are generally assumed to be very high, but they change depending on the individual process. So in the presented work beneath the capabilities of developed positioning and referencing system are investigated. Also the requirements given by cold forming process of micro rotary swaging are addressed.

KEYWORDS: micro parts production, micro cold forming, linked parts

1. INTRODUCTION

The presented research emerged in context of a project for the production of micro components in linked parts. Linked parts are a solution to overcome micro-production specific problems [1], like dominating adhesion forces and the challenge to handle the generally fragile parts. The basic idea is to rescale the parts to macro dimensions by keeping them linked during the whole production process. Like that a better handling and a continuous feeding is enabled [2]. In the ideal case the base material is used as the linkage and the individual parts are not separated during production. Depending on the base material and specific process requirements different types of linked parts are applied. Especially the two types shown in Fig. 1 are investigated; the ladder type a) and the line type b). The ladder type has the advantage that the parts are not in the flow of feed forces. It is preferably provided for foil material. The line type is used in the considered process chain. Here it consists of spherical parts which are interconnected by a wire. They are produced in a process, where a laser is focused on the wire
and when the material begins to melt both ends of the wire are moved towards each other. The dominating surface tension forms the melted metal to spheres [3].

![Diagram of linked parts]

These material accumulations are preforms for subsequent bulk forming operations. In general the tolerances of micro parts are assumed to be very small. The resulting requirements for the machining processes must be investigated individually. For feeding of the linked parts a conveyor is developed and adapted to a rotary swaging machine, Fig. 2. The system combines two feeding units, so that the linked part can be tensed and material flow caused by the forming process can be drawn to avoid a sagging and a de-positioning of the linked part.

![Diagram of feeding system]

In a process chain with subsequent processing steps the individual parts in the linkage must be positioned separately in each operation unit. So the single parts must be referenced separately to the feeding system. Further the small tolerances of these parts lead to the demand of a precise positioning in subsequent processing steps. This can be realized by mechanical references, if the frame structure can bear the contact forces, which is not ensured for micro dimensions. An alternative is the application of additional sensors for the part detection. Non-contact sensors and especially optical scanning sensors are preferably applied. The considered production process rotary swaging is a cold forming process. The suitability of the feeding concept for this process is analyzed under different aspects like the positioning accuracy and the maximum transmittable friction forces between the linked part and the feeding unit to take the axial processing forces.
2. FEEDING UNITS

The feeding units are designed to feed both types of the considered linked parts, Fig. 3. They work with two pairs of powered belts (1). The belts with an elastic coating are pressed against each other and the linked parts are fed through the gap in between (2). The force transmission is done by friction and can be adjusted (3) by the gap respectively the contact force between the belts and the friction coefficient. Another possibility is to transmit the feeding force by form fit. This can be realized by elastic deformation of the coating when the part are rigid enough and the elasticity is significantly higher in orthogonal than in tangential direction.

In the actual state the feeding units use the rotary encoders of the servo drives (4) for positioning. The relatively low angular resolution of these encoders leads to a theoretical linear position resolution of 100 µm. Nevertheless positioning tests have shown that the repeatability is better than 20 µm. By the integration of external high resolution encoders a linear resolution of the system better than 1 µm can be achieved.

The maximum transmittable force values are measured to be about 12 N for wire with 0.9 mm diameter and 18 N for a linked part consisting of wire of 0.36 mm diameter and spherical material accumulations of 0.9 mm diameter.

3. SENSOR SYSTEM

The production in linked parts is particularly favourable when subsequent production steps are applied to a part. For each positioning the individual part must be referenced separately to the feed system. There are multiple devices which can be used to measure the micro parts. For the fragile parts exclusively optical sensors are investigated.

Three different sensor systems are tested to realize a continuous measuring of the linked parts. All sensors follow the concept of light section. The sensor is fixed in position while the linked part is moved through the measuring field. By merging the single sections with the position information from the feeding unit a 2-D image of the linked part is generated, Fig. 4. These data are analysed by a part detection algorithm and the part position is determined.
An optical micrometer and a laser profile scanner were tested in [4,5,6]. The optical micrometer measures the diameter of the linked parts, Fig. 5. Hence it is appropriate for the rotational symmetric line type. At low feed rates of about 0.6 m/min a detection accuracy of 1±5 µm is achieved.

The laser profile scanner has the advantage that it generates 2½-D data, Fig. 6. This simplifies the part detection, but at metallic surface measuring errors caused by reflections can appear. During the tests it turned out that the reachable sample rate for the part detection is limited by the required exposure times, which again depends on the power of the light source and the sensitivity of the sensor. This system is able to detect the position of the parts at feed rates of about 1.4 m/min with a systematic error of 3 µm.

Due to the insufficient achievable feed rates of the two systems a third sensor system based on a high speed camera is developed, Fig. 7. The system is based on a line camera with 1024 pixels, a sample rate of 48.8 kHz and a powerful led light source. As the principle of light section is applied the line camera simplifies the data processing.

The detection algorithm requires a pre-processing of the line images. First a foreground-background filter is applied. Each pixel below a definite threshold is set to zero while the rest remains untouched. Like that only the important data need to be processed by the detection algorithm. To extract the information about the position of the micro parts different detection algorithms are developed. To be able to determine the centre of objects which create the same data over multiple measurements it is necessary to consider multiple
measurements at once. To keep the computing effort as small as possible a double-stage detection system is developed, Fig. 8. Each image line is first converted into a characteristic value. These values are not directly used to determine the position of the object but stored over a certain window. This window is used in the second stage to generate the pattern value which is used to detect the object position. In this layout only 2048 values need to be calculated in comparison to $1024^2$ for a two dimensional layout with a quadratic window.

To generate the characteristic values out of the raw data, Fig. 9e, four different algorithms are tested. The Fast Fourier Transformation converts an image line to frequency amplitude pairs, Fig. 9a. These pairs are compared to a predefined pattern and the similarity is calculated. Edge detection creates characteristic values by determining the diameter of the object inside the image line, Fig. 9b. Comparison method calculates the difference between the image line from the camera and a predefined image line, Fig. 9c. The fourth method shown in Fig. 9d uses the value and position of each pixel inside the image line to generate the characteristic value. These calculations are based on HU-Moments.

\[
HU = \sum_y (y - \bar{y})^2 g(y)
\]  

with:

- \(y\) = location of pixel in the image line
- \(\bar{y}\) = center of mass in the image line
- \(g(y)\) = value of pixel at position \(y\)

\[\sum (\overline{y}) \]

\[\sum \]

Fig. 9. Part detection algorithms: a) Fast Fourier Transformation, b) Edge detection, c) Comparison method, d) HU-moments, e) Raw data - scanned linked part

The four algorithms are compared regarding resulting detection quality, number of calculations and explicitness of the generated values. The detection quality represents the difference in the calculated values between the part and the line structure as well as the stability of the values. The explicitness is a measure for the number of different objects that can be identified. In experiments the HU-Moments proved to generate the best characteristic values. Because of the experiences from the characteristic value generation only the comparison method and HU-Moments are considered to calculate the pattern value. They are rated by the number of calculations and position detection quality. It is shown that the comparison as well as the HU-Moments determines the position of the micro parts equally well. But the comparison uses less computational power than the HU-Moments.
\[ \text{Comp} = \sum_x pp(x) - pv(x) \] \hspace{1cm} (2)

with:

\[ pv(x) = \text{value of the pattern value at position } x \]
\[ pp(x) = \text{value of the predefined pattern at position } x \]

To realize a real-time capable system a FPGA board (Xilinx, Spartan 3-an) is used. Due to the limitations in computational power only 2773 sequential operations are available to calculate each image line at the maximum sample rate of 48.8 kHz. For the calculation of the characteristic value over 7000 and for pattern value over 6000 operations are necessary. Through parallelization of the calculations it is possible to process over 64000 image lines per second. The architecture of the characteristic value calculation is illustrated in Fig. 10.

Fig. 10. Architecture of the characteristic value calculation.

The testing of the described system of line camera and FPGA has shown that it is possible to determine the position of parts smaller than 1 mm with an accuracy of ±3 µm at feed rates up to 4.3 m/min or of about ±20 µm at feed rates of 59 m/min. The development of a closed loop feed system is planned via a prepared CAN-interface card which can be implemented into the existing detection system.

4. PROCESSING - ROTARY SWAGING

Based on spherical preforms generated by laser melting many other types of geometries can be produced within the line type linked part. Here the micro rotary swaging is considered. Rotary swaging is an incremental cold forming process especially for rods and pipes. According to the German standard DIN8583 the rotary swaging belongs to the open die forging processes. The deformation of the work piece takes place in the swaging head in small steps by a radial oscillating movement of the tools, Fig. 11.
According to the part geometry different variations of rotary swaging can be applied. Fig. 12 shows schematically the infeed swaging and Fig. 13 the plunge swaging process. In common the radial movement of the forming tools (die segments) (2) is generated by the rotation of the driven mandrel (5) and the use of base jaws with cam on the top (3). By means of the roll motion between the base jaws and the cylinder rollers (6), the forming tools are pushed inwards at the same time and reduce with every overrunning of the cams the cross section of the work piece (1), which is axially fed or positioned in the swaging head. The rotation of the main shaft in relation to the work piece leads to a uniform forming over the circumference. The difference of the two process variations is essentially the kind of feed motion. In the infeed swaging process the work piece (1) is fed continuously in the swaging device and the reduction of the work piece by the reducing area of the forming tool (2) takes place over the whole feed length. The wire diameter is reduced from the initial diameter $d_i$ to the final diameter $d_f$ In plunge swaging the work piece is positioned in the swaging head and stands still during the forming process. The oscillating movement of the forming tools is overlaid with a radial feed movement due to the axial movement of the wedges (4), so that a local reduction of the work piece occurs according to the tool geometry. This can be a local diameter reduction or like in the considered case a forming of a preform $k_i$ to an end geometry $k_f$.

Rotary swaging is well implemented in the macro range especially in the automotive industry for example for the manufacturing of axles, steering shafts and so on from tubular blanks with the main goals of weight reduction and optimization of properties e.g. strain hardening. However for manufacturing micro parts it is not well investigated and it is known from other studies that transfer of knowledge from the macro to the micro range is not always
possible [7], for example due to size effects [8]. Since 2007 investigations are running for a better understanding of the micro rotary swaging [9,10,11].

For the forming of the linked parts plunge rotary swaging is considered. Besides of the use of tools with the targeted geometry the volume of the ball in the linked parts and their positioning during the forming are significant parameters for the part quality. To investigate the influence of the positioning on the processing results an experimental study is performed.

Therefore a high precision linear axis with position accuracy in the range of ± 3 µm is used as feed drive, Fig. 14. As sample wire segments with the described laser melted sphere on the tip, Fig. 15a, are used. The initial balls on the wire tip are generated with a bigger diameter than necessary to fill the tool cavities. The experiment evaluates the effect of placing the ball in 5 different positions within the forming tool. The center of the cavity in the forming tool is the ideal position, at which better results are expected. Two positions forward at +0.1 mm und +0.2 mm and two positions backward at -0.2 mm and -0.1 mm from the center point are chosen. After the ball is positioned at the selected point the forming is started. All the samples are made with the same processing parameters. Fig. 15 shows different parts before (a) and after forming (b+c).

![Fig. 14. Experimental setup.](image)

![Fig. 15. Positioning experiment results.](image)

The targeted shape can clearly be recognized (Fig. 15b). It can also be seen that a material flow took place on the axial direction in the free end of the ball. Failures occur during forming mostly at the external positions -0.2 mm and +0.2 mm and for balls with diameter higher than 850 µm. At +0.2 mm the wire, which is harder than the sphere, tend to pierce it. As result the shape is separated from the wire and builds a sort of ring. At -0.2 mm for more than half of the tested samples were cut off the wire. For bigger spheres with diameter more
than 1 mm wings (Fig. 15c) or the previous mentioned failures appear at all investigated positions.

For preforms with diameters between 750 µm and 850 µm the shape is generated well and the material flow ahead of the shape depends on the position of the ball. Fig. 16 illustrates the relation between preform positioning and extension of the part ahead of the formed part for preform diameters of about 800 µm.

The higher flow takes place at +0.2 mm and the shape is comparable for all the positions. Defects occur only at +0.2 mm and this might be due to the deviation in the volume and the eccentricity of the ball.

For smaller material accumulations with diameters between 750 µm and 650 µm forming takes place at all positions. As in the previous case some samples are cut off of the wire at +0.2 mm. The best results in term of dimensions and form accuracy were reached with ball diameters within this diameter range. Fig. 17 shows a qualitative comparison of the shape for different ball diameters and at different positions. Next to the positioning the diameter of the preforms is an important factor.

Fig. 16. Length of the Material for different position of the initial ball.

<table>
<thead>
<tr>
<th>d_{k,i}</th>
<th>Position Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8 mm</td>
<td>-0.02 mm</td>
</tr>
<tr>
<td>0.7 mm</td>
<td>-0.02 mm</td>
</tr>
<tr>
<td>0.8 mm</td>
<td>0 mm</td>
</tr>
<tr>
<td>0.7 mm</td>
<td>0 mm</td>
</tr>
</tbody>
</table>

Fig. 17. Influence of preform diameter.
5. CONCLUSION AND OUTLOOK

In this study the forming of material accumulations in linked parts by plunge rotary swaging is introduced. The production system consists of a double-stage conveyor system, a part detection system and a rotary swaging machine. Forming experiments are the base for the positioning accuracy demand in the cavity of -0.2 mm to +0.1 mm, which is about 10 times higher than the performance of the introduced feeding system with a repeatability of better than 20 µm. The applied part detection system consists of a high speed line camera sensor and a FPGA-based data processing algorithm that fulfills both the resolution and the real-time demands.

Regarding the achieved accuracy of the part detection and potential of the feeding system performed work is a major step for the continuous micro production in linked parts. The system is assumed to be applicable even to processes with higher accuracy requirements.

In future work this arrangement will be connected with a field bus to enable a fully automated process. Another major challenge is the application of a material flux model to the conveyors to ensure a constant position and constant tension in the linked part during forming.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for Subproject C5 "Teileverbunde" within the SFB 747 (Collaborative Research Center) "Mikrokaltumformen - Prozesse, Charakterisierung, Optimierung".

REFERENCES

Influences on the FEM-results of different parameters of the rotating straightening process

Graziella Kreiseler

Fraunhofer Institute for Manufacturing Engineering and Automation
Graziella.Kreiseler@fraunhofer.ipa.de

ABSTRACT
Rotating straightening is a process used to accomplish high quality straightening results for wire. In this paper, the process will be described with all major parameters. These parameters include the kinematics of the process, the geometry of the machine carrying out the straightening and material information. The influence of some of these parameters will then be calculated using FEM. In this paper, the friction between the straightening brackets and the wire and the influence of the bending of the wire before straightening will be discussed. In order to evaluate, how strong the parameters’ influences are, the resulting geometry of the wire is used.

KEYWORDS: rotating straightening, FEM, friction

1. INTRODUCTION
In 2011 there were 1.514 million tons of steel produced [1], 6-8% of this goes into the production of wire [2]. Wire is a semi-finished product used for various applications in industry, such as springs for watches, shopping carts and construction steel. In order to manufacture wire, the raw material is pulled through various drawing dies and then coiled up for easier transportation. This results in residual stress and bending in the semi-finished product. For further manufacturing bending and stress have to be diminished, so that the following processes can work with a straight input material which acts in a predictable way. This is accomplished by straightening. There are different kinds of straightening, such as roller straightening and straightening by stretching [3]. This paper will focus on rotating straightening, which is also known as wing straightening, as a method to reduce 3D-stress and bending. Rotating straightening achieves very good straightening results but is one of the least known straightening process in manufacturing science. Most of the knowledge about the process is achieved empirically and the process is controlled by skilled workers and their experience. In the research at the Fraunhofer Institute for Manufacturing Engineering and Automation a finite element model of the process is used to enhance this knowledge. The results from these simulations are compared to experimental results.
1.1. Rotating straightening

A common process to straighten wire is roller straightening, in which wire is lead through different straightening rolls with decreasing displacements. The rolls are in one spatial direction. For reducing more stress and bending there is a possibility to have rolls in different directions after the first set of rolls [4]. The straightness is achieved by repetitive bending with decreasing curvature.

Rotating straightening is able to diminish 3D stress by rotating the straightening tool around the wire. The wire enters the process from a coil. It is unrolled and then lead through rolls before it enters the actual process, which is shown in Fig.1. This takes place continuously, meaning the wire has a constant forward velocity \( \vec{v} \) imposed by the rolls pressed on the wire before and after the straightening toll. The pressure of the rolls can be adjusted in the process.

In Fig.1 the main components of the straightening tool are shown. There are three brackets, which have different displacements \( d \), two more brackets in front and after those are not displaced and not shown. The brackets rotate around the wire with a rotational speed \( \omega \). The brackets are fixed in a straightening wing. Due to the superposition of \( \vec{v} \) and \( \omega \) it is possible to diminish 3D stress and bending in the wire. After the process, the wire is chopped into pieces of a given pre-adjusted length. The length is adjusted with a bedstop and the chopping is triggered by a measuring wheel. This results in non-constant movement in the process, which is not modelled in this first evaluation.

![Fig.1. Straightening system](image)

Other parameters characterizing the wire are the plastic elastic behaviour, the microstructure or the surface of the wire. Parameters for the process are e.g. the condition of the straightening brackets, the material of the brackets or the settings of the parts used to impose the feeding velocity.
1.2. Model

The modelling and simulation for this research was done with ABAQUS® Explicit V6.12. The simulation consists of three steps, first the displacement of the brackets, then the actual straightening with the movement of the wire $\vec{v}$ and the rotation $\omega$ of the brackets around the wire. The last step of the simulation is a spring back analysis. Figure 2 shows the model during step one and two. The upper picture shows setup before the straightening process, then the state after the displacement of the brackets. The third row shows the results in the middle of step 2 with rotating brackets. On the bottom there is the result from step 2. The third step is not shown because there are only minor changes to the shape of the wire.

The model focuses on the major parameters of the process, as given by experienced workers. The geometry of the brackets, their distance and the diameter of the wire are taken from an actual process, which is used to compare the results to experimental results. The brackets are modelled as rigid bodies. The wire is modelled using a plastic/elastic material without any strain rate dependency or dependency on temperature. The temperature can be neglected because the process is always in the range of cold work [5] with a maximum temperature in the process of about 100 °C. The strain rate dependency was analysed and there is a difference of less than 2.68 % for the yield stress and 3.1 % for the tensile strength of the wire. Other geometric boundaries are replaced by boundaries conditions. The forward velocity of the wire in step 2 is imposed as a constant velocity at the beginning of the wire. The brackets are also fixed by geometric boundary conditions. Furthermore the cutting of the wire and the adjustment of the length with the bedstop is not modelled.

![Fig.2. Modell](image)

The velocity $\vec{v}$, the rotational speed $\omega$ and the different displacements are main adjustable parameters of the process and are modelled. The influence of other parameters is not that obvious and has to be investigated. Two of these will be presented in this paper, the
influence of the bending of the wire before it undertakes the process, and the fiction between the brackets and the wire.

To evaluate the results of the simulations the resulting node coordinates after step 3 were exported. By using Matlab R2012a the nodes were divided in sections along the wire and the centre of the section was calculated. In the next step of the evaluation the perpendicular vector to the line connecting the beginning and end of the wire was calculated. Since the wire is bent in all three spatial directions the length of the vector $|\Delta g_i|$ was used to compare the results.

2. INFLUENCE OF MATERIAL BEFORE STRAIGHTENING

As mentioned above, the semi-finished wire is rolled up in coils for better handling. At the beginning of the straightening process the wire is uncoiled. In modelling this presents some problems. The exact radius of the coil changes over time from a larger radius in the beginning to a smaller radius as the wire is uncoiled during the process. In the actual process, the diameter can change several decimetres. The wire is also able to move within large areas during the uncoiling. Unfortunately, the exact state of the wire, e.g. residual stresses, cannot be determined before the straightening process and there is some uncontrolled straightening done by the uncoiling device. Furthermore the geometric boarders of the machine parts in front of the straightening tool are wide.

In order to evaluate the influence of this unknown part of the actual process, there was a numerical experiment with a bent wire and with straight wire as the raw material entering the process, to see the effects on the results of the simulation. For the bent wire an average radius of 0.3 m was approximated from the real process. Figure 2 shows the results of these simulations. Four experiments were carried out, with bent and straight wire and with different displacements $d$ and $d/2$ of the straightening brackets. The figure shows the shape $|\Delta g_i|$ of the wire along the length wire $|l_i|$.

![Fig.2. Results after simulating with bent and straight wire](image)

88
Table 1 shows the maximum of $|\Delta g_i|$ for each of the experiments. The difference between the maximum for the bent and straight wire for the displacement $d$ is 0.088 mm for $d/2$ 0.03 mm. The difference between $d$ and $d/2$ for the straight wire is 0.417 mm and therefore much larger. For the bent wire the difference between $d$ and $d/2$ is 0.359 mm and also larger than the difference between straight and bent wire.

| experiment               | max. $|\Delta g_i|$ [mm] |
|--------------------------|--------------------------|
| Straight wire, $d$       | 0.624                    |
| Bent wire, $d$           | 0.536                    |
| Straight wire, $d/2$     | 0.207                    |
| Bent wire, $d/2$         | 0.177                    |

The results show, that the influence of the bending is not as strong as the influence of position of the straightening brackets. This underlines the assumption, that the wire is completely plastified in all three spatial directions during the process, which erases the original bending of the wire and only marginal initial stresses are still in the wire after the straightening. This thesis is also supported by the experience from the skilled workers. The settings of the process are not adjusted as the wire unrolls.

### 3. INFLUENCE OF FRICTION

Another influence, which was evaluated during the research at the IPA, is friction. The wire is in contact with certain areas of the straightening brackets and in this region there is friction. The effect of the wear caused by the friction on the brackets was described in [6]. The value of the wear on the brackets after several hundreds of meters was measured. Since these effects only take place after several hundreds of meters of straightened wire, the brackets were modelled as solid bodies.

The real brackets are made of cast iron and are a wearing part. In the simulation presented here the influence of friction on the wire is evaluated. For this there was taken a constant value of 0.16 for the friction coefficient between cast iron and steel [7] and the contact between the wire and the brackets was set to friction using the Coulomb friction model. This is a simplified approach to the phenomenon of friction. Following [7] there are a lot of influences on the friction between two contact partners. One major influence is the speed of the process, which is relatively high in this process. [8] also stresses the influence on the abrasive wear and the parts broken from the surfaces by abrasion, which can have a large influence on the magnitude of the friction coefficient. Also there are other more accurate friction models than the simplified Coulomb’s friction model, in which the friction force is equal to the friction coefficient $\mu$ times the normal force. In a first attempt even though there are shortcomings of this model it will be used to evaluate the influence of friction. Other influences as described in [9] will also be neglected in this first numerical experiment.
Figure 3 shows the results of the simulation with all the mentioned limitations. Similar to the results presented for the original state of the wire, the influence of the friction is weaker than the influence of the displacement of the brackets.

The maximum values for the experiment with and without friction are given in the following Table 2. The differences between experiment with friction and without for the displacement d is 0.066 mm and for d/2 0.026 mm. The differences for the displacements are much larger.

| experiment            | max. $|\Delta g_i|$ [mm] |
|-----------------------|----------------------|
| With friction, d      | 0.808                |
| Without friction, d   | 0.742                |
| With friction, d/2    | 0.230                |
| Without friction, d/2 | 0.204                |

In order to evaluate the influence further it is planned to do experiments in the actual process with and without lubrication.

4. CONCLUSION

In this paper an estimation for two parameters influencing the rotating straightening process is given. Following researches have to model these parameters not only separately but also in addition to one another and also different displacements will be modelled.

As mentioned above there are plans to compare the results of the simulation with and without friction to experimental results. For this there should be experiments without lubrication and experiments with lubrication. Lubrication diminishes the friction coefficient between the two surfaces and the results should be comparable to those gained by the simulation. [10]
After the further investigations and the comparison to experimental results, it should be possible to optimize the settings of the machine and look for solutions outside the normal areas. Key questions for this are:

- Are there uncommon settings which lead to good straightening results?
- Is there a possibility to change the dimensions of the straightening wing for better results?

REFERENCES:


Finite Element Analysis on the Friction Effects in the Gear Rolling Process

Alireza Khodae, Arne Melander

KTH Royal Institute of Technology, Sweden
khodaee@kth.se

ABSTRACT

Gear rolling as a manufacturing method for gear wheels has a number of advantages in comparison with traditional gear production methods which makes it an interesting topic for scientific research. The main benefits are reduced material consumption, fast cycle time of the process, improved strength of the product due to the alignment of the microstructure during cold forming and a good surface quality.

There are several factors which can influence the process quality and accuracy of the geometry in the final state. One of the most important parameters is the friction between the tool and the blank during the forming process. Therefore it is of interest to study the effect of different friction models in finite element simulations to evaluate its effects on final gear wheel shape.

In this paper the shear friction model is studied. Two friction factors will be studied in this case. Two dimensional FE calculations will be performed with the code DEFORM. The results of the FE simulations and the effect of friction factor on gear flank and tip geometry are presented. One example is presented of the effect on the so called “rabbit ear ”at the gear tip.

KEYWORDS: Gear Rolling, Finite Element Method, Friction, Friction model, Gear tooth geometry

1. INTRODUCTION

Gears are used in power transfer systems and for that reason significant attention is devoted to the design of sustainable manufacturing methods. Manufacturing techniques which do not involve material removal from the initial blanks have a special potential from the environmental point of view. In this paper gear rolling is studied where material removing operations are replaced by cold forming.

Gear rolling can be performed with flat or with round forming tools. A few studies [1-6] have been published where the manufacturing processes are studied with simulation techniques. This opens up the possibility of process planning tools for gear rolling where the process design can be optimised to meet requirements on high accuracy of the cold formed gear wheels.
The ambition of this work is to investigate the effect of the friction model, and the friction factor in the process of gear rolling. The results will be used in further process planning in industrial application of gear manufacturing with the rolling method. This research is done by formulating and running FE models describing the process of gear rolling. These simulations were designed to describe the rolling process of high gears with normal module around 4 mm, and addendum diameter around 100 mm.

Obviously, FE simulation without experimental verification cannot be reliable to be used as a reference. But it can give us a good understanding of the process before starting the experimental activities in the future.

Special attention will be devoted to evaluation of some geometrical tolerances for the simulated gear wheels. In the end the influence of friction factor on final geometry and product properties will be described. [11]

2. **FINITE ELEMENT MODEL**

2.1. *Process model within FE*

To simulate the process of gear rolling, DEFORM 2D, FE code has been used. The material model considered is ideal plastic and tool is modelled as a rigid body. Material for work-piece is 16MnCr5, which is common steel for gear wheels. Other material models can be useful, but they need more investigations in the present time. Authors just present the results with one model which have been studied for one year and have good relevance to the real application.

The work-piece has a mesh with 3000 linear quadrilateral elements. With these settings, the simulation time with Intel core i5 2GHz, 4 GB RAM computing power, is about 3 hours for one rolling cycle.

2.2. *Friction model*

Material flow in metal forming application is generated due to the displacement transferred from the tools to the deforming object. Therefore the frictional forces between the die and work-piece have great influence on the process mechanics. The most common friction model to describe the friction in such applications is the frictional shear stress model which is expressed as:

\[ \tau = f \sigma \]  

Where, in (1), \( \sigma \) is the flow stress, \( \tau \) is frictional shear stress and \( f \) is the friction factor. Generally for many of the metal forming applications this formulation gives adequate representation of friction. [10]

The cold rolling of gears is run in a lubricated condition in special rolling machines. In this paper we have considered different shear friction factors to see the effect of friction in the rolling of high gears with the help of finite element method.

Two different shear friction factors are considered. One for very good lubrication with \( f=0.1 \). The second friction factor is \( f=0.4 \) for dry cold rolling or poor lubrication.
3. RESULTS AND DISCUSSION

3.1. Friction factor effect on the geometry accuracy

To evaluate the final product geometry accuracy we need to compare it with the standard product geometry. In this part we present the standard product as the reference for measurements. The measurement results on the simulated gears will show the errors with respect to this standard.

Reference profile:

In this work, we have considered a gear wheel with the following parameters. This gear wheel middle section is shown in Fig.1.

Helical Gear:

Helix angle: 20°  
Pressure angle: 20°  
Normal module: 4 mm

Number of teeth: 21  
Addendum diameter: 103 mm  
Root diameter: 84.75 mm

Fig. 1. 2D view of the reference product and geometrical definitions

Geometrical variations:

Based on simulation results, we can measure the single pitch deviation and also do a span measurement over four teeth to illustrate the differences to the standard profile. These
measurements are performed at the end of 12th, 13th and 14th cycles of rolling to compare the effect of different friction factor but with all other process parameters constant. In addition to these two parameters one interesting result is to check how the addendum diameter and root diameter of work-piece are changing at the end of these cycles. Graphical illustration of these measurements has presented in Fig. 2 to Fig. 5.

In Fig. 2 we can see the effect of friction on the single pitch on the profile at the end of each cycle. It can be seen that the number of cycles to reach the standard pitch is larger for the low friction factor than for the high friction factor. Figure 3 shows that the trend of variation in span measurement is same as for the single pitch and with lower friction we reach the standard span value for a larger number of revolutions. This generally can be explained by the action of high friction to lock the material at the tool interface and that way generate more plastic deformation in the material. Also from Fig. 2 and Fig. 3 we can see that a small friction factor generates small errors compared to the standard profile. This can be especially observed in Fig. 3.

In Fig. 4 and Fig. 5 the results of measurement on the addendum diameter and root diameter of the work-piece are presented. A high friction factor will provide faster changes in the geometry. This can especially be observed in Fig. 4 for the addendum diameter. It can clearly be seen that the same rolling cycle number will lead to a larger diameter with the higher friction.

Considering the root circle, the Fig. 5 does not predict too much variation by adding more number of rolling cycle on the work-piece in the end of process. The main reason for this behaviour can be that in the last cycles of rolling most of material deformation is occurring on the flanks and close to the tips but not on the root area.

Figure 5 shows that with higher friction can we have closer values to our target than for lower friction. This can be explained since the higher friction can help to dig out the root during tool rotation.

![Graph showing single pitch measurement on the profile at the end of full cycles](image-url)
Fig. 3. Single pitch measurement on the profile at the end of full cycles

Fig. 4. Addendum circle measurement on the profile at the end of full cycles
Material response during rolling:

To study the material response during rolling several different parameters can be evaluated. We have chosen to present plastic strain and the accumulated damage.

Effective strain shows to what extent we have deformed the material. This will also let us know in which area of the profile we have the highest flow stress. Figure 6 and Fig. 7 show the effective plastic strain distribution in the end of 13\textsuperscript{th} and 14\textsuperscript{th} rolling cycles and with different friction factors.

From Fig. 6 and Fig. 7 we can see that more friction will generate more strain at the same number of rolling revolutions. This is predictable since friction will help to generate material displacement over the surface of the gear flank. The high strain values are near to the root and this will provide higher flow stress at the root. This could be relevant if the gear rolled component is used without additional heat treatment.

Damage generation in a material is very important to follow since it is an indicator of the risk of crack generation during forming.

Damage can be described with different models. In this work the “Normalized C&L” damage model was used.

\[ \int_0^\varepsilon \frac{d\sigma_{\text{max}}}{d\sigma} d\varepsilon \leq C \]  

\( \sigma_{\text{max}} \) is the maximum principal stress in the material. \( \sigma \) is the effective stress in material. \( \varepsilon \) is the effective strain. \( C \) is the critical value for the damage where cracking can be initiated.

Figure 8 and Fig. 9 show the damage distribution at the same position as was used to present the strains. We can see that the damage distribution pattern is similar for all the figures in Fig. 8 and Fig. 9. There is however a slight tendency that internal damage is higher for the low friction factor, compare to surface damage. The damage lever is however higher for the higher friction factor. This is mainly caused by higher strain levels.
Fig. 6. Effective strain on one sample tooth with $f = 0.1$
  a) in the end of 13\textsuperscript{th} cycle, b) in the end of 14\textsuperscript{th} cycle

Fig. 7. Effective strain on one sample tooth with $f = 0.4$
  a) in the end of 13\textsuperscript{th} cycle, b) in the end of 14\textsuperscript{th} cycle

Fig. 8. Damage on one sample tooth with $f = 0.1$
  a) in the end of 13\textsuperscript{th} cycle, b) in the end of 14\textsuperscript{th} cycle
3.2. Rabbit ear shape

Figure 10 and Fig. 11 show the final product tip shape based on two different friction factors. The tip acquires a special shape, which referred to as “rabbit ear” in the literature and is a typical phenomenon in the gear rolling process. The present simulations showed that the rabbit ear can differ because of the friction factor.

From these figures, we can conclude, with lower friction we have more symmetric rabbit ears than with higher friction. This is explained based on the amount of material displacement at the tips. In the rolling process for gears we have the maximum change of shape around the tips during the final cycles. Higher friction will generate larger folded shape for rabbit ears, and lower friction will form more symmetric rabbit ear shapes.

3.3. Process time

The results of FE simulation showed that the lower friction will need higher number of cycles to fully form the blank into the final product diameter. This will require longer time as well. Because the time for each cycle with these process settings, were 3.5 seconds and the maximum forming time were recorded as 49 seconds. Generally these data just can confirm the fast process time and shorter manufacturing time for the gear wheels if we compare with traditional methods like hobbing or milling.
4. CONCLUSION

In this paper two different scenarios for friction modelling in gear rolling was studied. The results of FE simulations showed that the lubrication condition has great influence on the product geometry and process time.

With the lower friction factor, which can simulate good lubrication and proper lubricant properties, it is shown that the geometries of the gear wheel flanks are more favourable. The effect of the friction on the tip geometry, where we have rabbit ears, where also studied, and we observed that lower friction will help to keep more symmetry on the tips and reduce the rabbit ear tendency.

We have shown that the process time will be influenced by the lubrication condition and the better lubrication will need slightly larger number of rolling cycles and thus a longer rolling time. But the effect was not very significant since cycle time is very short. Further studies on the material modelling and also verification of the FE results with real experiments are necessary.

ACKNOWLEDGEMENTS

The authors are indebted to the national Swedish Government Agency VINNOVA for financial support from the FFI program under the contract SMART.

REFERENCES


Theme 5

Assembly and Automation Technology

Electronic Component Cleaning in Remanufacturing
M. Bauer\textsuperscript{1}, S. Thäter\textsuperscript{2}, S. Schweinstig\textsuperscript{1} and R. Steinhilper\textsuperscript{2}
\textsuperscript{1}Fraunhofer IPA, Germany
\textsuperscript{2}University of Bayreuth, Germany

Vision-assisted and 3D model-based remote assembly
Abdullah Mohammed, Lihui Wang
KTH Royal Institute of Technology, Sweden

Minimising energy consumption for robot arm movement
Bernard Schmidt\textsuperscript{1}, Abdullah Mohammed\textsuperscript{2}, Lihui Wang\textsuperscript{1, 2}
\textsuperscript{1} Virtual Systems Research Centre, University of Skövde, Sweden
\textsuperscript{2} KTH - Royal Institute of Technology, Sweden

Detection of Battery Screwdriver’s Optimal Working Regimes
Ivans Grinevich, Natalija Mozga, Guntis Strautmanis, Oskars Lininsh
RTU Faculty of Transport and Mechanical Engineering

Supply chain collaboration on the competitiveness of Basque Country manufacturing companies.
Alberto de la Calle, Esther Alvarez, Arantza Múgica, Olatz Ukar
Universidad de Deusto, Bilbao, Spain
International Conference on Advanced Manufacturing Engineering and Technologies

Electronic Component Cleaning in Remanufacturing

M. Bauer¹, S. Thäter², S. Schweinstig¹ and R. Steinhilper²

¹Fraunhofer IPA, Germany
²University of Bayreuth, Germany
Marco.Bauer@uni-bayreuth.de

ABSTRACT

Nowadays in the automotive industry a move away from purely mechanical to electronic assemblies can be observed. In order to protect these electronic assemblies from environmental influences as well as from undesirable access, they often are sealed with different compounds e.g. silicone gels or epoxy resins. Therefore new challenges for the remanufacturing of electronic assemblies with regard to removal of sealants, a definition, creation, and analysis of technical cleanliness have arisen. These are taken up as a part of the research project "eCleanER - Electronic Component Cleaning Engineering for Remanufacturing". First research results about various procedures for the gentlest removal of sealants from electronic assemblies are for the first time described in this paper.

Keywords: remanufacturing, electronic components, sealants, solvents, technical cleanliness

1. INTRODUCTION

Natural resources are necessary for the survival and development of mankind. Nearly all resources are limited and almost depleted or can only be obtained with great effort [1]. By applying a second product life cycle of individual assemblies by remanufacturing about 85% to 95% of the material as well as the energy consumption can be saved in comparison to the production of new assemblies. Due to less material and energy consumption the remanufactured assemblies can generally be offered at prices about 45% to 65% of equivalent prices of new ones [2]. In the past few years terms and definitions of reprocessing, recycling or reuse were frequently used as synonyms for remanufacturing – sometimes without a clear differentiation. The differentiation is necessary due to the different flow of material. Reuse describes an assembly that can be used and implemented in another field of application without changing the existing properties. The term recycling specifies the conversion, as well as a substantial recovery of no longer needed or serviceable products into secondary material. Recycling is divided in recycling of materials and recycling of products (remanufacturing) [3] [4]. Remanufacturing is the ultimate form of recycling and includes all necessary steps to remanufacture used or damaged products to the quality of new products. In order to achieve these conditions, few parts which don’t comply the requirements of a whole product life cycle
have to be replaced. The six key steps of remanufacturing, on the example of an electronic control unit (ECU), can be seen in Fig. 1.

![Remanufacturing process chain of mechatronic systems](image)

**Fig. 1: Remanufacturing process chain of mechatronic systems [5]**

Thereby the remanufactured assemblies achieve qualities of new ones. Remanufacturers not unusually give the same warranty on their products as the manufacturers of new products. These manufacturers also realized the potential of remanufacturing. An example is Bosch with its eXchange-program [6].

The proportion of electronic parts in cars has risen extremely in recent years and will increase even further in the future, as more and more electronics play a key role in the implementation of many innovations in the car [7]. Especially engine and transmission control units usually got the highest requirements of reliability and endurance. These control units are more and more sealed to protect them from environmental stress. Sealant specifies a kind of casting technology, surfacing or surrounding a liquid sealant on a circuit board or a casing [8]. The different kinds of usage of electric and electronic assemblies as well as their differentiated environmental influences led to an enormous assortment of sealants. The selection criteria are mainly protection:

- from humidity
- from extreme temperatures
- from undesirable access
- of company know-how
- from intense mechanical stress.

The scientific and technical experience in the field of the remanufacturing of engine and transmission control units as well as the carefully substrate removal of sealants is limited. Furthermore those assemblies are, besides the entertainment electronics, the most expensive assemblies in the car. In an economic point of view an investment in the industrial remanufacturing is able to depreciate in short time.
2. INITIAL SITUATION

Remanufacturing companies, specialised in electronic assemblies are interested in remanufacturing of wire bonds, braze points and defective electronic components. A further offer of companies is the restoring and updating of software. For the remanufacturing of electronic assemblies these companies need to remove the sealants from the assemblies, especially from circuit boards. This is the third step in remanufacturing process chain and creates the technical cleanliness.

The previously described sealants are classified in single and multi component systems depending on their compound. Single component systems are premixed substances that can be used directly after opening the manufacturers packaging. Thereby the curing process of the ingredients is mainly initiated off a stated starting temperature. The curing process of multi component systems already gets initiated at room temperature. The substance of the sealant consists of two separately stored components. Those are a liquid matrix, additives and an accelerator, a catalyst or an activator. The mixed components have a low viscosity before the reaction. For this reason those components are suitable for small and angular geometries. The reaction is mostly endothermic with no activation energy at room temperature.

The goal of the following experiments is to separate the sealants from circuit boards. In this case the separation is the removal of substances or mixtures of substances by abolishing the force of adhesion and/or cohesion [9].

![Abolition of the force of adhesion (left) and cohesion (right)](image)

In this regard the force of adhesion is the force of the sealant on the surface of the circuit board (Fig. 2, left). It is caused due to the interaction of the molecules of the sealant and the circuit board surface. Especially the cleanliness of the surface (before sealing) has an essential impact on the force of adhesion. The force of cohesion is the intermolecular force within the sealant which is responsible for its coherence (Fig. 2, right). As there are no experiences on the field of the selective removal of sealants, the following experiments can be regarded as preliminary tests. Therefore all potential applicable chemical, thermal and mechanical processes shall be analysed upon their efficiency. The process of cleaning has to achieve good results also for geometrical complicated proportions. The conductor plates in some extent exist of overlapping components (e.g.: chips, capacitors, dissipators, inductors etc.). The cleaning processes have to detach the used sealant completely free of residues in order to reseal the components after remanufacturing. Fig. 3 shows on the left side the sealant of an opened electronic control unit and on the right side same unit with a partially removed sealant. Between the contacts of the components are still parts of the sealant remaining.
3. CHOICE OF REPRESENTATIVES

The automotive industry uses a lot of different sealants. Seven of those are picked representatively. This choice must not be done through a simple ranking among the occurrence of the different sealants and control units. Because various influencing factors like the profit margin, workload, kind of sealant or sustainability are not accounted. Instead of that the Analytical Hierarchy Process (AHP) as well as the Cost-Utility Analysis (CUA) are offering a good possibility for the choice of adequate representatives. The Analytical Hierarchy Process is an essential approach in the decision theory. This mathematical method was developed by Thomas L. Saaty in the 1970s. The intent and purpose of this method is to involve rational and intuitive alternatives in the process of decision [10]. For this reason the decider deploys comparisons for each couple, which are used to define a priority of ranking. Following criteria represent all important decision-making parameters.

3.1. Costs

The costs are split into material and labour costs. They represent the time exposure and monetarily effort that has to be adduced in remanufacturing. In this case the material costs are such as the kind and amount of solvent and/or other additives/operating fluids that are used during remanufacturing. Labour costs reflect the time exposure of individual manual work respectively the level of automation of the existing process.

3.2. Quantity

The quantity refers on the one hand to the number of remanufactured control units presently executed at the company, on the other hand to the parts expected to be remanufactured in the next 5 years. Relevant to the evaluation of the parts in the next 5 years is the amount of assemblies circulating and whether they are still produced/installed and accordingly since they were not installed anymore. Furthermore the share in the market plays a major role. The future market share is determined on the basis of the current market share in the OE as well as the remanufacturing sector plus the average endurance of the units.

3.3. Sealant

The criteria sealant is split into the feasibility for remanufacturing and the needed workload. The feasibility characterizes the present effectivity of the sealant removal. Parameters like layer thickness, consistency and mixture of the substances are playing a major
role. On one hand the needed effort summarizes the capability for automation, as well as the machine and tool usage. On the other hand it offers the opportunity to evaluate factors like the application of chemicals and the occupational safety and health.

Resulting from the ranking and selection process the top seven representatives, out of the wide range of automotive control units, are selected. Due to the fact of the enormous scope of the calculations and confidential information, detailed results cannot be shown in this paper. Regarding the following experiments, the henceforth representative sealants from the chosen control units are defined through the acronyms in Table 1. The material of the sealant is unknown.

Table 1: Description of the samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Characteristics</th>
<th>Type of sealant</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-1</td>
<td>Grey, rubber-like</td>
<td>Packing compound</td>
</tr>
<tr>
<td>P-2</td>
<td>Black, sticks on surface of housing</td>
<td>Joint compound</td>
</tr>
<tr>
<td>P-3</td>
<td>Black to grey, sticks on surface of housing</td>
<td>Joint compound</td>
</tr>
<tr>
<td>P-4</td>
<td>Skin-colored to white, hard</td>
<td>Sealant</td>
</tr>
<tr>
<td>P-5</td>
<td>Clear-transparent, jellylike, sticky</td>
<td>Sealant</td>
</tr>
<tr>
<td>P-6</td>
<td>Amber, rubber-like, supple</td>
<td>Sealant</td>
</tr>
<tr>
<td>P-7</td>
<td>Grey, rubber-like</td>
<td>Packing compound</td>
</tr>
</tbody>
</table>

4. EXPERIMENTS AND RESULTS

To accomplish the following preliminary tests, samples were picked by chosen representatives (Fig. 4). In this test, first experiments to overcome the forces of adhesion and / or the forces of cohesion are made.

Fig. 4: Glasses with samples from the representatives

4.1. Drying by sublimation

The effect of sublimation is used in the field of freeze drying. It describes the direct state change of a substance from the solid to the gas phase [11]. Hereby destructive drying processes, occurring at great heat, can be avoided and it is possible to produce a product with a very small portion of volatile substances. This experiment aims to remove all volatile substances out of the compound, to change its properties. The sublimation is executed in a freeze dryer. In this process the samples are treated at 0.470 mbar for 75 hours. This is enough time for the diffusion of all volatile substances out of the sealant and rises linear with the geometric sample size. The temperature of the samples is between -4 °C to -5 °C.
With this process the characteristics and an enduring change of the sealant cannot be reached. A reason therefore is a too low integral share of volatiles in the sealant. In consequence of the vacuum the structure of the polymer is not changed.

4.2. *Viscosity determinations during temperature increase*

A change of the temperature leads to a modification of the effective velocity on molecular level. The aim is to loosen the network structure of the sealants by heating them and to be able to remove the sealant easier from the circuit board. The chosen temperature range is between 20 °C to 120 °C and is limited by the specification of electronic components in the automotive industry. Thereby the samples are heated indirectly. The experiment is accomplished with a rotational rheometer and is splitted in two different measurements, the recording of the amplitude sweep and frequency sweep. The samples are analysed under a roughened cone-plate-system at an opening angle of $\alpha=1°$ regarding their viscous and elastic properties. The shear modulus $G'$ is characteristical for the elastic region and describes how a deformation is totally regenerated after an endless long time. The viscous part is specified by the viscosity $\eta$ and represents the load at which non reversal deformations take part at the substances.

The results of the experiments shows, that the interlacing of the molecular network structure can’t be melted or cut by a moderate temperature rise.

4.3. *Reduction of temperature*

To avoid damaging of the electronic components, the samples are cooled slowly down, for 12 hours from room temperature to -18 °C. In this process the normally at room temperature very gluey and flexible mass becomes brittle. The low temperature allows only a restricted motion of the molecular network structure. This leads to reduced adhesion forces of the sealant to the circuit board and the sealant is less gluey and sticky. Changes of the appearance were not determined. By slowly heating up to room temperature the characteristics of the sealant are restored to its normal. Indeed the temporary reduced temperature changes the molecular network structure to be very inflexible, however the temperature difference is not sufficient enough to cause significant and enduring structural and material changes in the sealants.

4.4. *Aqueous ultrasonic supported immersion cleaning*

The ultrasonic cleaning is based on the principle of cavitation. A converter produces ultrasonic waves which are carried through a transfer fluid to the sample that needs to be cleaned. Those waves create small gas bubbles on the surfaces, which implode immediately. In consequence of the rapid pressure change a force results on the surface of the treated circuit board and its sealant. This force effect is supposed to affect the surface of the samples at the barrier between circuit board and sealant and finally separate them from each other. Important factors on the effectiveness of the ultrasonic cleaning are the frequency and temperature. The frequency has a direct impact on the size of the gas bubbles and those on the resulting force. With the temperature different properties of the fluid (as vapour pressure, viscosity, surface tension) correlate.

The ultrasonic bath is heated up to 70 °C and filled with distilled water and a basic cleaner plus a degreasing amplifier. The sample is put into the ultrasonic bath for one minute at 80 kHz, for another minute at 37 kHz and further two minutes at 37 kHz. These frequencies and temperatures are given by the used ultrasonic bath. The times are determined by previous successful cleaning experiments. After the first minute at 80 kHz there is no external
transformation visible. After the decrease to 37 kHz little changes at the elasticity of the sealant can be detected. Further treatment at low frequency causes that the sealant is easy abatable from the surface of the circuit board. Additionally the elasticity rises and during tension horizontal cracks are producible. Partially some pieces with smooth breaking edges can be removed. This process is in practice only in some extent applicable, as most of the electronic parts are not functional anymore. This is on the one hand because of the application of watery cleaner and on the other hand because of the direct impact of the ultrasonic on the components.

### 4.5. Solvents and material analysis

In order to adapt all following chemical experiments on specific substances the representative sealants have to be analyzed. The use of the Nuclear Magnetic Resonance-Spectroscopy (NMR-Spectroscopy) promises fast and good results. Therefore the samples have to be completely dissolved in a deuterated solvent. For the reduction of costs, first attempts in dissolving the sealants of the chosen representatives, various common used solvents are picked. Experiments with different solvents are providing two important results. On the one hand first findings about the solubility of the compounds in different solvents regarding to the ablation from the circuit board surface are gathered. On the other hand a suitable solvent for the NMR-Spectroscopy can be found. The selection of the solvents, as seen in Table 2, is arranged for covering a wider area of solubility in Teas chart in order to localize compounds the representatives are made of as good as possible [12].

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Solvents</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>acetone</td>
<td>polar apriotic</td>
</tr>
<tr>
<td>B1</td>
<td>n-butanol</td>
<td>polar-protic</td>
</tr>
<tr>
<td>C1</td>
<td>chloroform</td>
<td>non-polar</td>
</tr>
<tr>
<td>D1</td>
<td>diethyl ether</td>
<td>non-polar</td>
</tr>
<tr>
<td>E1</td>
<td>ethanol</td>
<td>polar protic</td>
</tr>
<tr>
<td>E2</td>
<td>ethyl acetate</td>
<td>polar-apriotic</td>
</tr>
<tr>
<td>M1</td>
<td>methanol</td>
<td>polar protic</td>
</tr>
<tr>
<td>T1</td>
<td>toluene</td>
<td>non-polar</td>
</tr>
<tr>
<td>W1</td>
<td>WD-60</td>
<td>penetrating oil</td>
</tr>
</tbody>
</table>

The selection of solvents also includes penetrating oil, which isn’t a solvent in the narrower sense, but seems to have a great use in the removal of silicone joints. It is used to lower the adhesion forces between the sealant and the body and/or the circuit board surface instead of breaking the cohesion forces.

Small amounts are taken from the representatives (see Table 1) and distributed to the testing glasses equally. The sample masses are between 0.1 g and 0.3 g, depending on the characteristics of the samples. First effects caused by the interactions between solute and solvent can immediately be seen after adding the solvent to samples respectively after a few minutes. The results after a penetration time of 3 hours can be seen in Table 3.
Table 3: Results of penetration with current solvents

<table>
<thead>
<tr>
<th>Sample / Solvent</th>
<th>A1</th>
<th>B1</th>
<th>C1</th>
<th>D1</th>
<th>E1</th>
<th>E2</th>
<th>M1</th>
<th>T1</th>
<th>W1</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-1</td>
<td>□</td>
<td>□</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>P-2</td>
<td>□</td>
<td>□</td>
<td>●</td>
<td>●</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>P-3</td>
<td>●</td>
<td>□</td>
<td>●</td>
<td>●</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>P-4</td>
<td>□</td>
<td>▲</td>
<td>▲</td>
<td>▲</td>
<td>▲</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>P-5</td>
<td>▲</td>
<td>□</td>
<td>▲</td>
<td>▲</td>
<td>▲</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>P-6</td>
<td>▲</td>
<td>□</td>
<td>●</td>
<td>▲</td>
<td>●</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>P-7</td>
<td>□</td>
<td>□</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

- ■ completely dissolved (not achieved)
- ▲ partially dissolved
- ▲1 adhesion forces are lowered significantly
- ▲2 adhesion forces are lowered completely
- □ no visible interactions
- ● visible swelling

Chloroform has the greatest overall effect of the used solvents. An obvious swelling of the bulk can be recognized by P-2 and P-3. Additionally P-3 is bleached out. Because of the fact acetone shows same solubility behavior on P-3 and both solvents have similar values for $f_h$ in Teas Chart, a mixture of them has to be tested in next attempts. P-4 shows an obvious swelling, too. Furthermore small parts are dissolved and are visible as dispersion in the solvate. P-5 has almost been completely dissolved in chloroform. Only small constituent parts are dispersed in the solution. This leads to the conclusion the sample exhibits similar interaction forces like chloroform. Or in other words, it has to be found in Teas Chart close to chloroform. Another remarkable effect can be seen on sample P-5 in combination with WD-60. The penetrating oil completely surrounds the sample and lowers, respectively overcomes the adhesion forces so it doesn’t stick anymore to the walls of the testing glasses.

Some other solvents have shown effects too, as can be seen in Table 3. On the other side for some solvents no visible effects are noticed. It has to be pointed out that this first attempt was a try to locate the unknown samples and to see if there are solvents that affect the sealants.

Based on these first findings about the solubility of the samples, chloroform respectively deuterated chloroform is the best choice for the following analyses. To identify the dissolved components a NMR-Spectroscopy is made. Sample P-4 (Fig. 5) can be identified as a compound of paraffines (Peak 1), polyethylene glycol (Peak 2), alcohols (Peak 3) and aromatics like phthalates (Peak 5), which are used as softening agents in the sealant. Peak 4 is deuterated chloroform used as solvent in the NMR.
Fig. 5: NMR-Spectroscopy of P-4

The NMR-Spectroscopy (Fig. 6) identifies the material of sample P-5 to 99% to be poly siloxane (silicone). The first Peak on the right side (Peak 1), surrounded by the small satellite peaks is a lead to silicones (besides the internal standard, TMS). Peak 2 leads to carbonyl groups and ethers. Peak 3 is again the solvent deuterared chloroform.

Fig. 6: NMR-Spectroscopy of P-5

In the next attempts more solvents and especially mixtures of different compounds have to be used in order get a complete solution and to monitor the effects of the solvents on electronic control units. Therefore a variation of process parameters is possible.
5. CONCLUSION

Analyses in the fields of electronics, remanufacturing of electronic automotive parts and manufacturing of electronic control units have shown a very great demand for the removal of sealants from circuit boards and the corresponding technical cleanliness. First experimental results have shown, that ablating of sealants is a very ambitious task. Mostly the substances, especially in older compounds are unknown, sometimes also by new control unit manufacturers. To identify the substances, more nuclear magnetic resonance spectres will be gathered with fitted parameters and the optimized solvents. Therefore the follow experiments will be able to use adjusted and optimized mixtures of different solvents.

ACKNOWLEDGMENTS

The research project eCleanER is supported by a grant from the German Federal Ministry of Economics and Technology to University of Bayreuth and Fraunhofer.

REFERENCES

Vision-assisted and 3D model-based remote assembly

Abdullah Mohammed, Lihui Wang

KTH – Royal Institute of Technology, Sweden
abdullah.mohammed@itm.kth.se

ABSTRACT

Most of today's industrial robots work in dynamic environments that require flexibility and adaptability during the operating time. Therefore, there is a growing need to develop robotic systems that can adapt to different scenarios using advanced technologies. This paper proposes an approach for an online programming-free model-driven system that utilises web-based distributed human-robot collaboration architecture to perform distant assembly operations. The developed system uses an industrial robot to assemble components unknown in advance. It uses a robot-mounted camera to capture the silhouettes of the components from different angles. Then the system analyses those silhouettes and constructs the corresponding 3D models. Using the 3D models together with the model of a robotic assembly cell, the system guides a distant human operator to assemble the real components in the actual robot cell. The results show that the developed system can construct the 3D models and assemble them within a suitable total time.

KEYWORDS: remote assembly, vision-based 3D modelling, robot control.

1. INTRODUCTION

The last few decades witnessed growing needs for globalisation to leverage production cost vs. market share. Therefore it is most likely that distributed manufacturing will became more common and play a highly important role in factories of the future. In a distributed environment where human operators and production equipment are not collocated, mutual collaboration towards remote real-time manufacturing is in demand. One example is human-guided robotic assembly.

The objective of this research is to bridge the gap between humans and robots while fully utilising the strength of each in remote real-time assembly. A novel approach for the real-time assembly achieved by remote human-robot collaboration is proposed in this paper. In the developed prototype, 3D models are used to guide an off-site human operator during remote assembly. 3D models of real physical components to be assembled by a robot are created on the fly based on the shapes’ silhouettes captured by a robot-mounted camera. The camera is turned off during assembly to have a better communication performance. In this context, the robot is treated as a manipulator which mimics the human’s operations remotely.
2. RELATED WORK

During the recent years, researchers have developed various tools to program, monitor and control the industrial robots. The aim is to reduce possible robot downtime and avoid collisions caused by inaccurate programming, through simulation. However, these tools require pre-knowledge about a robotic system. Introducing unknown objects to the robotic system may cause a breakdown to the simulation due to no-longer valid robot programs.

Both laser scanners and vision cameras are common techniques to convert unknown objects to virtual 3D models. Modelling objects using stereo vision cameras was a main focus for several researchers [1-3], whereas others, including [4] adopted a pre-defined library of 3D models to match the real desired objects. However, the stereo vision camera-based approach suffers from two drawbacks: (1) the utilised equipment is expensive and less compact, and (2) it lacks the ability to capture and model complex shapes from fixed single viewpoint due to limited visibility.

2D vision systems can also be applied to model unknown objects. By taking a number of snapshots of an object from different viewports, the object can be modelled based on analysing the silhouette in each snapshot. For example, [5, 6] focused on modelling the object in high accuracy and with details.

Despite the fact that these approaches were successful in their reported applications, they are unable to model multiple objects in a single run. Besides, they lack the ability to model objects remotely.

In this paper, we propose a new approach to constructing 3D models of multiple arbitrary objects, simultaneously, based on a set of snapshots taken for the objects from different angles. This approach is implemented through a system that allows an operator to perform assembly operations from a distance.

3. SYSTEM OVERVIEW

The proposed system demonstrates the ability of identifying and modelling arbitrary incoming objects to be assembled using an industrial robot. The new objects are then integrated with the existing 3D model of the robotic cell in a structured environment – Wise-ShopFloor [7], for 3D model-based remote assembly.

The system consists of four modules: (1) an application server, responsible for image processing and 3D modelling; (2) a robot, for performing assembly operations; (3) a network camera, for capturing silhouettes of unknown/new objects; and (4) a remote operator, for monitoring/control of the entire operations of both the camera and the robot. The system is connected by an Ethernet and the Internet. Figure1 shows the details of the developed system.

The network camera is mounted near the end effector of the robot. First, the robot moves to a position where the camera is facing the objects from above to capture the top-view snapshot. The system then construct the primary models of the objects by converting their silhouettes in the top-view snapshot to a set of vertical pillars with a default initial height.

After that, the camera is used to take a sequence of new snapshots of the objects from other angles. Projecting the silhouettes of each snapshot back to the 3D space generates a number of trimmed pillars. The intersections of these pillars identify the final 3D models of the objects. Figure 2 shows a simplified 2D trimming process of one object after the top-view snapshot, where the bounding polygon including errors is used to approximate the actual object.
4. METHODOLOGY

4.1. Image processing

Image processing steps are performed to recognise the features of the captured objects through their extracted silhouettes. The details of those steps are explained below.

Capturing snapshots

An IP-enabled network camera mounted near the end-effector of the robot is used to take a sequence of snapshots. These images are then sent to the application server for processing.
Converting to grayscale

The colour images are then converted to grayscale for reducing computational complexity, by taking the average value of RGB values of each pixel in the images.

Adjusting brightness and contrast

Finding the right pixel intensity highly relies on the lighting conditions of the working environment and the settings of the camera. Therefore, the brightness and contrast are adjusted based on the lighting conditions of the developed system.

Gaussian smoothing

A zero mean Gaussian filter is used to remove the noise from the image and improve the accuracy of extracted silhouette. It is achieved by applying Equation 1, where output image \( H(i, j) \) is the convolution of input image \( f(i, j) \) and the Gaussian mask \( g(k, l) \).

\[
H(i, j) = f(i, j) * g(k, l) = \sum_{k=-[l]}^{(n-1)/2} \sum_{l=-[m]}^{(m-1)/2} f(i-k, j-l)g(k, l) \tag{1}
\]

The discrete form of convolution is performed which goes through each element in the convolution mask and multiply it with the value of the corresponding pixel of the input image; the sum of these multiplications is assigned to the pixel in the output image.

Image thresholding

This process identifies the silhouette pixels in the image by assigning a certain intensity values to them. It is started by scanning the image pixel by pixel while comparing its intensity value with a threshold value. Each pixel in the image will have either white or black intensity value depending on whether it is higher or lower than the threshold value.

Silhouettes labelling

This process is to assign a specific label for each silhouette in the image. The connected component labelling algorithm [8] is chosen due to its efficiency. The process starts by scanning the image pixel by pixel to find one that belongs to one of the silhouettes, followed by examining its neighbouring pixels. If one or more neighbouring pixels already have a label, the algorithm assigns the lowest label to the pixel. Otherwise, a new label is assigned. The outcome of labelling operation is a two-dimensional array where each element represents a pixel, and each silhouette is represented by a unique label.

4.2. 3D modelling

Camera Calibration

The mathematical model of the camera is defined using the pinhole camera model [9] due to its acceptable level of approximation. Constructing that model requires camera’s calibration to determine its parameters and identifies the physical location of it. The
calibration includes: (1) focal length, (2) optical centre, (3) radial distortion coefficients, and (4) tangential distortion coefficients. Figure 3 illustrates some of the parameters.

![Fig.3. Image plane in the calibrated camera](image)

The camera’s position and orientation with respect to the robot’s end-effector is described as well using a transformation matrix. Since the camera is mounted near the end-effector of the robot, the calibration needs to be performed only once as long as the camera has a fixed position and orientation with respect to the end-effector.

![Fig.4. Relationships between coordinate systems](image)

To construct the 3D models, a coordinate frame is defined and placed at the optical centre of the camera. The transformation matrix between the base and TCP (tool centre point)
of the robot are known as a priori, and added to the transformation matrix between the TCP and the camera’s centre. Together, they define the relationship between the base coordinate system of the robot and that of the camera. Figure 4 describes the locations and specifications of those coordinate systems.

Another 2D coordinate system in the image plane must be defined which specifies the locations of pixels in a captured image to simplify the processing.

**Construction of pillars**

The first snapshot taken by the camera provides the top view of the objects. The system extracts first the silhouettes of the objects from that snapshot; these silhouettes helps the system to construct an initial representation of the 3D models. These models are represented by a set of pillars in 3D space, each of which corresponds to one pixel with a non-zero value in the image plane. The construction of the initial pillars is accomplished by applying Tsai’s camera model [9], as described in Equation 2 to Equation 4.

\[
x' = \frac{x}{z}, \quad y' = \frac{y}{z}, \quad r^2 = x'^2 + y'^2 \tag{2}
\]

\[
x'' = x'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2 p_1 x' y' + p_2 (r^2 + 2x'^2) \tag{3}
\]

\[
y'' = y'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2 p_2 x' y' + p_1 (r^2 + 2y'^2) \tag{4}
\]

where \(k_1, k_2, k_3\) are radial distortion coefficients and \(p_1, p_2\) are tangential distortion coefficients. The projected point on the image plane represented by UV pixels coordinate is calculated using Equation 6.

\[
f_x = f \times s_x, \quad f_y = f \times s_y \tag{5}
\]

\[
u = f_x \times x'' + c_x, \quad v = f_y \times y'' + c_y \tag{6}
\]

Equation 5 introduced two different focal length coefficients: \(f_x\) and \(f_y\); this is due to the fact that the individual pixels on a typical CCD image sensor are rectangles. Therefore, pixel’s dimensions are defined by \(s_x\) and \(s_y\).

Figure 5 illustrates the construction of the pillars. The system introduces a temporary height for the pillars to overcome the lack of depth information in the top view snapshot. This height is based on the maximum height of the objects. The extra length of the pillars will be trimmed off later during the process.
Trimming of pillars

The initial pillars need to be trimmed iteratively to construct the 3D models that are close-enough to approximate the real objects. Figure 6 shows the trimming process of one pillar as an example.

The trimming process starts after analysing the second captured image and extracting the silhouettes. First, it projects the pillars one by one from the 3D space to the image plane. Since each pillar is represented by two 3D end points, the projection is only for the points. The projection creates two 2D points on the image plane, calculated by Equation-5. Then, a line...
that connects the two 2D points is created using Bresenham algorithm [10]. The process continues by extracting the pixels that are shared by the projected line and the silhouette of the object, which reveals a trimmed line. Finally, the trimmed 2D line is projected back to the 3D space to replace the old pillar, resulting in a trimmed new pillar. The same trimming process is repeated for all pillars and for all snapshots.

3D model generation

The trimmed pillars approximate the actual objects by straight lines. Despite the fact that the trimmed pillars represent a good approximation for the real objects, the pillars alone are neither intuitive nor computationally efficient for visualisation, due to the fact of non-solid geometry. Moreover, the modelled shapes need to be compatible with other 3D models in the Wise-ShopFloor [7]; therefore 3D models are preferable. This can be achieved by surface triangulation to build 3D models based on the end points of the trimmed pillars.

The pillars are first divided into several sections according to their alignments with each other. This division localises and simplifies the triangulation process to create uniform and correct surface triangle patches. The triangulation process is divided into four steps to create: (1) the top surface, (2) the bottom surface, (3) the sides, and (4) other triangle patches to cover the gaps that may have been left to create a closed surface of one model (or one object).

During the surface triangulation, three neighbouring end points are chosen from the pillars, whose order is crucial to the direction of the generated triangle patch as it determines where the surface normal will point to. This surface normal is used in the Wise-ShopFloor to define the visibility feature of the triangle patch.

5. CASE STUDY

The aforementioned modules and functions are implemented in Java and integrated to the Wise-ShopFloor. The application server runs in a computer with Intel Core i5 2.8 GHz processor and 4 GB RAM running on Windows Vista Home Basics operating system. Four simple objects are chosen for a case study to test the feasibility of the developed system. Figure 7 depicts the experimental results of the main stages of the image processing and 3D model generation.

Fig. 7. Experimental results of case study
Note that the black area in the first colour image is to cover a partial robot gripper automatically during the image processing. Its appearance in the captured image is due to a constrained mounting position of the camera; this position prevents the interference with the end-effector. In this case study, seven snapshots were taken for 3D models generation. Increasing the number of snapshots from difference angles could improve the quality of the models. But on the other hand, the processing time would be longer.

As for the remote assembly, a simplified scenario is tested. A remote operator “assembles” the models one on top of another to create a stack using a 3D robot model. Simultaneously, the real robot moves the actual objects accordingly. During the assembly, the needed control commands are transmitted from the virtual robot to the real robot, automatically with no extra robot programming.

Additional features have been developed to speed up the assembly operations by introducing automatic tools to pick and place the objects. By analysing the top view snapshot, the system is able to calculate the centres of gravity for the objects and provide the suitable robot’s trajectories to pick them.

The system has been tested for ten times and the average computation time for each processing step was calculated. Figure 8 shows the percentage of the processing time of each step with respect to the total time needed for image processing. It is found that although the system can process an image in ±4 s, the silhouette labelling consumes 36.4% of the total processing time. The reason is due to the nature of the labelling algorithm which scans the image pixel by pixel and examines all its neighbouring pixels.

Fig.8. Comparison of processing times

![Fig.8](image)

Fig.9. Modelling error vs. number of images processed

![Fig.9](image)
Seven snapshots were used for 3D modelling. The results of the pillar trimming in each snapshot are recorded. Figure 9 shows the accuracy which calculated by comparing the actual pillar height of an object with that of its 3D model after processing each snapshot. The accuracy of pillar trimming is quite high as the error converges quickly to a small value after snapshot 7 in about 24 s.

6. CONCLUSIONS

This paper presents a novel approach for remote assembly where an off-site operator can monitor and control a real industrial robot in a virtual robotic environment. The 3D models of the components to be assembled are constructed based on the snapshots of the components captured by a robot-mounted IP camera. Due to the real-time network speed constraint, the camera cannot be used during assembly – thus this vision-assisted and 3D model-based approach.

From the results of the case study, it is clear that our system can generate a set of 3D models in 24 s based on seven snapshots. The efficiency can be improved by parallel image processing during the travel time when the robot is moving for the next snapshot. Additional snapshots can be utilised to improve the modelling accuracy of more complex objects.

The accuracy of a 3D model may be attributed to: (1) camera calibration, (2) camera resolution, and (3) the distance between an object and the camera.

Targeting these practical issues, our future work will be centred on accuracy, repeatability, user-friendliness, and efficiency. More tests of realistic component assembly are also planned.

REFERENCES


Minimising energy consumption for robot arm movement

Bernard Schmidt¹, Abdullah Mohammed², Lihui Wang¹,²

¹ Virtual Systems Research Centre, University of Skövde, Sweden
² KTH - Royal Institute of Technology, Department of Production Engineering, Sweden, bernard.schmidt@his.se

ABSTRACT

Optimising the energy consumption of robot movements has been one of the main focuses for most of today’s robotic simulation software. This optimisation is based on minimising a robot’s joints movements. In many cases, it does not take into consideration the dynamic features. Therefore, reducing energy consumption is still a challenging task and it involves studying the robot’s kinematic and dynamic models together with application requirements. The primary focus of this research is to develop a model to reduce the energy consumption in robotic applications. An energy optimisation module reported in this paper was developed using Matlab®. By solving the kinematics and dynamics equations of the robot, the module is able to choose the solution with the minimum energy consumption of the robot’s movements. Moreover, placement of the targets in robot’s working area that minimise the energy consumption can be suggested. The results show the value of the reported approach as a tool for energy efficient robot path planning.

KEYWORDS: energy consumption, optimisation, path planning, robot dynamics.

1. INTRODUCTION

Since the last decade many industrial countries witnessed an increase in prices of both electricity and fuel. Recent statistics showed that one of the large consumers of energy is the manufacturing industry. The majority of the energy is usually consumed by robots in the manufacturing industry. In addition, the optimal usage of energy in robots plays an important role in minimising CO₂ emission in the production stage of a product’s lifecycle.

Industrial robots are often described as unsustainable machines that demand a high level of energy consumption. However, those robots provide precision; strength and sensing capabilities which can produce high quality end products [1]. Therefore, the robotic energy consumption became a major target for many research groups and robot manufacturers. Some of the researchers focused on defining tools to measure and analyse the robot’s energy consumption. For example, the work reported by [2] contributes to identifying energy efficient strategies in robotic applications. Others like [3] summarised different methods for energy-
efficient use of common industrial robots. On the other hand, many researchers proposed robotic trajectory planning approaches with optimised time and energy consumptions [4-7].

However, these approaches put a high priority on minimising the movement time of a robot, which may not necessarily minimise the energy consumption. The total energy consumed by the robot is usually affected by the required torque on each joint and inertia tensors of each link. Other researchers focused on optimising the robotic manufacturing system as a whole [8, 9]. Despite the aforementioned efforts, optimising robot energy remains a challenge and requires further investigations.

In this paper, we introduce an approach to minimising the energy consumption of an industrial robot’s movements. It is achieved by developing a module that optimises the robot’s joint configuration to follow a certain trajectory defined by an operator. We evaluated the performance of the system by comparing the results of the optimisation module with those of a commercial simulation software of the robot. Three scenarios have been used to study the results of the developed module; with and without a payload at the robot’s end-effector.

2. SYSTEM OVERVIEW

The developed system consists of five modules: the first one is responsible for preparing the desired trajectory of the robot with the desired velocity and acceleration. The second module is dedicated to solve the inverse kinematics of the trajectory. The third module calculates the forward and backward recursions to solve the inverse dynamics of the desired task. The energy consumption is calculated at the fourth module introducing a set of robot’s joints configurations to be optimised in the fifth module. Figure 1 illustrates the developed modules implemented in MATLAB® software.

![Fig.1. Overview of the developed system.](image)

3. METHODOLOGY

The implementation of the system starts by building the mathematical model of the robot to solve the kinematics and dynamics issues and calculate the robot’s feasible joint
configurations. The system then optimises the configurations to find the best one based on energy consumption.

Both the kinematic and dynamic models of the robot have been developed to analyse the forces and torques on the robot’s joints during the motion of the robot. As an example, the ABB IRB140 industrial robot has been chosen.

3.1. Denavit-Hartenberg notation

This development requires first establishing the kinematic parameters of the robot and assigning the joints frames. The definitions of the parameters are quite crucial for the kinematics and dynamics calculations in the upcoming steps. The descriptions of the robot joint’s frames were based on the Denavit-Hartenberg (D-H) notation. Starting from the base of the robot, the frames have been assigned to each joint. It is achieved by taking into account that the joint’s rotation is always around the z-axis of the assigned frame. Figure 2 shows application of the D-H notation on the robot model. Frame marked as 0 is assigned to the robot base (link 0) and is used as global/world frame. Frames 1, 5, 6 create the robot wrist and their origin is located in the same place, called wrist centre $P_w=[P_{xw} P_{yw} P_{zw}]^T$. Parameters a1-a4 and d1,d2 are used to describe translation between frames in applied D-H notation.

$$T_{0}^6T(\theta_1…\theta_6)=T(\theta_1)\cdot T(\theta_2)\cdot T(\theta_3)\cdot T(\theta_4)\cdot T(\theta_5)\cdot T(\theta_6) = [TCP^0R \ 0 \ 1]$$ (1)

3.2. Forward kinematics

The forward kinematics is performed to calculate the transformation matrix that describes the location and orientation of the end-effector of the robot with respect to its base. With the help of the D-H parameters the matrix is found by first calculating the transformation matrices of each joint with respect to the previous ones, and then multiplying those matrices to find the transformation matrix of the end-effector with respect to the base (1).
where: $^iJ^T$ is transformation matrix between link $i$ and $j$. $^6TCPT$ is transformation applied by tool attached to link6 of the robot. $^0TCP\ R$ and $P$ are respectively rotation matrix and translation vector. In forward kinematics transformations are functions of joint values.

### 3.3. Inverse Kinematics

Solving the inverse kinematics of the robot is an important step for calculating the inverse dynamics of the robot in the upcoming steps. It is archived under the assumption that the first three joints are responsible for the end-effector’s position and the last three are responsible for its orientation. With given position and orientation of the target location, the position of the wrist is calculated.

The calculation of the first three joints ($\theta_1$, $\theta_2$, $\theta_3$) is made geometrically, whereas the last three joints ($\theta_4$, $\theta_5$, $\theta_6$) are calculated analytically.

The process starts by solving the configuration of the first joint $\theta_1$ (2). It is achieved by looking at the robot from the top view as seen in Fig. 3; which clearly shows that the angle changes the robot wrist position in the x-y plane. Using atan2 function defined by (3) returns angle value within a range: (-π:π]. For joint 1 there are two possible values, that correspond to the configurations that robot arm is reaching the target toward or backward.

![Fig.3. The first joint angle projection on the x-y plane.](image)

$$\theta_1 = \begin{cases} \arctan \left( \frac{y}{x} \right) & x > 0 \\ \arctan \left( \frac{y}{x} \right) + \pi & y \geq 0, x < 0 \\ \arctan \left( \frac{y}{x} \right) - \pi & y < 0, x < 0 \\ +\pi/2 & y > 0, x = 0 \\ -\pi/2 & y < 0, x = 0 \\ \text{undefined} & y = 0, x = 0 \end{cases}$$

Then the values of the next two angles are calculated by looking at the plane Z0-XY0 formed by link2 and link3 of the robot in Fig. 4. It starts by calculating $\cos \theta_3$ value using equation (4) which is used later to calculate joint 3 value $\theta_3$ by implementing equation (5). Two values correspond to two robot arm configurations 1, 2 as shown in Fig. 4.
\[
\cos \theta_3 = -\frac{d_2^2 + (a_2 + a_3)^2 - \left(P_{zw} - d_1\right)^2 + r^2}{2 \left(a_2 + a_3\right) d_2}
\]

\[
r = \sqrt{\left(P_{xw} - a_1 \cos \theta_1\right)^2 + \left(P_{yw} - a_1 \sin \theta_1\right)^2}
\]

\[
\theta_3 = \begin{aligned}
&\arctan 2 \left(\sqrt{1 - \cos^2 \theta_3}, \cos \theta_3\right) \\
&\arctan 2 \left(-\sqrt{1 - \cos^2 \theta_3}, \cos \theta_3\right)
\end{aligned}
\]

The process then continues to calculate joint 2 value \( \theta_2 \) according to equation (6) for each value of \( \theta_3 \).

\[
\theta_2 = \beta - \alpha
\]

\[
\beta = \arctan \frac{P_{zw} - d_1}{r}
\]

\[
\alpha = \arctan \frac{(a_2 + a_3) \sin \theta_3}{d_2 + (a_2 + a_3) \cos \theta_3}
\]

When the first three joint values are obtained in all configurations, with forward kinematics the orientation of the wrist can be calculated (7) and the rotation that have to be applied in the wrist (8).

\[
^0_w R = ^0_3 R(\theta_1, \theta_2, \theta_3) ^3_4 R(\theta_4 = 0)
\]

\[
^R_{z \times z}(\theta_4, \theta_5, \theta_6) = ^0_w R^T \cdot ^0_C P \cdot ^C R_{TCP}^T \cdot ^T \cdot ^R_w = \begin{bmatrix}
    r_{11} & r_{12} & r_{13} \\
    r_{21} & r_{22} & r_{23} \\
    r_{31} & r_{32} & r_{33}
\end{bmatrix} = R_Z(\alpha) R_X(\beta) R_Z(\gamma)
\]
When Euler angles $\alpha$, $\beta$, $\gamma$ in configuration ZXZ are used to express rotation $R_{ZXZ}$, the dependency between these angles and joint values is given in (9). To obtain Euler angles from known rotation matrix calculations follow (10-11).

$$\begin{align*}
\theta_4 &= \alpha, \quad \theta_5 = -\beta, \quad \theta_6 = \lambda \\
\beta &= \begin{cases}
\tan^{-1} \left( \frac{\sqrt{r_{31}^2 + r_{32}^2}}{r_{33}} \right), & r_{33} \\
\tan^{-1} \left( \frac{-\sqrt{r_{31}^2 + r_{32}^2}}{r_{33}} \right)
\end{cases} \\
\alpha &= \tan^{-1} \left( \frac{r_{13} - r_{23} \sin \beta}{r_{31} \sin \beta} \right) \\
\gamma &= \tan^{-1} \left( \frac{r_{31}}{r_{32} \sin \beta} \right)
\end{align*}$$

(9) (10) (11)

For each configuration of the first 3 joints, two possible configurations (10) in the wrist can be obtained. That gives in total eight possible configurations of joint values from inverse kinematics.

### 3.4. Inverse dynamics

Recursive Newton-Euler Algorithm (RNEA) [10] is used to calculate the necessary parameters for the energy consumption calculation. To start using the method, the inertial tensor matrixes values need to be obtained first. Therefore the 3D CAD model of the robot is used with the help of SolidWorks® software to calculate the inertial data. Calculations have been made with assumption that the mass of each robot link is distributed equally through the volume of the link.

Target path is interpolated with small time intervals. For each point on the path inverse kinematics is performed to obtain joint positions, velocities and accelerations for all possible configurations. After that, forward-backward recursion algorithm is applied.

**Forward recursion**

This process is used to obtain the linear and angular motion of each link of the robot. It starts from the first link to the last one $N=6$, and with assumption that robot base (link0) is not moving, the initial conditions; angular velocity $\omega_0$ and acceleration $\alpha_0$ together with linear accelerations $a_0$, $a_0$ are all set to be 0.

Then the angular velocity $\omega_i$ and acceleration $\alpha_i$ for each link $i$ as well as linear accelerations at the end $a_i$ and the gravity centre $a_{ci}$ of the link are also calculated using equations (12).

$$\begin{align*}
\omega_i &= \dot{^i_i R^T} \cdot \omega_{i-1} + z_i \cdot \ddot{\theta}_i \\
\alpha_i &= \dot{^i_i R^T} \cdot \alpha_{i-1} + z_i \cdot \ddot{\theta}_i + \omega_i \times z_i \cdot \dot{\theta}_i \\
a_i &= \dot{^i_i R^T} \cdot a_{i-1} + \dot{\omega}_i \times r_{i-1,i} + \omega_i \times (\omega_i \times r_{i-1,i}) \\
a_{ci} &= \dot{^i_i R^T} \cdot a_{i-1} + \dot{\omega}_i \times r_{i-1,ci} + \omega_i \times (\omega_i \times r_{i-1,ci})
\end{align*}$$

(12)
**Backward recursion**

The forces and torques affecting the joints of the robot are calculated in this step. The process starts from the last link $N$ and ends at the first link. The angular velocity and acceleration values which are obtained in the previous step are used in the current stage. First, the gravity vector $g_0$ is expressed in frame for each link using equation (13).

$$g_i = iR^T \cdot g_0$$  \hspace{1cm} (13)

Then the forces $f_i$ and torques $\tau_i$ are calculated for each link using equations (14). At this stage, inputs are external forces and torques $f_{N+1}$, $\tau_{N+1}$ applied to the robot end effector.

$$f_i = iR \cdot f_{i+1} + m_i(a_{ci} - g_i)$$
$$\tau_i = iR \cdot \tau_{i+N} - f_i \times r_{i-1,ci} + iR \cdot f_{i+1} \times r_{i,ci} + \omega_i \times (I_i \cdot \omega_i) + I_i \cdot a_i$$  \hspace{1cm} (14)

### 3.5. Energy consumption calculating

The calculation of the power consumption is performed by two steps. First, the power consumed in each joint during a specific time interval $k$ is found, and then the total power consumption at that interval is calculated by summing all joints power consumptions as shown in equation (15). For short time intervals, it can be assumed that torque and angular velocity are constant.

$$P(k)_i = (\tau_i(k) \cdot \dot{\theta}_i(k))$$
$$P(k) = \sum_{i=1}^n P_i(k)$$  \hspace{1cm} (15)

The energy consumption is then calculated in equation (16), where $dt_k$ is the length of $k^{th}$ time interval of the path divided into $M$ intervals.

$$E = \int_{t_0}^{t_M} P(t) dt \approx \sum_{k=0}^M P(k) \cdot dt_k$$  \hspace{1cm} (16)

### 3.6. Energy optimisation

The developed system is designed to choose the joint configuration that requires the lowest energy consumption. Since the number of possible configurations is limited to a few solutions, an optimisation algorithm is implemented to analyse the consumption in each configuration in iterative way through the whole target path and chooses the best one.

### 4. CASE STUDY

The validity of the optimisation module was examined using three scenarios: straight line, square path and different locations. These scenarios are explained in detail in the following sections.
4.1. **Straight line scenario**

In this case, a straight line path beginning at (596, 0, 662) and ending at (596, -211, 580) mm was used as a desired trajectory for the end-effector of the robot. Two experiments have been done to evaluate the performance of the module, with and without payload; both in the robot simulation software (RobotStudio®) and in the optimisation module. Table 1 shows the results from the optimisation module compared with the RobotStudio® ones.

### Table 1. Straight line scenario results from the optimisation module

<table>
<thead>
<tr>
<th>Start conf.</th>
<th>Start joint values</th>
<th>End conf.</th>
<th>End joint values</th>
<th>Energy consumption [J]</th>
<th>Payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0,0,x)</td>
<td>0</td>
<td>(-1,0,0,x)</td>
<td>-21.29</td>
<td>19.6</td>
<td>0 kg</td>
</tr>
<tr>
<td></td>
<td>-13.38</td>
<td></td>
<td>-22.27</td>
<td>6.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.42</td>
<td></td>
<td>-55.28</td>
<td>56.72</td>
<td>6 kg</td>
</tr>
<tr>
<td></td>
<td>5.95</td>
<td></td>
<td>52.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2. **Square path scenario**

A square shaped path is used as another case to evaluate the developed system. The system is again tested without and with payload. Table 2 and Table 3 show the results from the optimisation module and those from RobotStudio®, respectively. The differences in energy consumption calculated by the implemented module and RobotStudio® are caused by the simplification of the used dynamic model.

### Table 2. Square path scenario results from the optimisation module

<table>
<thead>
<tr>
<th>Corners’ joint values</th>
<th>Energy consumption [J]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>-21.69</td>
<td>-26.10</td>
</tr>
<tr>
<td>-37.83</td>
<td>-27.45</td>
</tr>
<tr>
<td>-4.49</td>
<td>-22.54</td>
</tr>
<tr>
<td>-30.56</td>
<td>-32.60</td>
</tr>
<tr>
<td>46.61</td>
<td>54.73</td>
</tr>
<tr>
<td>22.08</td>
<td>20.27</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 3. Square path scenario results from RobotStudio

<table>
<thead>
<tr>
<th>Joint configuration</th>
<th>Energy consumption [J]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>without load</td>
</tr>
<tr>
<td>(-1,1,0,x)</td>
<td>80.5</td>
</tr>
<tr>
<td>(1,1,0,x)</td>
<td>150.0</td>
</tr>
<tr>
<td>(-1,-1,1,x)</td>
<td>104.9</td>
</tr>
</tbody>
</table>
Optimal trajectory for the square path scenario with payload is presented in Fig. 5, where joint values $\theta_i$ ($i = 1, \ldots, 6$) are obtained through inverse kinematics, joint velocity $\omega_i$ and acceleration $\alpha_i$ are calculated by differentiation, and torques $\tau_i$ are obtained with inverse dynamics calculations.

4.3. Different locations scenario

In this scenario square horizontal path is placed in different locations in robot workspace. For each location the minimal energy consumption for optimal configuration is calculated.

Figure 6 presents how the energy consumption is related to the location of the target path in the robotic space. White points represent the locations where the task cannot be performed. Data are visualised for slices $x=0$, $y=0$ and $z=0$. 


5. CONCLUSIONS

It is important to optimise the energy consumption for the economic aspect as well as to minimise the potential environmental impact. The same target can be reached by the robot arm with different configurations and each configuration needs different power consumption. Our module addresses this issue by selecting the optimal configuration to perform the task.

As presented in the second case, the energy consumption for performing the task depends also on the location where the target path is defined. In this case, the optimisation of the location of the path to minimise the energy consumption seems to be required to ensure energy saving.

Our future work will focus on the dynamic model of the robot, as it needs to be expanded to consider friction and the dynamics of the actuators that include motors, gears and couplings. The current implementation has a modular structure, and it can be easily updated with more accurate model as well as extended for other types of robots. Moreover, validation of energy consumption model with actual energy consumption through real experiment and measurements need to be performed. Furthermore, the developed energy optimisation module can be extended in the future to consider the orientation of the target path.

ACKNOWLEDGEMENTS

The authors of this paper would like to thank Andrea Arzamendi Moñux and Miren Marticorena Lizaso for their partial contributions.

REFERENCES

Detection of Battery Screwdriver’s Optimal Working Regimes

Ivans Grinevich, Natalija Mozga, Guntis Strautmanis, Oskars Lininsh

RTU Faculty of Transport and Mechanical Engineering
natalija.mozga@inbox.lv

ABSTRACT
In given article the screwdriver’s electric power consumption studies depending on the corresponding fixed threaded joint’s assembly time are made. So far there are no recommendations from the tool manufactures for screwdriver’s optimal operating modes (the conclusion is made based on the attached instructions which provide instrument manufacturers), when evaluating this aspect taking into account the electrical power consumption and assembly time for different types of fixed threaded joints (wood, metal, plastic, etc.) and also taking into account obtainable tightening moments. In the examined literature sources which deal with the opportunities concerning the optimization of the automated assembly processes of fixed threaded joints, not enough attention is paid to the reduction of the instrument’s energy consumption which is one of the main tasks in the industrial process because it directly impacts production costs.

KEYWORDS: optimization, assembly time, threaded joints, electric power consumption

1. INTRODUCTION

In nowadays automated manufacturing the question concerning electric power consumption reduction is of great interest, and it leaves its impact on the final product cost. One of the options for reducing consumption of electric power is an efficient screwdriver’s use. In the literature [2, 3, 4] is described the power estimation addiction from the beginning of speed and load torque, but there is no information about the energy dependence from the initial set rotation speed. This is due to the fact that the electric motor is usually used at a constant or slowly changing mode.

Knowing what power the electric motor develops and each regime’s time it is possible to determine the electric power consumption. Fixed threaded joint’s assembly time consists of a rotor head’s run-time, screwing time, tightening time and reaction time till the start button is released (ratchet mechanism’s operating mode).

In the literature sources [5] are given the run-time and acceleration time formulas, but the impact on overall power consumption depending on this time is not viewed. This is due to the fact that the electric motor is usually operated several tenths of minutes, minutes or even hours and on the total energy consumption’s background acceleration time impact is very
small. By running the screwdriver often with large initial set speed for a short period of time until a few seconds, the run-time energy consumption can take up to 80% of total electric power consumption throughout all operating period [1].

2. METHODS USED IN THE RESEARCH

The experiments show that it is possible to find screwdriver’s optimal mode of operation in order to prolong its life time, reduce electric power consumption and improve the quality of threaded joint.

For practical experiments in order to join two metallic plates with the size of 80x70x15 widely used 40 mm standard length bolts were selected: M5, M6, M8 and M10 with threading pitch and nominal diameters after standards ISO DIN 13. Based on made calculations and making round theoretically calculated allowable tightening torque values, we obtain: M5- 4 Nm, M6- 7 Nm, M8- 15 Nm and M10- 30 Nm.

The necessary hardware and software to process the measurements and data received in the course of the experiment are showed in the block diagram of the equipment (Fig. 1a).

The equipment consists of a pulse-width modulator (PWM), which is operated by a rotary switch and a potentiometer. The power supply of the pulse-width modulator is provided from the electrical power network of 220V. The electric motor of a screwdriver is operated from the accumulator battery, where the operating voltage and electric power are controlled by the pulse-width modulator. There is a resistor with the resistance of 0.005Ω to determine the consumed current connected in series between the accumulator battery and the PWM, whose voltage is connected to the second channel of the Pico-Scope 2205 dual-channel PC Oscilloscope. The voltage of the accumulator battery is connected to the first channel of the PC Oscilloscope. The Pico-Scope 2205 is connected to the computer via the USB port. The developed equipment (Fig. 1b) consists of the electric screwdriver by Bosch GSR 14.4 VE-2 13, which is fastened to the moving element 11 of the stand.

The stand is secured on the metal plate 1 of 200*300*20 mm. There is a c-clamp 3 for securing of the metal plates 4, where a screw-bolt is inserted, fastened to the base of the stand 2. The screw-bolt is fastened with a lock, which rests against the c-clamp. The Bosch screwdriver motor is operated by a control unit 6. The Bosch electric screwdriver is moved in a vertical direction with a handle 12. The dead load of 50 N of the moving element 11 of the stand has an effect on the nut during the experiment. The required sliding moment of the clutch was set with the rotary switch 9. The fastening period was set with the dynamometric wrench of an indicator type (TOHNICHI DB50N with ¼” square according to the DIN 3120, the high accuracy +/- 3% according to the DIN/ISO) after the unscrewing moment. The USB Oscilloscope (Pico-Scope 2205) will be used to measure the output.

The following corresponding parameters will be determined for the selected screw-bolts M5, M6, M8 and M10, using the Pico-Scope 6 PC Oscilloscope during the period of the conducted experiment: output (P, W) and assemblage period (t, s) that are used to determine the electric power consumption. There are 10 measurements taken in the selected measurement range of rotations (from 300 min⁻¹ till 1500 min⁻¹) at every rotation value. The measurement parameters are entered in the table if the practically obtained fastening period (determined after the unscrewing moment with the dynamometric wrench Tohnichi of the indicator type) does not differ for more than 10% of the theoretically estimated value of the fastening period. The initial rotation with the value 1500 min⁻¹ is not used for the screw-bolt M5 because the fastening period exceeds the permissible value at the lowest sliding position of the screwdriver’s clutch. It is so because in the dynamic system the fastening period of a thread connection depends not only on the sliding over moment of the screwdriver’s clutch (it
is set by turning the rotary switch of the screwdriver) but also on an additional moment, which comes into existence from the kinetic energy of the rotor head and nut. According to the kinetic energy theorem, the given kinetic energy turns into additional fastening period. The estimated moment for the screw-bolt M10 was provided only at 1500 min$^{-1}$ and the maximum sliding moment of the screwdriver’s clutch (the 25$^{th}$ position of the screwdriver).

![Diagram of Measuring Equipment](image)

Fig. 1. Measuring equipment: a) block diagram of the equipment; b) collective view of the equipment

The values of average energy consumption for every screw-bolt depending on the initial set rotational speed and assembly time are showed in Fig. 2. It is obvious that the screwdriver has such operational modes where it is possible to optimize the energy consumption after the assembly period (initial set rotations). Viewing the consumable power, distributing the main points of a parabola on the sample of the screw-bolt M6 (Fig. 2 - points 1, 2 and 3), each of these processes can be viewed separately (for other screw-bolts these processes are identical with an exception of M10, where fastening is possible only at maximum initial set rotations). It can be concluded that the highest energy consumption at the maximum initial set rotations is related with switching a comparator on, due to some small rotations, the screwdriver’s clutch cannot slide providing the required moment, therefore the electric current is increased until the clutch slides (Fig. 3 stage d). Similarly higher energy consumption occurs at maximum initial rotations, which is related with large electric power consumption to start running and higher losses of electric power (Fig. 5 stage a). The optimal operating mode of the screwdriver according to the electric power consumption is showed in Fig. 4, providing the required fastening moment (7 Nm) for the rigid connection of the screw-bolt M6.
Fig. 2. The consumption’s average values of electric power for corresponding bolts taking into account: a) assembly time and tightening torque; b) initial set rotation speed and tightening torque

Fig. 3. 1st point of the curve. Power consumption for the bolt M6 by the initial set rotation speed $300 \text{ min}^{-1}$

(a – run-out power, b – motor power when the set rotation is reached, c – power after the nut’s tightening, but before the comparator’s work, because the set operating level is not reached yet, d – power when the comparator has worked, e – power when the comparator is turned off, as the engine picked up rotation speed, and that’s why the current which is proportional to the voltage when the comparator starts working decreased)
Fig. 4. 2nd point of the curve. Power consumption for the bolt M6 by the initial set rotation speed 900 min\(^{-1}\)

(a – run-out power, b – motor power when the set rotation is reached, c – power when the nut has been tightened, but the start button has not yet been released)

Fig. 5. 3rd point of the curve. Power consumption for the bolt M6 by the initial set rotation speed 1500 min\(^{-1}\)

(a – run-out power, b – motor power when the set rotation is reached, c – power when the nut has been tightened, but the start button has not yet been released)

3. CONCLUSIONS

From the obtained results it is clear that creating a fixed threaded joint connection at low speed (300 min\(^{-1}\) and 500 min\(^{-1}\)) consumable energy is independent of the selected tightening moment because when the comparator is switching on, the current is close to its the maximal value. Also at high set rotation speed the energy consumption is not dependent from the selected tightening moment due to the high kinetic energy of the screwdriver’s rotation parts.

On the basis of the experimental results it can be concluded that the screwdriver has such operational modes that permit performing of the electric power optimization depending on the assembly period (assembly time is the time when the screwdriver is working and screwing the nut). The most economical mode of a screwdriver is at such minimum set rotations requiring no additional electric power (involving of the comparator), providing fastening of a nut with the corresponding moment.
Taking into account that in the computerized assembly usually using the maximum rotations of a rotor head of the screwdriver, performing optimization of the energy consumption after the assembly period, decreasing is obtained:

1) For the bolt M5x40 with tightening torque 4 Nm is till 16.4%;
2) For the bolt M6x40 with tightening torque 7 Nm is till 31.4%;
3) For the bolt M8x40 with tightening torque 15 Nm is till 35.5%.

ACKNOWLEDGEMENTS

The work has been supported by the Project „Production and Metrology of thin coatings aimed at the reduction of sliding friction surface wear and friction coefficient” Nr. TOP/12-03.

REFERENCES

Supply chain collaboration on the competitiveness of Basque Country manufacturing companies.

Alberto de la Calle, Esther Alvarez, Arantza Múgica, Olatz Ukar
Universidad de Deusto, Bilbao,
acalle@deusto.es

ABSTRACT

Globalization, demand and technological uncertainties are moving the competition battlefield from single companies to entire value chains. Thus, collaborative strategies among supply chain partners are getting superior relevance. This new scenario should be accompanied by an understandable integrated model that considers all these facts. This paper tries to shed light on the collaboration among supply chain partners. An integrated value chain framework is proposed. This research focuses on the validation of this framework and attempts to understand the current situation of collaboration in one Spanish region. Based on the feedback of 91 manufacturing companies of the Basque Country we measured the level of involvement of customers and suppliers across 14 different supply chain processes. In addition we also gathered the reasons that motivate them to collaborate with customers or/and with suppliers. The information is analysed using the proposed integrated value framework.

KEYWORDS: Customer and Supplier involvement, value, alignment.

1. INTRODUCTION

The business landscape has been altered as the globalization advances. The availability of high-quality products at competitive prices is easier in a borderless world. Easy imitation and low cost strategies are other strategies that globalization makes appear. Due to this fact, firms coexist in a speedy and highly changeable environment rooted in constant technology development, shorter product life cycles, and changing fashion and customer interests.

In order to remain competitive, companies strive for improving their value offers: goods and services, at reasonable prices. It forces companies to review periodically the development of their current strategies, looking for the development of sustainable competitive advantages.

Nowadays competitiveness is moving toward a value added industry. Focusing on manufacturing firms, the evolution of the value propositions has moved from a product focus (price, quality and availability) to a more holistic product and service approach. The key to attainment of market leadership is to raise the bar within the organisation through the creation of value at all levels. In this sense, many authors argue that all value creating parameters exceed an isolated company’s responsibilities, moving the competition battlefield to the entire
supply chains that deliver them [1]. Both academics and practitioners highlight the relevance of collaboration throughout the value chains.

However this is not an easy task. Companies still view the supply chain collaborative initiatives as a role in reducing costs and sustaining revenues [2]. Although cost reduction used to be a prime motivator for supply chain collaboration, it entails much more, i.e. enhance customer service, increase flexibility and increase the success of launching new products [3].

In spite of these problems, globalization, demand and technology uncertainties make managers going beyond their companies. The involvement of customers and suppliers is on the rise; however, failure rates also seem to remain high [4].

Enterprises need to understand how the entire value chain works and how it can be improved for accomplishing, if not going beyond, the customer expectations. This study contributes to supply chain management literature in three ways. Firstly, the study provides an integrated value chain framework. It considers not only the value generated by a single company but by the whole value chain. This makes explicit the common objective of generating product and service attractiveness. Secondly it contributes to understand the actual involvement level of customers and suppliers in different value chain processes. It is also evaluated the changes depending on different factors i.e. size, sector and position in the value chain. Finally it contributes to a deeper knowledge about what motivates companies to launch collaborative initiatives with the value chain members.

Based on the above considerations, we set out the following research questions (RQ):

- RQ1. Which framework should be considered in order to create and deliver value for overcoming customer expectations?
- RQ2. Which is the involvement level of customers and suppliers across different supply chain processes?
- RQ3. Which objectives are pursued by companies when collaborate with their customers and suppliers?

In order to answer these questions, we build, based on extensive literature review, an integrated value framework from the value chain perspective and empirically tests on field work developed among manufacturing companies from the Basque Country.

The paper is organised as follows. Section 2 presents the theoretical background for supply chain collaboration and value chain as perceived in the literature. This analysis led to the definition of an integrated value chain framework. This section responds to the RQ1. Section 3 describes the research methodology. Section 4 provides the description and results of the field work, trying to answer to the RQ2 and RQ3. Finally Section 5 presents the conclusions and future research developments.

2. THEORETICAL BACKGROUND

2.1. Supply chain collaboration

From its logistics origin, the supply chain management becomes into a broaden approach by which companies face increasing competitiveness [5]. Supply chain management seeks to enhance competitive performance by linking closely integrating internal functions within a company with the external operations of suppliers and customers [6]. Thus, supply chain management concept is not only focused on improving the product and information flows but on the functional areas that contribute to create and deliver value to the customers. Indeed, supply chain collaboration is advocated as the key to creating value in supply chain management [7].
Several research works also discuss that companies must look beyond their organisational boundaries and evaluate how the resources and capabilities of suppliers and customers can be used for creating exceptional value [8], [9]. However, the existing research on supply chain collaboration does not follow common definitions and dimensions [10] which led to inconsistent findings about the relationship between supply chain collaboration and performance.

Collaboration among business partners may vary depending upon the organizations’ prospective role in the supply chain [11], their current position in the chain, the industry in which they operates or the size and annual turnover of the companies.

2.2. An integrated value framework

Most of the current discussion on Competitiveness is moving around the concept of “customer value creation”. The product characteristics, such as innovative profile, quality and functionalities, are all important but also when and how they are delivered, and how the company supports customers after service. So for many companies value is in the product-service package. Considering this perspective it is possible to differentiate the supply and demand chains in each company of the value chain.

The value chain and the value system [12] explain how the value is added to the product through the supply chain, highlighting both each company’s capabilities and the links between companies. Following the value added perspective, a prerequisite is that company assumes that value is created and delivered. Due to the new competitive paradigm in which value is in the couple product & service, a new perspective is needed. On the one hand the supply chain comprises all those processes necessary to fulfill customer demand. The demand chain, on the other hand, comprises all the processes necessary to understand, create and stimulate customer demand [13]. In addition, the influence of the access to customers’ and suppliers’ intelligence is recognized.

Sherman [14] states that the supply chain must be redesign to deftly balance, align, and seamlessly integrate demand chain in search of sustainable and competitively superior performance gains. He proposes a common framework for any business and industry, named demand management structures divided into three areas: demand creation, demand fulfilment and demand performance structures. But it misses the collaborative perspective. This framework is built in a single company environment. Complementing this approach, Walters [15] maintains that upstream and downstream processes are equally important for improving the value chain performance.

There are many difficulties in order to maintain a balance and integrated view of demand and supply chains. Some of them are related to the management approach: although supply and demand chain management are of fundamental importance for all businesses, one of them is usually prioritized in many companies [16]. Lean and agile strategies are examples of response to demand features from the demand fulfilment chain. Open innovation, co-creation, innovation networks can be examples from the demand creation chain. Linking both kinds of strategies could be a hard or even unaffordable task for many companies.

More recently, a demand-supply chain management approach (DSCM) has been introduced as a framework for developing a superior competitive advantage by coordinating the demand and supply chain processes across intra and inter enterprise [17]. Another interesting approach is the global networked circle model developed on the basis of best-in-class manufacturing companies [18].

Following the literature review, we proposed an integrated value framework (Fig.1) which considers all the mentioned issues: the value is created from the demand creation chain
and it is delivered through the demand fulfilment chain. Both chains should be aligned and measured using the demand performance chain.

This framework establishes a common and understandable reference point for adding value through collaborative initiatives. It makes clear the co-responsibility of the value chain companies in attending the requirements and needs of the final customers. This means that both demand creation and demand fulfilment chains contribute to the customer satisfaction and consequently to the business performance.

The integrated framework is based on differentiated chains and results:

- The demand creation chain is based on Sherman [14], Hilletofth [19] and Jüttner and other [20] contributions. It comprises Market intelligence, R&D and Commercialization processes.
- The demand fulfilment chain deals with the material flow including the procurement, manufacturing, delivering and return processes. Using the Supplier Chain Operations Reference model (SCOR) terminology, they are the source, make and deliver processes.
- The demand performance chain is the responsible of the alignment between both approaches (value creation and delivery) and could be described as “fit” in the strategic management literature.
- Product and Service attractiveness are the results of managing the demand creation and fulfilment chains. They are what the customer perceives. Value measurement should be customized and adapted by the value chain members in order to be aligned with the customer expectations: i.e. complaints, customer satisfaction questionnaires.

The notion that companies should manage external and internal integration to develop competitive advantages is recognized. However, the initiatives and capabilities they should develop to do so, are not clear in the literature.

### 3. RESEARCH METHODOLOGY

As we proposed a new integrated conceptual value framework, we validated it with an advisory group formed by 4 academics (from business and engineering background) and 6 practitioners (5 managers from different manufacturing companies and one consultant). We
asked for their opinion about the conceptualization and the applicability of the framework as a good tool for the awareness of co-responsibility in the creation and delivery of value to final customers. We received constructive and positive comments which allow us to follow with the next steps: to confirm the sense and usefulness of the framework compiling the collaboration reality of manufacturing companies in a specific Spanish region.

Data for the field work were obtained through personal in depth interviews and a web survey. The unique restriction was that they must be manufacturing companies with at least one production plant located in the Basque Country. The reason for selecting this region is that it is one of the most industrialized regions in Spain (in 2012, 21.7% of the GDP was generated by manufacturing industry according to the Spanish Statistical Office data). The reason for focusing only in manufacturing firms is that these companies should attend both product and service attractiveness.

Finally, 91 manufacturing companies filled out the questionnaire. 33% were obtained via personal interviews and 67% via web. The questionnaire is divided in three parts: general information about the company such as size, turnover, position of the company in the supply chain, and position of the respondent. Second part deals with the measurement of the involvement level of customers and suppliers into 14 supply chain processes. This approach was used in study carried out in India [11]. Finally, they are asked about the objectives pursued by the manufacturing companies when collaborating with both upstream and downstream the value chain.

The position occupied by the company in the supply chain was codified using automotive nomenclature: OEM (overall equipment manufacturer), Tier 1 (first level supplier of OEM: critical parts), Tier 2 (second level supplier: components supplied to Tier 1), Tier 3-4 (third and fourth level supplier) or Raw Material provider (RWprov).

### 4. DESCRIPTIVE ANALYSIS AND RESULTS:

#### 4.1. Data base description

A profile of respondents is presented in Table 1 characterized by its chain position, size based on the number of employees and industry sector. There are represented a variety of industries, such as automotive sector – Auto-(24 companies), aerospace –Aero- (13), and other metal industries (39), energy (4), paper (3), Building materials (2), Fast Moving Consumer Goods (1), Pharmaceutical (1), Electronics (1) and Glass industry (1).

<table>
<thead>
<tr>
<th>Nº of employees</th>
<th>Total</th>
<th>Industry sector</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aero</td>
</tr>
<tr>
<td>&lt;10</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>10-50</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>50-250</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>&gt;250</td>
<td>39</td>
<td>5</td>
</tr>
</tbody>
</table>

| Position in the | Total | Industry sector |
| supply chain    |       | Aero | Auto | Metal | Other |
| OEM             | 0     | 0    | 0    | 0     | 0     |
| Tier 1          | 0     | 0    | 0    | 0     | 0     |
| Tier 2          | 4     | 6    | 9    | 4     | 23    |
| Tier 3-4        | 1     | 2    | 0    | 0     | 3     |
| RWprov          | 0     | 0    | 2    | 4     | 6     |
| Total           | 5     | 21   | 37   | 28    | 91    | 13   | 24   | 39   | 15    |

#### 4.2. Validation of the INTEGRATED VALUE FRAMEWORK

Personal interviews serve us to confirm that our proposed new integrated value framework is understandable. All interviewed managers share the opinion that the framework
compiles the value essence of a company, product and service, and also take into account the collaboration among supply chain partners. They agree with the need of alignment both chains. More than desired times companies dedicate much time to face daily problems that avoid them to think more in the future. There is also a trend that comes from the past, to reduce cost, production times and inventories, forgetting that there is another important value which could be created, improved, modified, renewed or substituted, in order to be more competitive.

There are three companies which organize their businesses following a kind of similar approach. They use different names but the idea was the same: the demand management chain (value creation) they call Product Management, while the demand fulfillment chain (value delivery) is named as Order-Payment cycle.

4.3. Customer-supplier collaboration in supply chain processes

In order to understand how the companies involve of both customers and suppliers, the value chain was broken down into 15 different processes. All of them were extracted from literature review, especially from Cooper et al. [21], Sahay [11] and from the SCOR model. After the revision of the advisory group, one process was eliminated (Merchandising) and two processes were re-rewritten in order to avoid confusion or misunderstanding. Four processes are identified in demand creation chain such as: Market research, Brand building, New-product development, and Marketing & Promotions management. Eight processes belongs to demand fulfillment chain: Inventory management, Purchasing management, Manufacturing planning, Warehousing management, Distribution management, Transportation management, Order processing/fulfilment and New process development. Two processes are included under Demand performance chain: Investment management and Environmental impact management.

The respondents were asked to indicate the involvement of both their customers and suppliers in each process on a scale of 1 (low involvement) to 5 (high involvement). Crossing the means (Table 2) of each company regarding their customer and supplier involvement level we configure the following matrix (Fig.2). The analysis develop with these data contributes to the second research question (RQ2).

<table>
<thead>
<tr>
<th>Process name</th>
<th>Customer involvement</th>
<th>Supplier involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>P1 Market/Demand research</td>
<td>2.66</td>
<td>1.29</td>
</tr>
<tr>
<td>P2 Brand building</td>
<td>2.42</td>
<td>1.30</td>
</tr>
<tr>
<td>P3 New product development</td>
<td>3.27</td>
<td>1.18</td>
</tr>
<tr>
<td>P4 New processes development</td>
<td>2.52</td>
<td>1.17</td>
</tr>
<tr>
<td>P5 Marketing/Promotions Management</td>
<td>2.29</td>
<td>1.31</td>
</tr>
<tr>
<td>P6 Investment Management</td>
<td>2.15</td>
<td>1.15</td>
</tr>
<tr>
<td>P7 Inventory Management</td>
<td>2.25</td>
<td>1.20</td>
</tr>
<tr>
<td>P8 Purchasing Management</td>
<td>2.27</td>
<td>1.23</td>
</tr>
<tr>
<td>P9 Manufacturing (planning)</td>
<td>2.86</td>
<td>1.29</td>
</tr>
<tr>
<td>P10 Warehousing Management</td>
<td>2.21</td>
<td>1.21</td>
</tr>
<tr>
<td>P11 Distribution Management</td>
<td>2.28</td>
<td>1.18</td>
</tr>
<tr>
<td>P12 Order Processing/ Fulfilment</td>
<td>3.61</td>
<td>1.26</td>
</tr>
<tr>
<td>P13 Transportation Management</td>
<td>3.16</td>
<td>1.22</td>
</tr>
<tr>
<td>P14 Environmental impact management</td>
<td>2.34</td>
<td>1.24</td>
</tr>
</tbody>
</table>

SD: Standard Deviation
Considering the processes the most important from the suppliers and customers involvement perspectives is the Order processing/fulfilment (P12). The rest of the processes differ in relevance in function of customers or suppliers collaboration. For example, Manufacturing planning (P9) is ranked as 2 as supplier involvement but as 4 for customer involvement. Another example is Purchasing management (P8) which is the 3rd in importance from the suppliers point of view but the 7th for customer involvement. The processes with the highest involvement rate for both customers and suppliers are: Order processing/fulfilment (65% of respondents ranked high both customer and supplier involvement), Transportation management (55%), New product development (51%), Manufacturing planning (34%).

Marketing and investment management are seen more as internal processes. But focusing on Marketing process if there is any kind of openness, companies prefer customers (38%). When companies look for partners on investment management process there is not a big difference on involving customers or suppliers.

The analysis of data reveals that the level of involvement is low for customers and suppliers. Specifically, 54% of the respondents marked low collaboration with the value chain (Table 2). On the other side, 20% assure that customer and supplier involvement is high. A study of the collaboration in the value chain across different sectors, supply chain positions and processes shows very interesting results.

Table 3. Supplier and customer involvement level.

<table>
<thead>
<tr>
<th>Supplier involvement</th>
<th>High</th>
<th>14%</th>
<th>20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>54%</td>
<td>12%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer involvement</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
</table>

For example, according to its value chain position all the companies identified as Tier 3-4 has a low level of involvement with customers and suppliers. OEM and Tier 2 has a similar level of supplier involvement (56% low and 44% high), but differ on customer integration. Tier 2 is closer to the customers that OEM, 21% of the OEM companies have a high level of customer involvement whereas in the case of Tier 2 this percentage is up to 43%. The principal reason is that most of the Tier 2 belongs to the automotive industry and their relation with customers used to be closer than in other sectors.
Considering the company size, the mean of customer involvement level decreases as the size of the company also decrease (Fig.3). Regarding suppliers collaboration it seems there is not that relation. Firstly because the companies with less than 10 employees involve their suppliers more than the companies with 10-50 employees do. And secondly because the companies with 50-250 employees involve their suppliers more than the big companies (>250) do.

![Fig.3. Involvement level in relation to the company size (number of employees)](image)

We run an analysis of variance (ANOVA) in order to determine whether there are any significant differences between the means of the 14 processes through the three Aerospace, Automotive and Metal sectors. Focusing on customer involvement, the results show that Manufacturing planning (P9) is significantly higher in Automotive than it is in Aerospace and Metal sectors. Regarding the supplier involvement, the New product development (P3) is significantly higher in Metal than in Aerospace.

### 4.4. Objectives of collaborating with the supply chain partners

Through the questionnaire we also try to look at the objectives pursued in collaborating with customers and suppliers, trying to shed light to the third research question (RQ3). We propose eight objectives and also the “Other” option was available. Each surveyed company should check their three main objectives when collaborating upstream and their three main objectives when do so downstream the value chain. The objectives can be classified by demand creation chain (market knowledge, success in launching new products and flexibility), demand fulfilment chain (asset management, process reliability and lead time) and demand performance chain (product quality and environmental impact). The objectives when companies look for collaborating with customers and suppliers differ significantly (Fig.4).

Attending the number of times that each objective was chosen we identify three groups of objectives. The first group with the most selected objectives is formed by flexibility (adaptation to demand) and quality, the second group is lead time, new products, market knowledge and process reliability. Asset management and environmental impact belong to the last group. Putting the emphasis on differentiating upstream and downstream collaboration objectives, it is clear that the objectives dealing with demand creation are more important in the collaboration with customers than with suppliers. On the other hand, demand fulfilment objectives are linked to upstream collaboration.
5. CONCLUSIONS

In the dynamic environment of most supply chains, collaboration is a fundamental strategy for sustainable and profitable growth. This study attempts to build a framework which integrates several issues that contribute to generate new competitive advantages from a value chain perspective. That is why the proposed framework is divided into three chains. It allows companies to have a holistic view of the whole supply chain, highlighting the idea of co-responsibility in value creation and delivery processes. The demand term in the three chains has been used on purpose. It reinforces the idea that business should be managed following a customer orientation.

Some interviewed managers confirm that their businesses are organized following a similar approach: product management and order-payment cycle. Because of daily urgencies, companies usually prioritize either demand creation or demand fulfilment. If the management defines their strategies focusing only on the efficiency of the supply chain it can lead to a downward spiral and has a limited potential of success. It could lead to the commoditization of products or to a short term competitive advantages. This is a problem because customers are increasingly looking for product and service providers. For this reason we propose the demand performance chain, which should be in charge of the alignment between demand creation and fulfilment chains (RQ1).

Through the field work it has been concluded that the demand creation processes (research and development, market intelligence and commercialization) are more related to the customer involvement. On the contrary, suppliers are involved in the processes identified in the demand fulfilment chain (purchasing management, manufacturing planning, distribution, inventory management and other processes related with the material flow). However it should be taken into account that that the level of customers’ and suppliers’ involvement varies across different processes and sectors (RQ2). It could be because the complexity of the processes differs from one industry to another.

The interpretation of the objectives (RQ3) pursued by companies when collaborating with customers and suppliers go in the same direction. Companies usually look for customer collaboration when they need to improve some aspect of the demand creation chain. But if they want to develop any improvement in the demand fulfilment processes they look for collaboration with suppliers.

Although the awareness about the relevance of collaboration between value chain partners is growing, there are still many things to do. We argue that the proposed framework suits to the current competitive requirements and it can help managers to view the competitiveness from a more holistic perspective.
REFERENCES


Material Science, Additive manufacturing and Welding

- A methodology for parameter setting in the laser cladding by process simulation
  Iván Taberner, Aitzol Lamikiz, Eneko Ukar, Silvia Martínez
  Dpt. of Mechanical Engineering, ETSI de Bilbao, UPV/EHU, Spain

- A novel strategy for the incorporation of optical sensors in Fused Deposition Modeling parts
  A. Kantaros, J. Giannatsis, D. Karalekas
  Laboratory of Advanced Manufacturing Technologies & Testing, Dept. of Industrial Management & Technology, University of Piraeus, Greece

- Cooling Rate Prediction in case of Pipelines Longitudinal Welds performed by Submerged Multi-Arc Welding
  Carmen Catalina Rusu, Luigi Renato Mistodie, Elena Scute
  Department of Manufacturing Engineering, Faculty of Mechanical Engineering “Dunarea de Jos” University of Galati, Romania

- A new approach to modelling friction stir welding using the CEL method
  M. Hossfeld, E. Roos
  Materials Testing Institute (MPA), University of Stuttgart

- Modelling the mechanical deformation of nickel foils for nanoimprint lithography on double-curved surfaces
  M. R. Sonne¹, J. Cech², J. H. Hattel¹ and R. Taboryski²
  ¹Technical University of Denmark, Department of Mechanical Engineering
  ²Technical University of Denmark, Department of Micro- and Nanotechnology
A methodology for parameter setting in the laser cladding by process simulation

Iván Tabernero, Aitzol Lamikiz, Eneko Ukar, Silvia Martínez

1 Dpt. of Mechanical Engineering, ETSI de Bilbao, UPV/EHU, Alda. Urquijo s/n, Bilbao, 48013 (SPAIN)
aitzol.lamikiz@ehu.es

ABSTRACT
The laser cladding process is becoming one of the most important and emerging processes for some applications such as rapid manufacturing of complex structures and repair of high value-added components. The main benefits of this technique are high flexibility for building layer-by-layer structures and the relative good quality of the bonding between the added material and the substrate, in comparison with conventional arc welding processes or other additive techniques such as plasma spray. However, the laser cladding process involves a high number of parameters that need to be adjusted in order to obtain satisfactory results. Furthermore, these parameters are of different nature (powder flow, laser parameters and added and substrate materials properties) and the selection of these parameters is one of the main obstacles to industrialize the process.

This work presents a numerical model to fit the process without needing of testing. The model is based on the coupling of a thermal model with a powder mass flow CFD model. The model result is the estimation the cladded structure geometry.

KEYWORDS: Laser cladding, simulation, optimization, parameters

1. INTRODUCTION

Nowadays, manufacturing sectors dedicated to high added value parts are involved on research and development processes in order to be able to compete in a more and more global market. This fact makes them to study new manufacturing processes, which allow obtaining more sustainable and cost-efficient production chains. In this sense, laser cladding process is becoming of high interest for these industries. This process is based on the localized melting of a substrate where a filler material is injected [1]. This material can be in form of wire or powder, and can provide to the substrate some additional properties. Laser cladding generates clads with good mechanical properties and a strong bonding with the base material, avoiding a large heat affected zone. This characteristic makes laser cladding an alternative of processes currently used for parts coating [2]. Moreover, overlapping clads it is possible to generate layers of material which can fill damaged parts or create hard coatings on wear resistant surfaces. Moreover, these layers allow manufacturing 3D complex geometries in order to
generate near shape parts that only require a finish machining [3]. This fact supposes a great advantage for sectors such as die and mould industries, where integration of this process could reduce the high amount of waste material in form of chips that involves traditional machining. Other sector in which this process is becoming of interest is the aeronautical sector, where it is a feasible alternative to conventional welding processes such as TIG (Tungsten Inert Gas) or plasma welding.

Part manufacturing by laser cladding is supported by a CAD design and a later CAM stage that obtains a CNC program with cladding trajectories [4]. Once the part is cladded, a finishing machining step is necessary to obtain final geometry. Process parameter selection is the main problem of laser cladding because of there are more than 20 different parameters with influence on the clad geometry and quality. This drawback is currently solved by a high costly experimental work expended in parameter optimization, previous to process application [5]. Thus, high added value parts manufacturers take process tuning as a great obstacle for laser cladding implementation in their industries. Process modelling is a suitable solution to minimize the invested time in tuning work, avoiding unproductive time on the machine. Generally, laser cladding modelling involves high complexity and several model types can be found in the bibliography. On the one hand, empirical models use experimental data in order to obtain a mathematical expression, which explains process behaviour [6]. These models, even though they are valid for process behaviour studying, do not reduce the experimental costs. Then, analytical models try to solve process equations avoiding experimental tests [7]. Solution of these type of models entails a high complexity and requires a great number of assumptions which reduce their validity to the studied process window. Finally, the semi-empirical models [8], despite need a series of adjustment tests, allow to generalise the results for wide ranges of the process parameters without complex developments. In any case, most works focused on laser cladding modelling subdivide process in different steps: material flow study, thermal problem on the substrate and clad growth. Although first two steps are only related to study process behaviour regarding process parameters, it is necessary their study in order to deal clad geometry modelling, which is the main objective of laser cladding modelling. If clad geometrical characteristics are previously estimated, it is possible to optimize the offset between layers in order to maintain constant beam spot size in the cladding zone [9]. Moreover, it allows adjusting clad overlapping in order to obtain high quality structures.

This work is focused on a modular clad geometry model, which integrates powder flow and thermal field steps, resulting in a global modelling of the process. Model is based on the assumption that all powder material reaching melt pool is used to generate the clad. Therefore, a powder flow model is necessary in order to estimate the material concentration on the cladding zone and a thermal model which estimates the melt pool shape. Both, powder flow model [10] and thermal model [11] has been previously developed and validated, and an attenuation model [12] has been also added in order to estimate the amount of energy that laser beam have lost due to its interaction with powder before reaching the substrate. Global model has been validated becoming as a useful tool for laser cladding process tuning for industrial applications.

2. LASER CLADDING MODELLING

Laser cladding modelling results in a high complexity task, so, as commented, it is recommendable the study separately the different phenomena occurring during the process. Firstly, a CFD model simulates powder flow in the nozzle outlet in order to estimate powder concentration. Then, an attenuation model takes data from CFD model and estimates the net-energy onto the substrate. This energy generates the melt pool onto the part surface and a
thermal model can calculate its shape. Finally, a geometric model uses all data provided by other models and estimates the clad geometry for each time step and for each point of the programmed trajectory. Fig. 1 shows the global scheme of the purposed laser cladding model, which estimates clad geometry.

Fig. 1. Scheme of global model including all developed models

In order to simplify the model, the following assumptions are considered:

1. Powder concentration and injection time on each point are the main aspects influencing clad height growth.
2. Powder losses on melt pool area are not considered, i.e. all powder reaching melt pool is trapped and will be part of the clad.
3. Clad width is directly related to melt pool width, therefore, clad only grows on the melt pool area. Powder outside this area rebounds and is considered as lost material.
4. Powder concentration over the melt pool is obtained from a CFD model developed using FLUENT® 6.5. This model has been previously developed and validated in [10].
5. Real energy reaching the part surface is calculated by a previous attenuation model developed and validated in [11].
6. Thermal field on the substrate and melt pool shape is estimated by a thermal model previously developed and validated in [12].
2.1. Powder Flux Model Development

The powder flux model has been implemented on FLUENT© 6.5 and solves the Navier-Stokes equations for a turbulent flow using a standard $\kappa$-$\varepsilon$ approximation. The discrete phase is treated by an Euler-Lagrange approach where the continuous phase is treated as a continuous homogeneous medium and Navier-Stokes equations are solved for each time step. The discrete phase is evaluated tracing a certain amount of particles in the previously calculated fluid field, and it can exchange mass, momentum and energy with the fluid phase. The main assumptions used for the model solving are:

1. Protective and dragging gases on nozzle inlets are considered normal flow to the surface with constant velocity.
2. Powder flow is considered as a steady-state problem, i.e. powder concentration is no time dependent.
3. Particle motion is calculated by a force balance, which takes into account drag, inertial and gravity forces for each particle.
4. It can be considered that a low amount of particles within gas flow is present, so, collision between particles is neglected. Only influence of particle interaction with nozzle walls are considered in the model.
5. Great changes are not expected in the density; therefore, a segregated model with constant density was selected.
6. Particle flow influence on the continuous phase is ignored due to the low mass and concentration of the particles.
7. Experimentally adjusted Rosin-Rammler distribution is used in order to introduce variable particle size in the model.
8. It can be considered that laser radiation and other heat sources have minimum influence on powder flow, therefore, these sources has not been included in the CFD model.

Model has been validated for different materials and particle sizes both qualitatively and quantitatively. Validation results for different material on powder flow focal plane are shown in Fig. 2. It can be observe how estimated and measured powder distribution adjustment can be considered acceptable since errors are always below 10%.

![Inconel 718 - $Z_{focal}$](image1.png)
![AISI D2 - $Z_{focal}$](image2.png)

Fig. 2. CFD model validation results for focal plane. Left) Inconel 718. Right) AISI D2
2.2. Attenuation Model Development

Real energy reaching substrate is an essential input for the global model due to it creates the melt pool onto the part surface. This energy is not the same which laser source generates because laser beam interacts with powder flow before arrive at part surface. So, laser cladding model must incorporate a model capable to estimate the amount of energy attenuated by interaction with powder flow. This work presents an attenuation model based on a typical shadow model widely extended in the bibliography [13]. It is therefore a semi-empirical model that requires a previous characterization of the interaction between the laser and the material, for subsequently being generalized to a range of parameter combination of the process. The proposed model is based on the following assumptions:

1. Model considers attenuation proportional to shadow generated by particles and interaction time between laser and powder.
2. Shadow between particles is neglected because can be considered low powder concentration within gas flow.
3. All data regarding particles, such as concentration or velocity, is provided by previously developed and validated CFD model.
4. Particles are considered spherical and their projection on the plane is approximated to a circle. Rosin-Rammler particle size adjustment made for the CFD model is also used in the model.
5. Density and particle size are considered to be independent of temperature maintaining constant values.
6. It is also considered that the attenuation is produced after a plane in which the mass flow distribution profile is consolidated and takes a Gaussian form.

Attenuation model has been validated for different materials, particle sizes and laser power, showing good agreements between measured and estimated attenuation values. Figure 3 shows an example of validation test where it can be observed an estimation error below 7% for AISI D2 and 900 W laser power.

![Validation Tests - AISI D2 - 900 W](image)

Fig. 3. Attenuation model validation test for AISI D2 and 900 W power.
2.3. Thermal Model

Once attenuation model estimates the real energy that is radiated onto the surface it is possible to deal with the thermal field estimation in the substrate, in order to define the melt pool. A thermal model has been developed in Matlab® 7.8 which solves the heat transfer equation by finite difference algorithm. Thermal model is based on the following assumptions:

1. Model only takes into account heat transmission by conduction. Convection and radiation are considered by a losses parameter which also groups all possible energy losses such as reflection or plasma formation. A 0.2 value is commonly used in laser cladding processes on steel milled parts. So, the equation solved by the model takes the form (Eq. 1):

$$a \nabla^2 \theta + \left(1 - A \right) \frac{q_r}{\rho c_p} = \frac{\partial \theta}{\partial t}$$

(1)

2. Initial temperature is fixed with 298 K and heat flux from surfaces no irradiated is considered null.
3. Influence of temperature on material properties such as conductivity or density is considered introducing in the model these non-linearities. Solid state phase transformations during heating and cooling steps are considered as energy consumptions. The numerical approach treats phase transformations with an additive principle, i.e. the percent phase transformed in each simulation step ($f_i$) is an accumulation of the phase transformed in the previous step ($f_{i+1}$). The kinetic equation of Johnson-Mehl-Avrami is used to model diffusive transformation during heating, estimating the fraction of transformed material over time for isothermal conditions. To model non-isothermal conditions, the fraction of transformed material balance is done in differential time steps.

Fig. 4. Thermal model validation test for different laser powers.
2.4. Geometric Model Development

Finally, once it has been estimated both powder distribution on the surface and thermal field in the substrate, it is possible to deal with generated clad geometry modelling. Geometric model is based on a mass balance between injected material into the melt pool and generated clad. Therefore, model considers the data provided by CFD and thermal models and calculates for each point clad growth for the input parameters established. Clad is discretized using parallelepipeds of \( W^2 \) section and \( H \) height (Fig. 5), that only grows there where the melt pool is formed.

![Clad discretization](image)

Fig. 5. Clad discretization starting from powder concentration and melt pool shape.

The model is based on a mass balance between injected material at each point and mass of the parallelepiped generated in that point (Eq. 2).

\[
m_i^i = m_p^i
\]  
(2)

The injected mass can be calculated multiplying the material flow in that point by the time during which powder is being injected on it (Eq. 3).

\[
m_i^i = \phi_i \left( \frac{Kg}{m^2 \cdot s} \right) \cdot W^2 \left[m^2\right] \cdot t\left[s\right]
\]  
(3)

Where injection time can be calculated as length of discretization step divided by the cladding nozzle velocity (Eq.4).

\[
t = \frac{\Delta s\left[m\right]}{V_f\left[m/s\right]}
\]  
(4)

Finally, the mass of generated parallelepiped is obtained multiplying its volume by the material density (Eq. 5).

\[
m_p^i = \rho \cdot \left( \frac{Kg}{m^3} \right) \cdot W^2 \left[m^2\right] \cdot H_i\left[m\right]
\]  
(5)

These expressions allow developing the mass balance presented in Eq. 2.
\[ \phi_i \cdot W^2 \cdot \frac{\Delta s}{V_f} = \rho \cdot W^2 \cdot H_i \] (6)

Simplifying the Eq.6 allows to obtain the geometric model formulation that allows to estimate the clad height for each point of the domain (Eq.7).

\[ H_i = \frac{\phi_i \cdot \Delta s}{\rho \cdot V_f} [m] \] (7)

CFD model provides powder concentration ($\Phi_i$) on each point of the resolution domain which is enclosed by data of thermal model, and trajectory discretization ($\Delta s$), material density ($\rho$) and velocity ($V_f$) are input parameters of the process. Therefore, the model is not only capable of estimating the geometry of the added clads but it also can estimate local effects such as changes in the trajectory direction or clad overlapping.

![Model results](image)

Fig. 6. Model results. Left) Single Clad. Centre) Direction change. Right) Clad Overlap.

Model has been experimentally validated for Inconel 718, a nickel-based superalloy commonly used in aeronautical industry. The average errors observed in validation tests for both estimated height and width are below 8% of measured values (Fig. 7). These errors show the validity of the model as tool in the adjustment step of the laser cladding process and its industrial implementation.

![Validation results](image)

Fig. 7. Validation for different input parameters. Left) Height results. Right) Width results.

3. CONCLUSIONS

Laser cladding is becoming of a great interest process for high added value parts manufacturing sector due to it makes possible most sustainable and cost efficient production chains. The main problem of process industrialization is the necessary experimental work for process tuning. A solution to reduce the time expend for optimal parameter obtaining is
process modelling. This work presents an innovative model which estimates clad geometry starting from data provided by other models. Firstly, a CFD model calculates the powder distribution in nozzle outlet. Data from CFD model is used, on one hand, by an attenuation model to calculate the real energy reaching the substrate, and on the other hand, by a geometric model in order to estimate clad growth. Attenuation model takes the powder distribution on planes where the flux is consolidated and calculates step by step the attenuation suffered by the beam until reaching the substrate. Thermal model uses this attenuated distribution to estimate the temperature field in the substrate for each time step. Finally, geometric model using both the powder distribution and temperature field onto the substrate surface, calculates the added clad geometry for each step of the trajectory. Model has been experimentally validated obtaining average errors below 8% for clad height and width estimation.

ACKNOWLEDGEMENTS

The authors wish to thank the Spanish Ministry of Science and Innovation for financial support provided through the SURFACER project (Ref: DPI2010-20317-C02-01).

REFERENCES

A novel strategy for the incorporation of optical sensors in Fused Deposition Modeling parts

A. Kantaros, J. Giannatsis¹, D. Karalekas²

Laboratory of Advanced Manufacturing Technologies & Testing, Dept. of Industrial Management & Technology, University of Piraeus, Greece
¹ggian@unipi.gr, ²dkara@unipi.gr

ABSTRACT
Embedded sensors in Additive Manufacturing (AM) can provide valuable information both for studying the actual fabrication process as well as for controlling and monitoring the fabricated part or product during its functional life. Although this application is in principle possible, there are significant challenges in the embedding process, mainly involving specific constraints posed by the sensor and the AM technology, that need to be overcome in order for it to be fully realised. In this context, a strategy for embedding Fiber Bragg Gratings (FBG) for strain measurements in Fused Deposition Modeling (FDM) parts is discussed, in the present paper. The proposed strategy involves the simultaneous fabrication of the specimen and of the designed FBG supporting and aligning fixtures, which allows the fast and accurate placement of the optical sensors minimizing thus possible interference effects. Some preliminary results of strain measurements from FBG fibers embedded in FDM parts are also presented.

KEYWORDS: Additive Manufacturing, Fused Deposition Modeling, Monitoring sensors

1. INTRODUCTION

The field of Additive Manufacturing technologies, although not quite new, has attracted a lot of recent attention as it opens new possibilities and poses new challenges in manufacturing functional parts. The unique nature of controlled and gradual material addition in Additive Manufacturing (AM), that enables the fabrication geometric structures of almost any complexity in relatively short time without the need for specific production tools, has been the main reason for their establishment as an indispensable tool for physical prototyping¹ during product development and design. Ongoing research is mainly focused on the study and improvement of existing processes and the development of new materials that will further enhance the quality and performance of AM parts, as well as the economic efficiency of the corresponding fabrication processes thus enabling their employment as direct manufacturing

¹This is the reason that the term “Rapid Prototyping” is also associated with them.
processes [1]. Recently a point has been reached where they can be considered as a viable alternative to conventional manufacturing technologies (casting, forming and cutting) both for the fabrication of complex parts/components in small numbers [2] as well as for the production of short-run manufacturing molds and dies [3]. In parallel new types and configurations of AM processes that enable the production of functionally graded [4], composite [5] and multi-material structures [6] are developed for a variety of applications such as the production of biomedical products, automobile/aerospace components, tissue engineering and biofabrication [7].

A particularly interesting field of research is that of constructing AM parts with embedded components, in order to simplify or eliminate assembly of products, or sensors, to provide data sensors for control and monitoring. In a study by Kataria and Rosen [8] it has been shown that it is possible to construct complex functional Stereolithography parts with embedded mechanical and electrical components. Relevant applications include also the construction of a non-assembly, multi-articulated robotic hand with inserts employing Stereolithography [9], as well as the construction of actuated assemblies using the PolyJet technology [10]. Process planning issues are quite important in this context, especially for flexible components, in order to maintain the shapes of the inserts during processing and to control precisely the part geometry [11].

Embedded sensors can help monitoring certain aspects of the functional behavior of a prototype or part, thus providing data for validation and improvement of its design, as well as a better understanding of its functional characteristics. The incorporation of various types of sensors, such as wireless sensors, optical fiber sensors and thermo-mechanical sensors, have been examined in various AM technologies [12]. Fiber Bragg Grating (FBG) sensors are a common type of optical sensors that can be employed for the measurement and monitoring of the strain, stress and temperature in specific locations of a 3D structure [13]. In the context of AM study, FBG sensors have been employed both for the study of the actual fabrication process and of the corresponding materials, as well as for experimental testing of specimens. In the case of materials investigation, FBG sensors have been employed for the study of solidification strains in photocurable resins for microstereolithography [14, 15]. FBG sensors have been incorporated also in Shape Deposition Modeling (SDM) [16] and Selective Laser Sintering (SLS) specimens in order to provide data for a study of their mechanical behavior for validation and design purposes [17, 18].

In the present paper a new strategy for the incorporation of FBG sensors in Fused Deposition Modeling (FDM) specimens is presented. The main purpose for inserting of FBGs is to monitor the accumulation of residual stresses, due to solidification shrinkage [19], in FDM parts during and after fabrication. Based on the results of the aforementioned studies it is obvious that incorporating sensors in AM is feasible. Sensor embedding may, in fact, seem rather straightforward, due to the controlled additive nature of the processes (usually in the form of layer addition) which can be interrupted, in specific points, for the placement of the sensors and subsequently continued. There are, however, significant challenges, mainly involving the accuracy and speed of the embedding step, that need to be addressed in order for it to be fully realized. In particular, the embedding step must be carried out as fast as possible, in order to minimize possible negative effects on the material addition process due to interruption. In the context of residual strain measurement, accuracy in positioning the FBG is also of great importance, in order to assure the uniformity of strains around the grating section of the fiber and minimize possible edge effects in measurement, as described in detail by Colpo et al. [20]. Furthermore, other constraints or limitations posed by the specific characteristics of both the sensor and the AM technology, such as the material properties and the machine chamber environment temperature, should also be taken into account.
2. FDM PROCESS

Extrusion-based technologies form a class of AM in which parts are fabricated (built) by successive deposition of layers which are formed by the controlled flow of the semi-liquid raw material through one or multiple nozzles in the deposition “head” assembly (Fig. 1). Currently there are several variants of extrusion-based systems that employ different types of raw materials and deposition mechanisms for the fabrication of parts [21]. The most common extrusion-based technology is FDM, in which thermoplastic polymers are deposited through the moving deposition head in order to form the required layers. Polymer is supplied in the form of filament, which is liquefied through heating in the nozzle. The most commonly used polymer in FDM is ABS (Acrylonitrile Butadiene Styrene) but several other thermoplastic materials can be employed. In most FDM systems a second nozzle is also employed for the deposition of a support material that is used for the construction of a temporary support structure. This structure is required for supporting overhang layers and safe attachment of the part on the platform, and it is removed manually after part fabrication in the machine finishes.

![Fig.1. Schematic representation of the FDM fabrication process (source: http://www.custompartnet.com/wu/fused-deposition-modeling).](image)

Based on FDM technology and its close relative Fused Filament Fabrication (FFF) several commercial and open-source systems have been developed and installed worldwide. Compared with other AM technologies, FDM main advantage lies on the use of industrial type materials with good mechanical properties, which are quite easily available and at relatively low cost. However, there are certain disadvantages mainly associated with the accuracy and speed of the fabrication process. A specific problem that affects both the dimensional accuracy and functional behavior of the part is the existence of internal stresses which are caused by the contraction (shrinkage) of deposited fibers as they solidify during layer formation. The effect of shrinkage is usually apparent as a contraction of the lowest layers and an upward curl near
the edges of the part, especially in parts fabricated with high layer filling density. An example of the curl effect is presented in Fig. 2, where a significant distortion in the shape of an orthogonal slab specimen is apparent at the edges, which amounts to an upward deviation of approximately 1 mm (10% of height) of the edges from the horizontal plane. Besides geometric deformation, the accumulation of internal stresses may also lead to other fabrication defects due to the detachment of the part from the support structure or the platform, as well as affect its mechanical strength in certain loading conditions causing layer delamination and failure.

![Fig.2. FDM part with significant curl deformation.](image)

3. **FBG EMBEDDING PROCESS**

In order to effectively measure the amount of internal strains in specific locations of parts, the accurate and precise incorporation of FBG sensors in FDM specimens has been investigated. The extrusion-type FDM system under investigation is a relatively low-end industrial AM system (3D printer) of small build size (203 x 203 x 305 mm), which is mainly targeted at the design modeling and prototyping sector, but is nevertheless capable of relatively good product accuracy (minimum layer thickness: 0.178 mm) and quality. ABS-P400, a type of ABS specifically developed for FDM systems, has been employed as the raw material for the production of specimens. The dimensions of the prismatic specimen, in which a single FBG sensor was embedded, are 10 x 10.4 x 40mm (width x height x length). The diameter of the optical fibers employed is 0.125mm and the grating length of the FBG section, operating at 1550nm Bragg wavelength, is 1mm.

Accurate strain measurements and the reliable calculation of the corresponding stresses from them, requires that the FBG fiber section is positioned at the center of the specimen, in order to minimize possible edge effects, and that it is perfectly aligned with the specimen’s long (axial) dimension, in order to obtain measurements in the longitudinal direction. This implies that the FBG section should be positioned, as close as possible, to the center of the middle plane of the specimen (Fig. 3). This means that employing a layer thickness of 0.254 mm the specimen building comprises of 41 layers and the FBG fiber should be embedded in the 21\textsuperscript{st} layer.
In order to position the sensor exactly a holding fixture is required. Positioning of the fixture and fiber, before part building begins, is not possible since it would interfere with the movement of the deposition nozzle. Fiber embedding should, therefore, be carried out by interrupting the building process at the specific layer. Since part building takes place in a closed heated chamber of approximately 75°C, the embedding step must also be performed as fast as possible, in order to minimize the possible effect of temperature loss. Furthermore, the optical fibers should be aligned and sufficiently fastened/secured, so that they are not carried away or injured by the nozzle movement during material deposition. If not, fiber damage or excessive fiber bending could occur, resulting to measurement inaccuracy and even signal loss. Material deposition in layer filling follows by default a crisscross pattern, depositing thin ABS filaments in +45°/-45° raster lines, alternating at successive layers.

![Fig. 3. Schematic representation of the specimen with the embedded FBG fiber.](image)

In order to satisfy the above constraints a modular fixture design has been devised. The fixture design is composed by two fixture assemblies left and right of the specimen, each of which comprised of two components/parts (Fig. 4). The fixture bottom part (basis) is fabricated simultaneously with the specimen, in order to provide a reference support structure for the exact placement of the upper part (grip), where the optical fiber is attached.

![Fig. 4. Schematic representation of the fixture assembly.](image)

Fabrication of the two required grips (also by FDM) and fiber attachment is performed prior to specimen building, employing an alignment structure that mirrors the exact arrangement of the specimen and fixtures at the moment of process interruption (Fig. 5a). In this arrangement, the two grips carrying the fiber are assembled on the bases and fiber/FBG section alignment is performed. In order to attach the fiber on the grip, the fiber is passed through a 'tunnel' in the grip, after one of the fiber connectors is removed. The fiber is then
fastened on the grip employing a small amount of thermoplastic adhesive (hot-gun glue), in order to secure firm attachment during handling and positioning. After fiber alignment and fastening, specimen fabrication at the machine begins. The fabrication process is interrupted manually at the 20\textsuperscript{th} layer in order to place the fiber-carrying grips. The process is then continued and the remaining layers of material are deposited (Fig. 5b).

![Fig. 5(a)](a)  ![Fig. 5(b)](b)

**Fig. 5.** The aligning and fiber fastening step (a) and the final specimen containing the FBG fiber (b).

Following the above described strategy several test fabrication studies have been conducted. The tests showed that the time required for fiber placement is in the magnitude of a few seconds; hence the perceived process interruption effects could be considered minimal. Further testing of the embedded FBG sensor showed also that it is capable of providing strain measurements which according to some preliminary results are in the range of 6,700 microstrains. This value is comparable to the values obtained through similar experimental strain measurements in ABS parts fabricated by an open-source 3D printer [19]. A subsequent preliminary investigation showed that the incorporation of multiple parallel fibers at various heights can also be accomplished following the above described strategy, with a simple modification of the fixture design that allows the arrangement of multiple bases, as shown in Fig. 6.

![Fig. 6](a)

**Fig. 6.** Specimen with three embedded fibers.
An interesting extension of the present study would be to evaluate the applicability of the proposed methodology for embedding FBG sensors in actual functional parts of more complex geometry. These sensors could be then applied for dynamic control and monitoring of the part/component during its functional life. Part complexity as such should not be expected to impose some difficulties since the basic principle of decomposing a part in layers, which is the main reason for geometric versatility in AM, permits also the incorporation of sensors in specific layers, indifferent of the overall geometric complexity. There are, however, some perceived constraints that should be expected. One such constraint concerns the relative size of the sensor, which must be sufficiently small compared to the layer thickness and the size of the part, in order to assure the complete embedding of the grating section and the elimination of edge effects on measurements. Furthermore, following the proposed methodology, it is possible to place multiple fibers, but only in layers/planes parallel to that of the machine platform (vertical to build direction). Finally the size of the chamber combined with the need for multiple alignment structures may impose a limit on the number of sensors that could be incorporated.

4. CONCLUSIONS

Embedded FBG sensors in FDM parts can prove to be a valuable source of information for studying the actual fabrication process as well as for monitoring and investigating the part’s behavior in subsequent testing or during its functional use. In the present paper a strategy involving the concurrent fabrication of the part and the fiber aligning fixture set-up is presented. The proposed placement/embedment technique offers enhanced versatility in placing a number of sensors in the FDM part, relatively fast and with sufficient placement accuracy. The reliability and accuracy of the measurements depends on the positioning of the grating/gauge in the center of the specimen, as well as on the accuracy of the fiber’s alignment with the longitudinal specimen axis. Both requirements are checked and satisfied in the present study during the alignment and positioning phases, to eliminate the appearance of any edge effects and of birefringence. Recording from the embedded FBG fibers showed that they can provide reliable measurement of strains, generated during the building process of the FDM component.

REFERENCES


Cooling Rate Prediction in case of Pipelines Longitudinal Welds performed by Submerged Multi-Arc Welding

Carmen Catalina Rusu, Luigi Renato Mistodie, Elena Scutelnicu

Department of Manufacturing Engineering
Faculty of Mechanical Engineering
“Dunarea de Jos” University of Galati
Romania,
carmen.rusu@ugal.ro

ABSTRACT

The submerged arc welding (SAW) process is a frequently used in pipelines manufacturing. In industry, the multi-arc/multi-wire welding is widely applied, in order to increase productivity compared with the classical arc welding processes. The common base materials used in pipelines fabrication are the micro-alloyed low carbon steels, originally termed High Strength Low Alloy (HSLA) steels, which have good mechanical proprieties such as strength, toughness, fatigue and collapse resistance and strain tolerance. However, during the SAW welding process, a greater attention on the heating rate, peak temperatures and cooling rate should be given. These effects, that strongly influence the weld and heat affected zone (HAZ) microstructure, are very difficult to predict by classical methods and the development of mathematical models should be made.

The paper presents the results of three-dimensional finite element modelling of heat transfer in the welded joints performed by submerged multi-arc welding process. Temperature field, thermal cycles and cooling rates are analysed and discussed. Finally, important conclusions regarding the thermal history and cooling rate prediction, in case of submerged multi-arc welding, are drawn.

KEYWORDS: pipelines fabrication, HSLA steel, submerged arc welding, cooling rate, 3D finite element model

1. INTRODUCTION

The submerged arc welding processes (SAW) in well known as the most applied procedure in pipelines fabrication for oil and gas transportation and/or distribution. The longitudinal welds of pipelines are, preferable, performed by submerged arc welding process. In recent years, there have been developed new submerged arc welding processes employing multiple arcs, especially used in the manufacturing of large welded pipelines, existing nowadays up to 6 electrode wire variants of the process [1]. The using of multiple submerged arcs ensures qualitative welded joints, with better mechanical properties, and an increased
The productivity increasing is limited by the heat input generated during the submerged arc welding (SAW) process. In order to get better mechanical characteristics of the welded joints, the heat input should be limited or distributed on multiple arcs [2].

- Thermo-physical material proprieties depend on temperature;
- Modelling of welding heat source is based on the Goldak model [8];
- Convection and radiation heat loss are considered;
- The base material is considered isotropic;
- The initial temperature of the base material is considered 20°C.

Heat flux losses by convection and radiation have been introduced in the heat transfer model [2]. Convection heat transfer coefficient of 12 Wm⁻²K⁻¹, emissivity of 0.7 and Stefan–Boltzmann constant of 5.67x10⁻⁸ Wm⁻² K⁻⁴ have been set.

The welding speed is considered constant during the whole process. The heat input is kept constant while the heat source moves continuously, heating and melting new regions in front of it and keeping its influence on the weld pool, as too. In order to simulate the welding
process, defining the time function as a sum of triangle functions was necessary. The succession of the triangle functions describes the positions of the heat sources. Considering the total welding time and the position of the heat source along the joint longitudinal axis, the number of the triangle functions can be calculated. The temperature field is the result of the total thermal effects during the welding process starting from the initial moment \( t=0 \) until the final moment \( t=t_n \).[9]

The element type is 3D SOLID 8-20 and the final model includes 101848 elements. The welded joint model was assumed to be symmetric and, therefore, only one half of the model was subject to the analysis. According to Goldak [8], the heat flow equation which describes the heat flux on the regions of the heat sources is the following:

\[
Q(x, y, z) = \frac{6\sqrt{3} \cdot \eta \cdot U \cdot I}{a \cdot b \cdot c \cdot \pi \cdot \sqrt{\pi}} \cdot \exp \left( -\frac{3x^2}{a^2} - \frac{3y^2}{b^2} - \frac{3z^2}{c^2} \right)
\]

where \( Q \) is the heat flux developed by each heat source; 
\( a, b, c \) – the thermal source dimensions; 
\( U \) – arc welding voltage; 
\( I \) – amperage; 
\( \eta \) – arc welding efficiency.

Two double-ellipsoidal heat source, with different dimensions, have been considered to the development of this complex model. The heat distribution within the moving heat sources is given by Eq [1], while the applied heat distribution is shown in figure 1. A schematic view of heat sources illustrated in figure 1 shows their position on the steel sheets [10].

For the prediction of cooling rate in case of double submerged arc welding process two configurations are considered and analysed. The first case is double arc/double wire (Fig. 2) configuration in which the heat sources simulate two distinct melting pools. The second case is double arc/three wires configuration (Fig. 3), considered to have three heat sources, which simulates the welding arcs, the first two heat sources simulating the overlapping heat effect generated by the first two wires creating the same melted pool and the third source simulating the heat effect generated by the third wires. The distance between the melted pools is 100mm.
Fig. 2. Double arc/double wire configuration. Fig. 3. Double arc-three wire configuration.

The welding parameters for the double arc-double wire and double arc-three wire welding process introduced in the numerical analysis are presented in the table 1. These parameters are chosen as real parameters that authors used in the experimental research. In case of welded gas pipelines achieved by double arcs SAW, API-5L [11] standard recommends X gap. The gap geometry for butt welding pipelines is presented in figure 4.

Table 1. Welding parameters used in the process FEA simulation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>double arc/double wire configuration</td>
<td>double arc/three wire configuration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top surface</td>
<td>WA 1</td>
<td>800</td>
<td>35</td>
<td>0.7</td>
<td>WA 1</td>
<td>800</td>
<td>35</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>WA 2</td>
<td>650</td>
<td>36</td>
<td>0.7</td>
<td>WA 2</td>
<td>800</td>
<td>35</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>WA 3</td>
<td>650</td>
<td>36</td>
<td>0.7</td>
</tr>
<tr>
<td>Bottom surface</td>
<td>WA 1</td>
<td>750</td>
<td>34</td>
<td>0.6</td>
<td>WA 1</td>
<td>750</td>
<td>34</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>WA 2</td>
<td>600</td>
<td>34</td>
<td>0.6</td>
<td>WA 2</td>
<td>750</td>
<td>34</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>WA 3</td>
<td>600</td>
<td>34</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Fig. 4. Gap geometry.

3. RESULTS AND DISCUSSIONS

Temperatures distribution at different time steps, when two/three thermal sources are applied for the case of double arc/double wire (DA/DW) is presented in figure 5 a, b and c and for double arc/three wire design (DA/TW) in the figures 6 a, b and c. At the beginning of the process, in the non-stationary process phase, the extent of the HAZ is less due to the lower temperature of the plates and the instability of the process (Fig. 5a and 6a). In the quasi-stationary phase - when the welding process is stabilized - the peak temperature reached in the area of the first thermal source, in case of DA/DW configuration is about 1588°C and about 1995°C (Fig. 5b and 7) in the area of the second thermal source. For the DA/TW design the peak temperature reached in the area of the first arc (heat sources 1 and 2) is about 1955°C and about 1594°C (Fig. 6b and 8) in the area of the second arc (third heat source).

For the two configuration proposed, the numerical results were processed and charts were drawn as figure 7 and 8 illustrates. The thermal history of few nodes, perpendicularly located at different distances on the longitudinal axis of the welded joint, was analysed. Obviously, the peak temperature is reached in the nearest node which is right on the axis of the weld. Also the shape of the thermal cycle indicates the influence of the second or third thermal source on the temperature history and further on the temperature field. In the heating phase, the approach of the first source leads to a temperature increase in all nodes selected.
The peak temperature is reached when the heat source is positioned in the node located on the welded joint axis. In the nearest node from the longitudinal axis of the joint, the temperature gradient is very high meaning that the temperature decrease is highest.

![Fig. 5. Temperatures field distribution at different times of the welding process in case of DA/DW configuration.](image1)

![Fig. 6. Temperatures field distribution at different times of the welding process in case of DA/TW configuration.](image2)

After that, the heat source moves away and a more and less decrease of temperature is noticed in the nodes located near the axis of the joint. In the points located far away, the temperature continuously increases due to the low heating speed and the approach of the second source. In these conditions the cooling phase is missing. In the case of DA/DW design, when the second welding is approaching, the temperature starts increasing and the peak
temperature, reaches 1995°C (Fig. 5b and 7). From this moment the cooling by convection starts in all points analysed and the temperature decrease is correlated with the cooling speed.

Fig. 7. Thermal cycles of the nodes located at different distances from the weld axis in case of DA/DW configuration

Even if the first thermal source induces a higher thermal input on the welded sheets, the maximum temperature reached are lower than the ones induced by the second thermal source and this is due to the fact that the first thermal source encounters cold areas of the plates and the second one encounters hot areas (Fig. 6a and 8).

For the case of DA/TW configuration, the thermal history of the same nodes as the DA/DW design - perpendicularly located at different distances on the longitudinal axis of the welded joint - was analysed. As in the first case, the peak temperature is reached in the nearest node which is right on the axis of the weld. Also the shape of the thermal cycle indicates the influence of the third thermal source on the temperature history and further on the temperature field. In the heating phase, the approach of the two overlapped heat sources lead to a temperature increase in all nodes selected and the peak temperature of 1955°C is reached when the first two heat sources are positioned in the nodes located on the welded joint axis. After that, a decrease of temperature is noticed in the nodes located near the axis of the joint because of the outlying of the first two sources. As in the first case, the cooling phase is missing. When the third heat source is approaching, the temperature starts to increase again, but it doesn’t exceed the peak temperature reached when the first double sources are applied, but reaches the value of 1594°C.

It can be noticed that the thermal cycles, from the two configurations analysed, are different. This outcome can be explained by the fact that in case of DA/TW configuration, the first two thermal sources, that form the same melted pool, induce a higher thermal input and overcomes the phenomenon of temperature decrease, due to the cool plate areas. As conclusion, the use of DA/TW configuration can be considered as a good way to achieve a better penetration and melting of the welded joint that leads to a productivity increase of the submerged welding process and a slower rate of cooling.

Numerical data related to heating and cooling rates for the DA/DW and DA/TW configuration were processed and plotted (Fig. 9 and 10). They are correlated with the complex cycle developed during the submerged double arc welding process. Because the thermal cycles are different, it is expected that the heating/cooling rates are different, too. The heating/cooling rate was computed as difference between temperatures values reported to an interval of one second.

The results show the major influence of the wires number and arcs configuration on the heating/cooling rate in case of double submerged arc welding process. It can be noticed that the highest cooling rate is achieved when the DA/DW configuration is proposed and decrease with the wires numbers applied. Another important issue is the distance between arcs. This
fact must be considered in the real welding conditions and the material between heat sources should not suffer a cooling under the limit $A_{c1}$, region where the steel is partially recrystallised [6].

![Fig.9. Thermal history of welded joint for DA/DW configuration](image1)

![Fig.10. Thermal history of welded joint for DA/TW configuration](image2)

4. CONCLUSIONS

A three dimensional finite element model of longitudinal butt welded joints performed by SDAW was developed and presented in the paper. For the proposed welding process, two configuration of the welding procedure are simulated and analysed: double arc/double wires (DA/DW) and double arc three wires (DA/TW). After reviewing FEA related to the thermal history and cooling rate prediction - in case of submerged double arc welding process - some important conclusions should be highlighted:

- FEA can be a useful instrument for temperatures field prediction and cooling rate prediction in pipelines welded joints;
- A specific heating-cooling-heating-cooling complex cycle is developed during SDAW for both DA/DW and DW/TW configurations;
- The thermal cycles and history for the two configurations are different due to the different heat input induced in the sheets and the numbers of heat sources applied;
- The highest cooling rate is achieved when the DA/DW configuration is proposed and decrease with the wires numbers applied on the first arc;
- The use of DA/TW configuration can be considered as a good way to achieve a better penetration and melting of the welded joint that leads to a productivity increase of the submerged welding process and a slower rate of cooling;
- The second arc plays a positive role, reflected into the cooling speed decrease of the welded joint.

In order have a complete overview related to the pipelines steels behaviour, a deep investigation of mechanical and metallurgical changes, so that the optimum number of wires chosen in the real welding technology, corresponding to the achievement of the base material best characteristics, should be set in accordance with data achieved on temperature field, thermal cycles and cooling rate described in this paper.

ACKNOWLEDGEMENTS

This work was supported by the Romanian National Authority for Scientific Research, CNDI–UEFISCDI, through grant 27/2012, project number PN-II-PT-PCCA-2011-3.1-1057.
5. REFERENCES


A new approach to modelling friction stir welding using the CEL method

M. Hossfeld, E. Roos

Materials Testing Institute (MPA), University of Stuttgart
email max.hossfeld@mpa.uni-stuttgart.de

ABSTRACT

Although friction stir welding (FSW) has made its way to industrial application particularly in the last years, the FSW process, its influences and their strong interactions among themselves are still not thoroughly understood. This lack of understanding mainly arises from the adverse observability of the actual process with phenomena like material flow and deposition, large material deformations and thermomechanical interactions determining the mechanical properties of the weld.

To close this gap an appropriate numerical model validated by experiments may be helpful. But because of the issues mentioned above most numerical techniques are not capable of modelling the FSW process. Therefore in this study a Coupled Eulerian-Lagrangian (CEL) approach is used for modelling the whole FSW process. A coupled thermomechanical 3D FE model is developed with the CEL formulation given in the FE code ABAQUS® V6.12. Results for temperature fields, weld formation and the possibility of void formation are shown and validated.

KEYWORDS: FEM, friction stir welding, coupled eulerian lagrangian, microstructure, experimental validation

1. INTRODUCTION

Friction stir welding (FSW) is a solid-state joining process mostly used for the joining of aluminium alloys. Invented in 1991 at TWI in England [1, 2], FSW has made its way to industrial application particularly in the last years [3]. At first, this is because of the capability of producing welds with excellent properties like very good static and fatigue strength, low distortion and almost plain surfaces even in the as-welded condition [4, 5]. Furthermore, the possibility to join dissimilar materials such as aluminium and steel or aluminium and copper enables tailored blanks for lightweight designs or low resisting high current connections. But beside this there are other uprising advantages of FSW as in today’s production environmental issues become more and more important. FSW consumes only about 2.5% of the energy of laser welding [6]. Also unlike other welding processes, FSW is free of toxic fumes and without the need for filler, gas shield or post weld heat treatment of the (usually almost plain) process zone. In addition high strength FSW welded joints enable automotive light weight constructions with a better material utilisation degree and decreased fuel consumption.
2. FRICTION STIR WELDING PROCESS

2.1. Operational principles, tool geometry, parameters and resulting microstructure

Even though the FSW process implies complex interactions between material properties and flow, heat transport and process forces, the basic operational principles are quite simple. The process mainly consists of a combination of frictional heating of the material and a stirring motion caused by a rotating tool. While friction and also plastic work dissipation heat, soften and plastify the material, the stirring motion mixes the material across the interface, resulting in a characteristic microstructure.

Tool geometry

Because all these aforesaid tasks have to be fulfilled by the FSW tool, tool geometry and material are always issues of permanent optimization [7–9]. The tool consists of a cylindrical shoulder and a protruding pin. The shoulder performs two main functions. First, the application of the bigger part of the process forces such as downward force and torque. Second, it prevents the plastified material from being pushed out of the actual processing zone under the shoulder. Therefore, the shoulder is mostly carried out concave. Particularly with respect to welding thin sheets, the shoulder additionally contributes notably to material intermixure. The main functions of the pin are both stirring of the material and an additional heating of the workpiece. Usually the pin is carried out as truncated cone or as a cylinder. For a better intermixure or for certain applications also other tool geometries are used, for instance other pin base geometries like squares or triangles, sometimes provided with flutes, threads etc. or shoulders with spiral threads.

FSW process

The FSW process can be divided into three essential phases (Fig. 1). First, the usually slightly tilted tool is plunged into a joint until the shoulder contacts surface of the material. After a short time of pre-heating, called dwelling, the rotating tool is moved along the joint line until the desired length is reached. In a last step the tool is removed from the weld. The remaining exit hole can be avoided through a retractable pin. For the sake of completeness it is mentioned at this point that the process is also capable of three-dimensional welds and also spot welds.

Fig. 1. The friction stir welding process. Illustrated without clamping.
Process parameters

The main process parameters of the friction stir welding process are rotational spindle speed $\omega$, traverse speed $V_x$, tool angle $\alpha$ and downward force $F_N$ (Fig. 2). The process may also be driven displacement-controlled. In this case the downward force $F_N$ is a reaction of the depth of immersion. This depth is usually measured from the workpiece surface to the lowest part of the shoulder and called heel plunge depth.

It should be mentioned that not only the parameters determine important process factors such as generated heat, material transport or processing zone compression, but also they interact strongly. For instance a slightly increased angle of the tool may result in a significant increased downward reaction force in a displacement-controlled process.

By determining aforesaid process factors, the process parameters also determine directly the weld evolution and geometry, its microstructure, surface and the quality of the welded joint.

Fig. 2. Parameters of the friction stir welding process and microstructural zones.

2.2. Microstructure of friction stir welded joints

Friction stir welded joints consist of four characteristical zones showing different microstructure and so mechanical properties. These are

- An inner processing zone with very strong influences of pin and also shoulder, usually called stirring zone or Nugget
- A Thermo-Mechanical Affected Zone (TMAZ)
- The Heat Affected Zone (HAZ), where influences of the temperature are observable
- The base material which is thermal and mechanical quasi unaffected

Figure 2 shows the different zones and their locations. In extreme cases, zones can grow or shrink significantly and almost merge together, e.g. when welding thin or very thick plates. Then a metallographic examination may not be as definite as in Fig. 3.
The shape of the Nugget zone depends remarkably on the tool geometry and the thickness of the plate. It is slightly unsymmetrical because of the tool rotation direction. Due to the severe solid state deformation during FSW this zone is fully dynamically recrystallized. This refines the grains up to diameters about 5-10μm and thus results in a high-strength weld. This phenomenon is called Hall–Petch strengthening [10, 11].

The TMAZ surrounds the nugget zone on both sides. The TMAZ is also subjected to remarkable plastic deformation and heat input. In this zone the material is not stirred and not dynamic recrystallized, but the microstructure is usually heavily distorted and the affected grains are elongated, particularly at the transition to the nugget zone (Fig. 3). These elongated grains can reach easily lengths about 100μm and widths about 5μm. Furthermore, some recrystallization seeds can occasionally be found in the TMAZ.

The TMAZ in turn is enclosed by the HAZ. Although this zone is not subjected to plastic deformation the microstructure of the weld may be altered by the process heat input. Depending on the given aluminium alloy and its original condition this may range from slight grain growth to accelerated age hardening. For example a grain growth from ~35μm to ~40μm was observed in the HAZ for welded rolled AW 5182-0 sheets.

**The influence of the microstructure on strength properties**

As mentioned before, mechanical properties vary among the different zones. Beside the effects mentioned above, there are additional influences like coherent and incoherent dispersoids, generation of dislocations, misorientation of grain angles etc. [3, 12, 13]. Moreover, crack paths predefined for instance by oxide particles or dispersoids have definite effect on fatigue life [4, 14]. Crack growth rates are strongly driven by microstructure of the weld [15].

In summary, both the processing zone and the additional welding zones are stress concentrators as a consequence of material inhomogeneities.

![Fig. 3. Barker's etching on three characteristical microstructural zones of FSW.](image)
3. NUMERICAL MODELING AND SIMULATION OF FSW

Although the basic principle of the process is simple, the modes of action in FSW are complex and not thoroughly understood. This lack of understanding mainly arises from the adverse observability of the actual process. To close this gap an appropriate numerical model validated by experiments may be helpful to investigate influences on the FSW process and their strong interactions. The challenging issues when modelling the FSW process are already contained in the process description above:

- large deformations and distortions
- mechanical properties of the materials are functions of temperature, strain, strain rate etc.
- non-linear phenomena like contact, friction etc.
- fluid-structure interaction
- detailed representation of the tool geometry if so with flutes, threads etc.
- representation of the weld geometry
- modelling the intermixture of the two sheets

Numerical models for FSW have been published by various authors, modelling techniques and goals [16–25]. Notable results were reached by use of the arbitrary Lagrangian–Eulerian formulation (ALE) by Ulysse [16], Schmidt and Hattel [17] or Guerdoux and Fourment [18]. When a lagrangian formulation is used, the main problem is the highly distorted mesh resulting in stability problems and time increment issues. Most authors deal with that by continuously remeshing or local mesh refinement, e.g. Guerdoux and Fourment [18]. In addition the plunging step of the FSW process is often not modelled because this easily causes excessive mesh distortion [17, 22]. Furthermore, very often only one instead of two sheets is used to represent the workpieces.

To avoid these issues, in this paper the Coupled Eulerian-Lagrangian (CEL) method by Noh [26] is investigated to model the FSW process. The CEL method uses a lagrange-plus-remap algorithm. When the mesh distorts during a lagrangian increment, the mesh is restored by calculating the material flow between elements and subsequent remapping [26, 27].

3.1. Numerical Model

The model was build up and simulated with the CEL formulation included in the FE code ABAQUS® Explicit V6.12 [27, 28]. The simulation represents all three phases of the FSW process shown in Fig. 1.

Geometry, assembly and boundary conditions

Figure 4 shows the assembly including boundary conditions. The tool is modelled as a linear-elastic lagrangian body. It has a length of 58 mm from which 35 mm belong to the fixture. The shoulder has a diameter of 12 mm with a concavity angle of 7°. The diameter of the pin is 5 mm and its length is 3 mm. The tool rotates with an angular speed of 209.5 rad/s while its tilt angle is 2°. The tool is represented in the model through 19,720 elements of type C3D8T. The two sheets lay within the eulerian mesh partial filled up with material. A gap between the sheets is possible for a more realistic setup (see Fig. 4 and Fig. 9). This feature may also be used for sensitivity analysis etc. The eulerian mesh with 255,300 elements of type EC3D8RT is bigger than the contained sheets. This enables the material flow and by this the
formation of the weld and burrs. When material passes over the borders of the eulerian mesh, the material and its properties are irretrievably lost.

![Tool, Lagrangian formulation](image1)

![Workpiece material area, void](image2)

![Workpieces, Euler formulation, filled](image3)

**Fig. 4. Schematic model assembly. Illustrated without fixture and without tool shank.**

The whole lower surfaces of the sheets are fixed in the z-direction. Also, all planes normal to the y-direction are fixed to avoid spreading. First, the rotating tool *plunges* into the two sheets ($V_z = 6$ mm/s). After the subsequent *dwelling* phase planes normal to the x-direction are charged with a velocity ($V_x = 600$ mm/min) to represent the feed motion. During the other phases, this velocity is set to zero. Heat dissipation from the welding process is modelled by heat conduction (tool shank and lower surface) and heat convection (upper surface and tool itself). These are given in Table 1 in the appendix. Furthermore, it is assumed that 100% of the friction dissipates in heat from which 90% flows into the aluminium sheet [29]. The environmental temperature is set to 20°C. The same thermal boundary conditions are used like in [18], see Table 2.

**Material properties**

The material parameters are extracted from literature and given in Tables 2 and 3 in the appendix. The elasticity of the aluminium alloy EN AW 6061 – T6 is modelled by an elastic–plastic Johnson–Cook [30] material model (1). Additional temperature dependent material parameters are given in Table 3. They are assumed to be isotropic.

$$
\sigma_y = \left[ A + B (\varepsilon^{pl})^n \right] \left[ 1 + C \ln \left( \frac{\varepsilon^{pl}}{\varepsilon_0} \right) \right] \left[ 1 - \left( \frac{T - T_{ref}}{T_{mel} - T_{ref}} \right)^m \right]
$$

*Contact*

The contact between tool and workpiece is represented by Coulomb’s law of friction with $\mu = 0.3$. A separation of tool and workpiece is possible. Schmidt and Hattel suggest this as a preliminary criterion for evaluating the success of the material deposition process during the simulation and important for a prediction when the suitable thermomechanical conditions and welding parameters are present [17].

In the model the contact was defined with the „*ALL* with self“ contact algorithm. This algorithm determines the contact condition for every single node. When contact between nodes is detected friction, frictional heating etc. are respected.
4. RESULTS

Figure 5 shows the calculated temperature fields starting with the transient phases plunging, dwelling and the transverse movement of the tool until the process reaches the steady state. During plunging and dwelling burr formation behind the tool (heel side) are higher than in front because of the slightly tilted tool. For the same reason heat generation and temperatures are highest on the heel side of the tool.

Over the whole process the temperature field is almost symmetrical with a very slight asymmetry on the advancing side caused by the higher deposit of material in this area. This slightly higher burr formation on the advancing side during welding with transverse movement matches reality [4, 33] and is also shown in Fig. 8.

Furthermore, Fig. 5 shows that high temperatures are locally limited and the temperature gradient is very high in the welding direction. Both facts are plausible and are confirmed by metallographic examination (compare Fig. 3 and Fig. 6), own experiments and literature, e.g. [4, 17, 18, 33, 34]. Also shown is the development of a turning at the advancing side of the tool during welding. Although the simulated temperatures are high considering the material properties of AW 6061, they are in good accordance with the experimental temperatures of Assidi et al [34]. Further capabilities of the model are shown in Fig. 6–9. Figure 6 shows the calculated equivalent plastic strain, temperature and the result of metallographical examination. The relationship between temperature and plastic strain and microstructure evolution is obvious. For a better comparison the contour of Nugget and TMAZ is layed over the numerical results (white). The high magnitude of accumulated plastic strain matches the results of Assidi et al [34] as well as the elevated temperatures at pin tip and shoulder. As expected the CEL formulation is capable of handling these large deformations.

![Fig. 5. Timeline with calculated temperature fields and weld formation during the different phases of FSW. Tool and clamping are not shown. Temperature in degree Celsius.](image-url)
Figure 7 shows a void formation at the bottom behind the tool due to failed material deposition. Moreover, Fig. 7 shows a metallographical examination of the appendant experiment. Figure 8 shows a stop action experiment with velocity vectors from the simulation with burr formation. These velocity vectors may enable a particle tracing for investigating the material flow during FSW. The possibility of a welding setup with a gap between the sheets is shown in Fig. 9.
5. CONCLUSIONS

In this study the CEL method by Noh was used for modelling the FSW process. A fully coupled thermomechanical 3D FE model for the whole process was built up in ABAQUS® Explicit V6.12. Some numerical results were validated. The following conclusions can be drawn:

- The results of the FE model, especially temperatures, weld geometry and plastic strain are in good accordance with own experiments and literature.
- An estimation of the microstructure evolution seems possible by using the numerical results for temperature and plastic strain.
- Most of the presented requirements of modelling FSW and its phenomena can be met by CEL. Main advantages of the developed FE model are the capability of large deformations, material flow, the possibility of burr and void formation and free surface tracking respectively. Furthermore an improved set-up with two material sheets and intermixture is possible and was shown.
- A dense mesh is important for the sufficient tracking of free surfaces like burrs or the formation of voids. Material loss over the border of the eulerian region should be avoided because it may alter the results.
- The CEL formulation in ABAQUS® is capable of reaching the steady state within a relative short time. Even with a mesh dense enough to track free surfaces properly the calculation time is about 2 days on a 3,4 GHz Intel® i7 processor.
- By the use of the velocity vectors a particle tracking may be possible.
- Because Coulomb’s law assumes that the frictional force is strictly proportional to the normal force and independent of the real contact area or the magnitude of the applied normal force [31] it has to be used with caution. To improve this friction law the use of a shear stress limit according to Orowan [32] may be applied. However it is not clear which stress limit is reasonable regarding influences like strain hardening etc.
- Like already mentioned by Guerdoux and Fourment [18] the simulation highly depends on appropriate constitutive and frictional models, while particularly the last is missing to this date for FSW. Non-linear friction phenomena like the influence of high pressure, velocity and temperature or even the influence of aluminium oxides require intense research. For example it is most likely that at the beginning of the FSW process friction is governed by aluminium oxide and that after a run-in period not only one shear layer determines friction.

Fig. 9. Same setup as in Fig. 5 but with 0.3mm gap between the sheets
ACKNOWLEDGEMENTS

The authors acknowledge support given by Deutsche Forschungsgemeinschaft (DFG) Project DFG RO 651/16-1.

APPENDIX

Table 1. Thermal boundary conditions of the model from [18].

<table>
<thead>
<tr>
<th>Thermal exchange between</th>
<th>Tool [W/m²K]</th>
<th>Backing plate [W/m²K]</th>
<th>Tool shank [W/m²K]</th>
<th>Ambient air [20°C] [W/m²K]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workpiece</td>
<td>50,000</td>
<td>2,000</td>
<td>-</td>
<td>30</td>
</tr>
<tr>
<td>Tool</td>
<td>-</td>
<td>-</td>
<td>20,000</td>
<td>20</td>
</tr>
<tr>
<td>Backing plate</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2. Johnson-Cook constants for AW 6061 – T6, from [35].

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>582</td>
<td>293.4</td>
<td>121.26</td>
<td>0.002</td>
<td>0.23</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Table 3. Temperature dependent material properties for AW 6061 – T6, from [23, 36].

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>37.8</td>
<td>162</td>
<td>945</td>
<td>2.685</td>
<td>68.54</td>
<td>274.4</td>
<td>23.45</td>
</tr>
<tr>
<td>93.3</td>
<td>177</td>
<td>978</td>
<td>2.685</td>
<td>66.19</td>
<td>264.6</td>
<td>24.61</td>
</tr>
<tr>
<td>148.9</td>
<td>184</td>
<td>1004</td>
<td>2.667</td>
<td>63.09</td>
<td>248.2</td>
<td>25.67</td>
</tr>
<tr>
<td>204.4</td>
<td>192</td>
<td>1028</td>
<td>2.657</td>
<td>59.16</td>
<td>218.6</td>
<td>26.60</td>
</tr>
<tr>
<td>260</td>
<td>201</td>
<td>1052</td>
<td>2.657</td>
<td>53.99</td>
<td>159.7</td>
<td>27.56</td>
</tr>
<tr>
<td>315.6</td>
<td>207</td>
<td>1078</td>
<td>2.630</td>
<td>47.48</td>
<td>66.2</td>
<td>28.53</td>
</tr>
<tr>
<td>371.1</td>
<td>217</td>
<td>1104</td>
<td>2.630</td>
<td>40.34</td>
<td>34.5</td>
<td>29.57</td>
</tr>
<tr>
<td>426.7</td>
<td>223</td>
<td>1133</td>
<td>2.602</td>
<td>31.72</td>
<td>17.9</td>
<td>30.71</td>
</tr>
</tbody>
</table>
REFERENCES


Modelling the mechanical deformation of nickel foils for nanoimprint lithography on double-curved surfaces

M. R. Sonne\textsuperscript{1*}, J. Cech\textsuperscript{2}, J. H. Hattel\textsuperscript{1} and R. Taboryski\textsuperscript{2}

\textsuperscript{1}Technical University of Denmark, Department of Mechanical Engineering
\textsuperscript{2}Technical University of Denmark, Department of Micro- and Nanotechnology
*mrso@mek.dtu.dk

ABSTRACT
In the present work, a manufacturing process for transferring nano-structures from a glass wafer, to a double-curved insert for injection moulding is demonstrated. A nano-structure consisting of sinusoidal cross-gratings with a period of 426 nm is successfully transferred to hemispheres on an aluminium substrate with three different radii; 500 µm, 1000 µm and 2000 µm, respectively. The nano imprint is performed using a 50 µm thick nickel foil, manufactured using electroforming. During the imprinting process, the nickel foil is stretched due to the curved surface of the aluminium substrate. Experimentally, it is possible to address this stretch by counting the periods of the cross-gratings via SEM characterization. A model for the deformation of the nickel foil during nanoimprint is developed, utilizing non-linear material and geometrical behaviour. Good agreement between measured and numerically calculated stretch ratios on the surface of the deformed nickel foil is found, and it is shown, that from the model it is also possible to predict the geometrical extend of the nano-structured area on the curved surfaces.

KEYWORDS: Nanoimprint lithography, double-curved surface, injection moulding, nickel foil, finite element model, deformation modelling.

1. INTRODUCTION

Functional nanostructures on double-curved surfaces have attracted increasing attention in the industry. Examples of functional nanostructures are well known from nature, where organisms and plants possess optical, adhesive, and self-cleaning capabilities [1]. The scientific literature is rich in examples of advanced materials emulating the well-known super-hydrophobic effect of the lotus leaf [2] and adhesive surfaces of the gecko’s feet [3]. Structural colors and iridescence are most often observed in invertebrates such as butterflies and beetles, but also in the feathers of birds. Parker did in 2000 [4] study the evolution in animal structural color for 515 million years, and in 2008 Kinoshita et al. [5] studied the physics of such structural colors. Structural color and iridescence may originate from two fundamentally different structures, 1) interference from multiple layers with different index of refraction, or 2) photonic crystals [6] where light is scattered from regularly spaced nanostructures (e.g. pillars or ridges) on the surface. The work presented in this paper is a part of a larger project
called NANOPLAST. The aim of this project is to produce injection moulding tool inserts with nano-patterned functional surfaces as mentioned above. In order to manufacture these structures on the double-curved surfaces of the injection moulding tool inserts, a technology called nanoimprint lithography (NIL) with flexible stamps is an opportunity. In literature, imprinting on simple curved surfaces using flexible materials has recently been demonstrated. Bender et al. (2006) [8] investigated the resolution, dimension stability and reproducibility of the Soft UV-Nanoimprint technique. Y-P Chen et al. did in 2008 [9] present their work on fabrication of concave gratings by curved surface UV-nanoimprint lithography, where a preshaped film was used to provide a uniform pressure distribution throughout the whole concave substrate. Ji et al. (2010) [10] used UV-SCIL (substrate conformal imprint lithography), with a flexible PDMS stamp in their work on photonic crystals patterning in LED manufacturing. In all these cases, the resolution limit due to distortion of the stamp when the pressure was applied, and complications regarding deformations of the flexible stamp, were addressed as major concerns when dealing with flexible stamps on not planar surfaces. Bending and stretching of the flexible stamp is also the main problem when doing imprint on the curved injection moulding tool inserts. As the nanostructures are created on the flexible stamp in its 2-D planar shape, these structures are deformed when the stamp is stretched. In order to take the stretching and deformations of the stamp into account when the stamp is designed, prediction of the 3-D shape is essential. This can be done by numerical simulations. Within the field of nanoimprint lithography, there is however a lack of modelling and simulating the deformation process of the flexible stamp on curved surfaces. First recently, Sonne and Hattel [11] presented a model for the constitutive and frictional behaviour of PTFE flexible stamps for nanoimprint lithography. A nickel foil is often used as the imprinting flexible stamp. The reason for that is the way it is manufactured by electroplating, where the nickel is grown on a silicon wafer with the desired nano-structures on its surface. The nanostructures on the nickel foil are then transferred to the curved surface of the actual substrate via embossing into a polymer resist. The aim of this paper is to address the mechanical deformations of the nickel foil during this nanoimprinting manufacturing process. A series of experiments is performed, where nickel foil is used to make nanoimprint of a predefined nanopattern consisting of a sinusoidal cross-grating with a period of 426 nm, on curved aluminium inserts consisting of hemispheres with different radii. A finite element model is developed for simulating the imprinting process, predicting the distortions and strains acting in the nickel foil during manufacturing. From the deformation of the cross-grating during imprinting, the stretch of the foil is measured and compared with the results obtained from the numerical model.

2. THE MANUFACTURING PROCESS

The route from deciding a specific nano-structured surface to having the final tool insert for the injection moulding machine is long and complicated, and includes many different engineering disciplines. In Fig. 1, the manufacturing process used in this work, is schematically shown. First a glass wafer is coated with a thin layer of photoresist (A). On the backside of the glass wafer, an absorber to catch light is deposited (B), and the photoresist is then exposed with two interfering light sources (C), which create a perfectly regular cross-grating with a predefined period. The photoresist is then developed, leaving the desired nano-structure on the glass wafer (D). A thin layer of chromium is then deposited on this surface (E) in order to make it conductive, so that the nickel foil via electroplating can be formed (F). The insert tool is planarized with a resist called Hydrogen Silsesquioxane (HSQ), which when cured becomes a hard, durable and tough film, ideal for injection moulds. The nickel foil is then by a huge pressure up to 800 bar pressed into the HSQ (H), making the imprint with the
nano-structures. The HSQ on the mould insert is then finally cured, ready to be used for injection moulding of polymer parts (I).

Fig.1: Schematic illustration of the manufacturing process, from preparation of the 426 nm pattern on the wafer (A-D), to electroforming the flexible nickel foil (E-G), further on to nanoimprint lithography into a material called HSQ on top of the final metal substrate (H-I).

3. APPLICATION

The manufacturing process is now applied on three aluminum inserts, which on the surface have a concave part of a hemisphere with varying radius; 500 µm, 1000 µm and 2000 µm, but with a constant height of 200 µm, see Fig. 2. These inserts are designed with the purpose of investigating, how the nickel foil behaves, when it is applied on curved surfaces with a very small radius of curvature. The deformation of the nickel foil is investigated both experimentally and numerically.

Fig. 2: Aluminum inserts with concave parts of a hemispheres with three different radii, but a constant height.

3.1. Experiment

As mentioned, the tested nanopattern is a sinusoidal cross-grating, formed by two exposures using a laser interference method on a 1.5 µm thick layer of photoresist, spin-coated on a 100 mm glass wafer. A polymer backside absorber was used to prevent back reflections of the laser beams into the photoresist. The exposed resists are developed, and the recorded amplitude grating pattern is converted to the surface corrugated phase grating, which is then electroplated with nickel, after coating with the a 60 nm chromium seed layer. The desired thickness of the electroformed nickel foil depends on the accumulated charge in the electroplating process. For this experiment, films of thickness 55±4 µm are used. The
nanopattern period on the nickel foil was measured with an Atomic Force Microscope (AFM) to be 426.2±0.5 nm. Aluminum alloy 6061 was used as material for the tool inserts. The HSQ resist is then spray-coated on to the aluminium surface with an ultrasonic nozzle. Spray coated films consists of hard, almost solvent-free particles and have to be re-saturated with solvent prior to full cross-linking in order to ensure that the resist will flow during the nano imprint. The nano imprint is performed using a special fabricated tool, that via a piston, creates a pressure, which presses the nickel foil into the HSQ surface on the aluminium insert tools, see Fig. 3.

After imprint (60 s at 650 bar) the HSQ film is thermally cured at 400 °C for 1 hour to ensure that the resist is fully cured, as reported previously in literature [12]. A Scanning Electron Microscope (SEM) (FEI Nova NanoSEM 600), operated at low vacuum mode is used to characterize the nano imprinted surfaces. Using top-view and SEM stage movement, the X and Y periods of the replicated nanopattern are measured on the spherical surfaces, in the alternating positions, while increasing the radial off-center distance. The uncertainty for measuring the X and Y periods is calculated to be 0.64 % for the SEM with adjustable stage. The X and Y periods are then used to calculate the mean period for each of the investigated radial positions. Note that geometrical correction needs to be applied to compensate for the increasingly curved position. The real observed period \( \Lambda_{\text{REAL}} \) is then calculated using this correction. This is presented later in Fig. 9.

\[
\Lambda_{\text{REAL}} = \frac{\Lambda_{\text{MEASURED}}}{\cos(\vartheta)} = \frac{\Lambda_{\text{MEASURED}}}{\sqrt{1 - \left(\frac{l}{R}\right)^2}} \tag{1}
\]

, where \( \vartheta \) is

\[
\vartheta = \arcsin \left(\frac{l}{R}\right) \tag{2}
\]

and the \( l \) and \( R \) are distances on a spherical surface, according to Fig. 4.
With respect to repeatability of the experiments, there are as described previously many manufacturing steps in this process, which might lead to a lack of precision on the final product. But as the manufacturing process of the nickel foil, and measurements of the final nanostructures are done with very accurate instruments originally used for production of integrated electrical circuits, the factors affecting the measuring uncertainty are very small. The largest uncertainty comes from the SEM measurement, with a precision of 0.64 %. This is the factor used for the error bars on in the results (see Fig. 9).

### 3.2. Numerical model

In order to support the experimentally obtained results a numerical model is developed. To get a better understanding of the mechanical development during the imprinting process, deformation of the nickel foils used for the three cases with double-curved surfaces in terms of spheres with radii of 2000 µm, 1000 µm and 500 µm, respectively, are modeled through numerical simulations via finite element analysis (FEA). Modeling this deformation is not trivial, as many different non-linearities have to be taken into account in order to get proper results; Large deformations makes the geometrical calculations non-linear, the material acts elasto-plastic, and finally, contact between substrate and nickel foil adds another non-linearity to the system of equations, which also has to be addressed. The nickel foil is fabricated via electroplating, where the glass wafer in a nickel bath solution works as cathode and by means of a voltage potential nickel is deposited onto the wafer. What influence this deposition of nickel has on its mechanical properties is not investigated in this study, instead the mechanical behavior of nickel was found in literature [13]. Nickel has elastic properties in terms of Young’s modulus \( E = 200 \) GPa, and a Poisson’s ratio \( v = 0.31 \). The overall mechanical behaviour in terms of the stress related to the total strain is shown in Fig. 5.

![Stress-strain curve of nickel](image)

Fig. 5: Stress-strain curve of nickel given in both engineering and true values [13].
From this figure it is seen that the material shifts to deform plastically at a yield stress around 200 MPa. For calculation of the displacements, strains and stresses, a standard mechanical model based on the solution of the three static force equilibrium equations is used, i.e.

\[ \sigma_{ij,j} + p_j = 0 \] (3)

for \( i,j = 1..3 \), where \( p_j \) is the body force at any point within the nickel foil and \( \sigma_{ij} \) is the stress tensor. Hooke’s law and linear decomposition of the strain tensor as well as large strain theory, taking the large deformations into account, are applied. The equivalent plastic strain is calculated on basis of standard \( J_2 \) flow theory, with a von Mises yield surface, where the material starts yielding when the equivalent stress (\( \sigma_e \)) reaches the yield stress (\( \sigma_y \)). The equivalent plastic strain increment couples to the deviatoric stress tensor through von Mises yield surface and the associated flow rule. Contact between substrate and nickel foil is in this case modeled via standard Coulomb friction, governed by the equation

\[ F_f \leq \mu F_n \] (4)

, where \( F_f \) is the force of friction exerted by each surface on the other. It is parallel to the surface, in a direction opposite to the net applied force. \( \mu \) is the coefficient of friction, which is an empirical property of the contacting materials. \( F_n \) is the normal force exerted by each surface on the other, directed perpendicular (normal) to the surface. The Coulomb friction \( F_f \) may take any value from zero up to \( \mu F_n \), and the direction of the frictional force is opposite to the motion that surface would experience in the absence of friction. For the contact between the aluminum substrate and the nickel foil, the friction coefficients was found in literature to be \( \mu = 0.35 \) [14]. The layer of HSQ is in this case not taken into account, as it was not possible in literature to find experimental data of its frictional behavior.

The system of equations mentioned in the previous section (eq. (3)) is numerically discretized through a finite element formulation and solved using the general purpose finite element code ABAQUS. In this case a 2-D axisymmetric model is used, where only aluminum substrate and nickel foil is taken into account, see Fig. 6.

![Fig. 6: Geometry in the numerical model.](image)

On top of the nickel foil a pressure is applied, which should end up having a value of 650 bar after 60 seconds. However, this cannot instantaneously be applied in the model as this will result in a diverging solution. Instead the pressure needs to be increased gradually, so static equilibrium in subsequent increments has time to adjust.
4. RESULTS

From the experiments, it is from SEM and AFM now possible to characterize the nano imprinted 426 nm sinusoidal cross-grating on the surface of the aluminium tool inserts, see Fig. 7. It is seen that the nano-structures are very well replicated onto the metal surface. From the SEM the period of the nano-structures can now be measured, and comparison with the numerical model can be carried out.

Fig. 7: Characterization of the nano-structured surface of the aluminium tool inserts with a) SEM and b) AFM, on the $r = 2000 \, \mu m$ curved surface.

From the results of the numerical calculations, it is first of all possible to see the deformed shape of the nickel foil, when it is subjected to 650 bar. In Fig. 8, the deformed shape of the nickel foils in the three different cases is shown with a contour plot of the maximum principal logarithmic strain.

Fig. 8: Contourplots of the deformed shapes of the nickel foils pressed down to the aluminum substrate in the three different test cases: a) curvature = 2000 \, \mu m, b) curvature = 1000 \, \mu m and c) curvature = 500 \, \mu m.
In order to compare the simulation results with the experiments, the maximum principal logarithmic strain is extracted from the model through a path on the surface on the nickel foil. With a base period of the nanostructure of 426 nm, it is possible to calculate the period on the stretched nickel foil by using equation (1). In Fig. 9, a comparison between numerical and experimental results of the grid period is shown. It is seen that the results from the simulations well matches the measured values of periods from the experiments, after they have been geometrically corrected as explained in section 3. The experimental results only show the stretching of the nickel foil, where it was possible to measure the periods of the nanostructures, whereas the results from the FE model further shows the “stretch history” all the way down the spherical cap. From the simulations it is also possible to back out the contact pressure between the nickel foil and the aluminum substrate, see Fig 10. From this figure it is discovered that the contact pressure in the region of the spherical cap is much higher than the applied 650 bar. Moving away from the center, the pressure increases to a peak value and then rapidly decreases to zero. This is where the nickel foil stops touching the aluminum substrate, which in the numerical model results in some pressure singularities. This point is also possible to observe in the experiments, where the nanopatterns on the aluminum insert is no longer replicated from the nickel foil. It is seen, that there is a good agreement between the model and the experiments in where this waste region is placed on the aluminum substrate, which furthermore verifies the setup of the numerical model, and that it with the model also is possible to predict the limitations of making nano imprints on double-curved surfaces, applying the suggested manufacturing technology.

5. CONCLUSION

In the present work, a manufacturing process for transferring nano-structures from a glass wafer to a double-curved aluminium insert for plastic injection moulding has been demonstrated. A nano-structure consisting of sinusoidal cross-gratings with a period of 426 nm was successfully transferred to hemispheres with three different radii, 500 µm, 1000 µm and 2000 µm, respectively. The nano imprint was performed using a 50 µm thick nickel foil, manufactured with electroforming into a glass-like resist called HSQ, with a huge pressure of 650 bar over 60 seconds. During the imprinting process the nickel foil was stretched due to the curved surface of the aluminium substrate and it was experimentally possible to address this stretch by counting the periods of the cross-gratings via SEM characterization. A numerical model for simulating the deformation of the nickel foil during nano imprint was also developed, utilizing non-linear material and geometrically behaviour. Good agreement between measured and numerically calculated stretch ratios on the surface of the deformed nickel foil was found, and from the model it was also possible to predict the limits of the nano-structures on the curved surfaces, with decreasing radii.
Fig. 9: Comparison of measured and numerically calculated periods of the nanostructure in the three different test cases: a) + b) 2000 µm, c) + d) 1000 µm and e) + f) 500 µm.
ACKNOWLEDGEMENTS

As a part of the NANOPLAST project, this work is financially supported by The Danish National Advanced Technology Foundation, which is highly acknowledged.

REFERENCES

International Conference on Advanced Manufacturing Engineering and Technologies

Theme 7

Operation Management

- A Genetic Algorithm For Simultaneous Scheduling Problem In Flexible Flow Shop Environments With Unrelated Parallel Machines, Setup Time And Multiple Criteria
  Jairo H. Torres Acosta, Viviana A. Perea González, César A. López Bello
  Department of Industrial Engineering, Faculty of Engineering. “Francisco José de Caldas”
  District University. Bogotá, Colombia

- The ICTs in the Extended Enterprises
  Arantza Mugica¹, Ibon Zamanillo², Olatz Ukar¹, María Ortiz³
  ¹University of Deusto
  ²University of the Basque Country
  ³Deustotech Institute of Technology

- Reference Model of Manufacturing Resources
  Richard Hedman, Robin Sundkvist, Peter Almström, Anders Kinnander
  Department of Materials and Manufacturing Technology, Chalmers University of Technology,
  Gothenburg, Sweden

- Towards a Downtime Cost Function to Optimise Machine Tool Calibration Schedules
  Abubaker Shagluf, Andrew Longstaff, Simon Fletcher, Paul Denton
  Centre for Precision Technologies, University of Huddersfield, UK

- A case study of long-haulage transport in China for a sustainable transport and manufacturing business
  Danfang Chen, Fredrik Tengdelius
  KTH Royal Institute of Technology, Sweden
A Genetic Algorithm For Simultaneous Scheduling Problem In Flexible Flow Shop Environments With Unrelated Parallel Machines, Setup Time And Multiple Criteria

Jairo H. Torres Acosta, Viviana A. Perea González, César A. López Bello

Department of Industrial Engineering, Faculty of Engineering. “Francisco José de Caldas” District University. Bogotá, Colombia.
jatorres@udistrital.edu.co

ABSTRACT
This paper considers a simultaneous scheduling problem in a static deterministic Flexible Flow Shop environment, where all the stages are made up of unrelated parallel machines. Furthermore, sequence- and machine-dependent setup times are given. A weighted approach is used to create a multi-objective function that minimizes the convex sum of three measures: makespan, number of tardy jobs and the idle time of the productive resources, which according to the theoretical review it hasn’t been considered. For this problem, a mixed integer program is formulated. This is, however, a combinatorial optimization problem which is too difficult to be solved optimally for large problem sizes, hence a Genetic Algorithm combined with a heuristic model is used to find good solutions in reasonable time. The theoretical verification of the proposed procedure allows to conclude that the combined utilization of this approaches outperform the performance of its individual application.


1. INTRODUCTION

In a general sense, Scheduling can be defined as a decision-making process applicable to any manufacturing or service industry that allocate resources to the tasks or operations that have to be done in a given period of time [1]. Given the limitation of those resources, it is necessary an efficient allocation so that the company can achieve one or more business goals.

In some cases which are the midpoint of this paper, the parallel machines are not exactly identical. This is because some may be older than others or work at lower speeds; some may be better equipped for what they are capable of doing a higher quality work. If this situation is presented, then some jobs can be performed in any of the $m$ machines, whereas others can only be performed by a specific group of the $m$ machines.

Given the NP-Hard nature of these types of problems, exact algorithms are efficient only in small instances. Such an approach can take hours or even days to derive a solution. For that reason, researches have focused on using heuristics and approximation algorithms to tackle these situations.
Although Flexible Flow Shop problem has been widely studied in the literature, most of the studies are concentrated on problems with identical processors [2, 3, 4, 5, 6, 7, 8, 9, 10]. Nevertheless, in real world situations it is common to find newer or modern machines running side by side with older and less efficient ones. Even though the older machines are less efficient, they are kept in the production lines because of their high replacement costs and as is logical, they can perform the same operations as the new ones, but require a longer operating time to execute them, for what the jobs with a higher processing time must be placed in the most efficient machines.

In this paper these types of scenarios are considered, in which the workstations are equipped with machines whose speeds depend on the jobs assigned to them.

Weng et al. [11] addressed the unrelated machines problem, with sequence-dependent setup times and an optimization criterion based on a weighted mean completion time. Tavakkoli-Moghaddam et al. [12] proposed a novel, two-level mixed-integer programming model of scheduling $N$ jobs on $M$ parallel machines that minimizes the number of tardy jobs and the total completion time of all the jobs in an unrelated parallel machines context. Given the complexity to find optimal solutions for these kinds of problems, an efficient genetic algorithm is proposed.

Similarly, Jungwattanakit et al. [3] analysed a flexible flow shop scheduling problem, where at least one production stage is made up of unrelated parallel machines. The objective is to find a schedule that minimizes a convex sum of makespan and the number of tardy jobs in a static flexible flow shop environment.

Liaw et al. [13] considered the problem of scheduling a given set of independent jobs on unrelated parallel machines to minimize the total weighted tardiness. A branch-and-bound algorithm that incorporates various dominance rules is presented. Finally, Choi and Wang [14] presented a novel decomposition-based approach, which combines both Shortest Processing Time rule and genetic algorithms to minimize the makespan of a Flexible Flow Shop environment with stochastic processing times.

Once we did a survey of the problem, the rest of the paper is organized as follows: The problem description is described in Section 2, whereas the mathematical model formulated through mixed integer programming is introduced in Section 3. Computational results are discussed in Section 4 and conclusions are given in Section 5.

2. PROBLEM DESCRIPTION

The Flexible Flow Shop, system is defined by a set $O = \{1, ..., t, ..., k\}$ of $k$ processing stages. At each stage $t$, $t \in O$ there is a set $M^t = \{1, ..., i, ..., m^t\}$ of $m^t$ unrelated parallel machines. A set $J = \{1, ..., j, ..., n\}$ of $n$ independent jobs that have to be processed on machines of the sets $M^1 ... M^K$ is defined [3]. Each job $j, j \in J$, has its release date $RD_j \geq 0$ and a due date $DD_j \geq 0$. Assuming unrelated machines, the processing time $p^t_{ij}$ of job $j$ on machine $i$ at stage $t$ is equal to $ps^t_j/v^t_{ij}$ [1, 5]; where $ps^t_j$ is the standard processing time of job $j$ at stage $t$ and $v^t_{ij}$ is the relative speed of job $j$ which is processed by the machine $i$ at stage $t$.

Furthermore, the following processing restrictions are considered: (1) Jobs are processed without preemptions on any machine; (2) each machine can process only one operation at a time; (3) Operations of a job have to be realized sequentially, without overlapping between stages and (4) Job splitting is not allowed.

Considered setup times are classified into two types, namely sequence- and machine-dependent. A sequence-dependent setup time occur when two jobs are assigned to the same machine for what the changeover from one task to another required a specific amount of time.
Meanwhile, a machine-dependent setup time occurs only if a job is the first job assigned to a specific resource.

The scheduling problem under consideration includes a weighted multi-objective function (see Equation 1) that allows through a relative importance factor, aggregate various objectives into a single one. The selected performance measures include makespan, number of tardy jobs and the idle time of the productive resources.

With all of the aforementioned properties, the scenario can be notated as $R_m^{FF}|r_j, d_j, p_{ij}, s_{jk}, ch_{ij}, nwt| C_{max}, n, ld$.

3. MATHEMATICAL MODEL

In this section, a mixed integer linear programming formulation based on [3] and adjusted to the problem under consideration is proposed.

3.1. Notations

The following symbols represent the notation that will be used throughout this paper.

$t$  Stage index, $t = 1,2,3,\ldots, k$

$j, l$  Job indices, $j, l = 1,2,3,\ldots, n$

$m^t$  Number of parallel machines at stage $t$

$i$  Machine index, $i = 1,2,3,\ldots, m^t$

$RD_j$  Release Date of job $j$

$DD_j$  Due Date of job $j$

$In_j$  Inventory holding cost incurred when job $j$ is completed before its due date $DD_j$.

$Pe_j$  Penalty for failure to customers incurred when job $j$ is completed after its due date $DD_j$.

$s_{ji}^t$  Setup time between job $j$ and job $l$ at stage $t$

$ch_{ij}^t$  Setup time of job $j$, if job $j$ is assigned to machine $i$ at the first position at stage $t$

$ps_j^t$  Standard processing time of job $j$ at stage $t$

$v_{ij}^t$  Relative speed of machine $i$ at stage $t$ for job $j$ (dimensionless)

$p_{ij}^t$  Processing time of job $j$ on machine $i$ at stage $t$

$Op_{ij}^t$  Operating time of job $j$ on machine $i$ at stage $t$

$Op_i^t$  Operating time of machine $i$ at stage $t$

$C_j^t$  Completion time of job $j$ at stage $t$

$Init^t$  Initiation time at stage $t$

$Cmax^t$  Maximum completion time at stage $t$

$X_{ij}^t$  Binary variable. 1 if job $j$ is scheduled immediately before job $l$ on machine $i$ at stage $t$. 0 otherwise.

$C_{max}$  Makespan

$T_{dj}$  Tardiness of job $j$

$Ad_j$  Earliness of job $j$

$Ct_j$  Cost of tardiness of job $j$ at stage $t$

$Ca_j$  Cost of earliness of job $j$ at stage $t$

$Id_i^t$  Idle time of machine $i$ at stage $t$
Binary variable. 1 if job is tardy. 0 otherwise.

\( n \) Total number of tardy jobs in the Schedule.

3.2. Mixed Integer Program

The problem can be formulated as follows:

**Objective Function:** Minimize the convex sum of makespan, number of tardy jobs and idle time of the productive resources.

\[
\text{Min} \quad \lambda_1 C^\text{max} + \lambda_2 n + \lambda_3 \sum_{i=1}^{m^t} \sum_{l=1}^{k} l d_i^t
\]

**Subject to:**

\[
\sum_{i=1}^{m^t} \sum_{j=0}^{n} X_{ijl}^t = 1, \quad \forall t, l,
\]

(2)

\[
\sum_{i=1}^{m^t} \sum_{l=1}^{n+1} X_{ijl}^t = 1, \quad \forall t, j,
\]

(3)

\[
\sum_{i=1}^{n} X_{ij0l}^t = 1, \quad \forall t, i,
\]

(4)

\[
\sum_{j=1}^{n} X_{ij(n+1)}^t = 1, \quad \forall t, i,
\]

(5)

\[
x_{ijl}^t = 0, \quad \forall t, i, j,
\]

(6)

\[
\sum_{j=0}^{n+1} X_{ijl}^t = \sum_{j=1}^{n+1} X_{ijl}^t, \quad \forall t, i, l,
\]

(7)

\[
x_{ijl}^t \in \{0, 1\}, \quad \forall t, i, j, l; j = 0; l = n + 1,
\]

(8)

\[
p_{ijl}^t = \sum_{l=1}^{n+1} \frac{p_{ijl}^t}{v_{ijl}} X_{ijl}^t, \quad \forall j, i, t
\]

(9)

\[
C_{ijl}^t - C_{ijl}^{t-1} \geq s_{ijl}^t + \sum_{i=1}^{m^t} P_{ijl}^t + \left( \sum_{i=1}^{m^t} X_{ijl}^t \right) - 1 \right] M,
\]

\[
\quad \forall t, j, l; j \neq l,
\]

(10)

\[
C_{ijl}^t \geq 0, \quad \forall t, j,
\]

(11)

\[
C_{ijl}^t - C_{ijl}^{t-1} \geq \sum_{i=1}^{m^t} \sum_{j=1}^{n} s_{ijl}^t X_{ijl}^t
\]

\[
\quad + \sum_{i=1}^{m^t} c_{il}^t X_{i0l}^t + \sum_{l=1}^{m^t} P_{il}^t, \quad \forall t, l,
\]

(12)
Constraints sets (2) and (3) are sequence restrictions and ensure that only one job is assigned to each sequence position at each stage. Constraints set (4) are sequence restrictions for dummy job 0 and ensure that only one job will be assigned to the first position. Constraints set (5) are sequence restrictions for dummy job \( n+1 \); this job closes the system, for what these conditions ensure that only one job will be assigned to the last position. Constraint (6) assures that after the job has been finished at any stage it cannot be reprocessed at the same stage. Constraint (7) forces to construct a consistent sequence at every stage, that is, that every destination node must be at the same time an origin node for the next scheduled task in the sequence on each machine at each stage, until the dummy job \( n+1 \), which closes the system. Constraint (8) specifies the decision variables as binary variables. Constraint (9) determines the operating time of every job, which is dependent on the standard processing time at stage \( t \) and the relative speed of the machine. Constraint (10) is a set of disjunctive constraints. It states that, if jobs \( j \) and \( l \) are scheduled on the same machine at a particular stage with job \( j \) scheduled before job \( l \), then job \( j \) must complete the processing before job \( l \) can begin. This constraint set forces job \( l \) to start by at least the processing time of job \( j \) plus the setup time for changeover the tasks. The value of \( M \) is set to a very large constant, greater than the sum of all job processing times and setup times, which relaxes the constraint. Constraint (11) declares the completion time of every job at each stage as a positive variable. Constraints set (12) are precedence stage constraints, and establish that a job cannot start its processing on stage \( t+1 \) before it finishes at stage \( t \). Constraint (13) applies only to stage 1, ensuring that a job cannot start its processing in that stage before its release date. Constraint (14) applies only to jobs that are assigned to the first position on each machine, that is, the job
cannot start its processing before machine availability. Constraint (15) links the makespan decision variable. Constraint (16) calculates the earliness or tardiness of each job, that is, the deviations from the due date. Constraints sets (17) and (18) calculate the associate costs of tardiness and earliness of each job, that is, penalty for failure to customers and holding inventory, respectively. If the difference between the completion time of a job and its due date is positive, then the job is tardy. Constraint sets (19) and (20) identifies if each job has this characteristic through the Boolean variable $U_j$ and calculates how many works are tardy in the schedule. Constraints set (21) calculates the operation time of job $j$ in the machine $i$ at stage $t$, as the sum of its processing times and setup times. Constraints set (22) calculates the operation time of each machine $i$ at stage $t$, as the sum of the processing times and setup times of all jobs assigned to it. Constraint sets (23) and (24) determinate the minimum initiation of a job at a particular stage, which is defined by the minimum completion time of a task at the previous station. Finally, constraint (26) calculates the idle time of every machine at every stage, as the difference between the operation time of the stage and the operation time of the machine.

### 3.3. Computational Results

The model was executed in the commercial mathematical programming software, GAMS with an Intel Core i7 2.20GHz CPU with 6GB of RAM.

To evaluate its performance, 7 initial objective functions were constructed, with different $\lambda$ values, where $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $\lambda_1 > \lambda_2 > \lambda_3$, inequation that has been maintained given the higher importance factor assigned to makespan. For the number of tardy jobs and idle time of the resources, the choice is indifferent, but it has been notice through test problems, that the running time of the program exceeds the computer’s memory when $\lambda_3 = \lambda_2$ and $\lambda_3 \geq 0.01$, so that it was necessary to built 6 new objective functions considering this restriction. Moreover, we used a specific amount of machines for each stage in the interval of 2 and 5, so it can be noticed the jobs waiting queue. The quantity of machines can vary from one stage to another, so that initially we take into consideration 7 possible combinations. We have found that an optimal solution can be obtained for instances with up to seven jobs and five stages, requiring maximum 73355.49 s of CPU time, which was considered acceptable.

### 4. GENETIC ALGORITHM

GA, inspired on Darwin’s evolutionary theory, has been recognized as a general search strategy and an optimization method which is often useful to attack combinatorial problems. It was introduced in the 70s by Holland [15] and Davis and Coombs [16] were first to propose it for solving scheduling problems [17]. Its implementation starts with the generation of an initial population of chromosomes, each of whom encodes a solution of the problem and its fitness value is related to the value of the objective function for that solution [17]. Along each iteration or “generation” genetic operators like crossover and mutation are applied to search potential better solutions.

Although is common to randomly generate the chromosomes, if the general process is left free into a random search space, the algorithm convergence is extremely slow [18]. Thus, we decided to include a heuristic method related with the geometry of the problem under consideration and based on the dispatching rules $ERD$ (Earliest Release Date) for the stage $t = 1$ and $SPT$ (Shortest Processing Time) for $t = 2, \ldots, k$, which accelerate its convergence and provide an initial solution to the scenario, while the remaining initial solutions are still randomly generated. The algorithm sticks to the following steps:

1. Sequence the jobs according to their arrival time or Release Date.
2. Assign the jobs to the fastest available machines at stage $t = 1$ using the job sequence.
3. Assign the jobs to the machines at subsequent stages using the FIFO rule.
4. Return the best solution per stage.

It is necessary a re-scheduling between work cells, considering the no-wait restriction, which ensures that there’s no delay between consecutive stages in a job execution. Fig. 1 shows the structure of GA used to find near-optimal solutions to the problem.

**Fig.1. Structure of GA.**

### 4.1. Coding

To represent the chromosomes, we used operations denoted by integer numbers which includes the number of each task and the available machines at each stage in which the processing of the jobs can be done.

### 4.2. Generation of initial population

Like we said, in this paper we include an initial solution through a heuristic method while the other individuals still randomly generated.

### 4.3. Evaluation function

In this paper, we deal with the minimization of the convex sum of makespan, number of tardy jobs and idle time of productive resources.

### 4.4. Crossover

For this problem, a one-point crossover is applied. First, we randomly select the stage to apply the genetic operator, following the no-wait condition and then the cut point in the array for interchange the genetic information.

### 4.5. Mutation

For this problem, we choose a Pair wise Interchange (PI) mutation, which randomly selects the stage to apply the operator and then the gen in the array to make the change.
5. COMPUTATIONAL RESULTS

In our tests, we used problems with up to 20 jobs × 10 stages. The standard processing times are generated uniformly from the interval [10,100]. The relative speeds are distributed uniformly in the interval [0.7, 1.3]. The setup times, both sequence- and machine-dependent are randomly generated between [0,1], whereas the release dates are generated uniformly from the interval between 0 and half of their total standard processing time mean. The due date of a job is set as [3]:

\[ d_j = r_j + \sum_{f=1}^{k} ps_j^f + \text{Total of mean setup time on all stages} + (n - 1) \times \text{Mean processing time of a job on one machine} \times U(0,1) \]  

Firstly, we have studied the GA algorithm with a random initial population to determine the favourable parameters of the approach. In our test, we investigated the influence of different population sizes and rates of mutation and crossover. For all problems we considered instances with \( \lambda_2 > \lambda_3 \times \lambda_3 \leq 0.01 \), according to the results obtained with the MIP. Table 2 shows the tested parameters.

<table>
<thead>
<tr>
<th>Population Sizes</th>
<th>30,50,70,90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crossover rates</td>
<td>0.2,0.4,0.6,0.8,1</td>
</tr>
<tr>
<td>Mutation rates</td>
<td>0.2,0.4,0.6,0.8,1</td>
</tr>
</tbody>
</table>

For the population sizes, we have found that there are not statistically significant differences, but in general a population size of 30 is slightly better. For the mutation and crossover rates, initially we tested the algorithm with a fixed crossover rate of 0.6 based on the results of [3]. It can be notice that there are not statistically significant differences; however we highly recommend a mutation rate of 0.2.

Finally, we tested the crossover rates founding that there are not statistically significant differences, being 0.8 and 1 highly recommended. We decided to use a pure crossover. Fig. 2 shows the obtained results with an instance of 7 jobs × 5 stages and 15 different machine configurations.

![Objective Function Values](image)

Fig. 2. Objective Function Values: Comparison between MIP, Heuristic, GA and Proposed Procedure

It can be noticed that the proposed procedure, which includes as an initial population the heuristic method is slightly better than a Genetic Algorithm with a randomly generated initial
population and the described heuristic method only (see section 4). We have found variation percentages in the objective function of 1.95% maximum regarding the MIP.

Fig. 3 shows the proposed procedure’s performance compared with the heuristic method and the Genetic Algorithm in an instance of 20 jobs × 10 stages and 11 different machine configurations.

![Objective Function Values: GA vs. Proposed Procedure](image)

As we showed in Fig. 3, in all cases it’s better to include in the GA the schedule obtained by the heuristic method as an initial solution. We have found variation percentages of 15% maximum.

### 6. CONCLUSIONS AND RECOMMENDATIONS

In this paper, we have dealt with the construction of a procedure which combines a Genetic Algorithm and a heuristic method for minimizing a convex combination of makespan, number of tardy jobs and idle time of the resources for the flexible flow shop problem with unrelated parallel machines. From our computational experiences, we can give the following conclusions and recommendations:

- The performance evaluation applied to the MIP showed that the GAMS running time increases severely when the importance factor of the idle time is equal or superior to 0.01, restriction that suppress big deviations of this metric regarding the other approximation methods. This increment can be a consequence of the highly combinatorial scenario under analysis, for what giving more importance to this metric apparently represent a more computational complexity regarding the other two. For our experience we can presume that this particular case is strongly NP-Hard. However, this assumption is beyond the scope of this paper.

- It can be noticed that although we have found optimal solutions to the problem through Mixed Integer Programming, its performance can be limited to small instances and requires maximum 73355.49 s of CPU time to converge. In comparison, the proposed method yields good solutions with objective function deviations of 1.95% maximum in 567 s of CPU time. We have found that the complexity of the proposed procedure follows a time complexity $O((n(\max p_j))^k)$.

- The computer evaluation of the 20 jobs × 10 stages instance allowed evaluating the real performance of the proposed procedure. Its fast convergence and best results regarding the Genetic Algorithm with a randomly generated initial solution showed that a Genetic Algorithm combined with a heuristic method outperform the individual application of these approaches.
Further research can be done using different dispatching rules such as LPT, ERD, EDD, MST and S/P to find an initial solution to the problem, verifying which one yield better solutions. Moreover, a comparison with different metaheuristics such as simulated annealing, tabu search, ant colony or a mixture of them can be done to find best solutions in less computing time.

**REFERENCES**


The ICTs in the Extended Enterprises
Arantza Mugica¹, Ibon Zamanillo², Olatz Ukar¹, María Ortiz³

¹University of Deusto
²University of the Basque Country
³Deustotech Institute of Technology
arantza.mugica@deusto.es

ABSTRACT

The incorporation and use of Information and Communication Technologies (ICT) in the Basque Country’s small and medium enterprises (SMEs) business environment is not an option but it is a necessity. This need arises because the evolution of the market which is more and more technologically advanced and ever more competitive.

In this work the small and medium enterprises (SMEs) situation in the province of Gipuzkoa (Spain) versus Extended Enterprise model concept is presented, in order to enable a better company’s management system that will allow to manage not only their own value chain but also to take into account the other agents that make up a company, from raw material suppliers to the end customer.

KEYWORDS: Extended Enterprise, Value Chain, ICTs

1. INTRODUCTION

Over the past few years, companies have not been left out of these advances and they have been incorporating new tools and systems based on new technologies. But the integration degree of the so-called Information Society in their organizations has been very different.

The Information and Communication Systems in the companies have been set out to improve access to the data produced in the company departments and for integrating them in a single system [1, 2]. The evolution of these systems and their intention to integrate the information generated outside its borders allows the emergence of a new business concept that is known as Extended Enterprise.

The Extended Enterprise concept (figure 1) includes two key elements for successful achievement. On one hand, the business vision, the business strategy that goes beyond what is produced in the company’s own facilities, extending it to all the processes that take place outside but that are the key to the survival of not only the customers but also the suppliers. On the other hand, The Extended Enterprise concept refers to the use of new technologies to generate synergies between all the actors of the same value chain, considering that the success of one of them it is also the success of the other members of the chain.
These organizations seek to integrate business processes beyond the corporate barriers considering the other elements involved in the value chain, not only suppliers, customers, external agents, etc., but also its employees.

These companies look for establishing long-term relationships with business partners, but it is critical to define each organization role in the value chain.

The proper operation of each member of the value chain improves the other chain members’ success, so all the members’ of the chain benefit from it [4,5,6,7,8]. The extended enterprise concept requires companies to anticipate and to provide cost implications of the decisions taken in a link of the chain for all other levels, or other links.

In many research works that companies must look beyond their organisational boundaries is also discuss and how the resources and capabilities of suppliers and customers can be used for creating exceptional value is [9,10].

2. OBJETIVES

SMEs have had to change their way of organizing. They have had to establish relationships with other companies and agents. These relationships have result in the emergence of inter-companies networks which involves cooperation between companies.

Thus, this research work main objective it to analyze the situation of Small and Medium Enterprises (SMEs) in the province of Gipuzkoa versus Extended Enterprise model, in order to enable future improvements in the management of companies. The new management systems allow companies not only to manage their own value chain but also take into account the other agents that make up a company, from raw material suppliers to the final customer. So, the incorporation of the Extended Enterprise and the enlarged value chain concepts in the SME aims to increase their competitiveness.

So, the network aim is the profit for the whole value chain. The goal is to improve the quality, efficiency and effectiveness, thus the members or the chain links must work like a network sharing the information and knowledge. This way, each member of the chain is only to be engaged in the matters at which it is best -cross-border issues. Thus, the competitive advantages of each member allow the competitive advantages for all the members of the supply chain.

The development of Information and Communication Systems has sought the unification and coordination of the data produced in the company different departments, and nowadays all of them can be integrated in a single system. The evolution of these integration systems and how the information generated outside company borders is integrated is introducing a "new
deal” in this field known as Extended Enterprise, which was created as an extension of the supply chain management.

The extended enterprise requires foreseeing the consequences as a result of decisions taken at a particular link of the chain on all other levels or links. Thus, each organization must be clear about their role in the value chain.

3. METODOLOGY

This exploratory study is limited to companies in the province of Gipuzkoa in the Basque Country (Spain). The use of information technologies by organizations depends on several variables, and their use is indispensable to apply the Extended Enterprise concept.

Thus, the universe under study (Table 1) will be limited by the following features:
- Organizations placed in Gipuzkoa.
- Organizations with a number of employees greater than 20 and less than 250.
- Organizations that do not belong to agriculture and fisheries sectors.

Table 1. Companies involved in the research work.

<table>
<thead>
<tr>
<th>Industry &amp; energy</th>
<th>Construct.</th>
<th>Trade, hotels and transport Banking, insurances &amp; serv. companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-49 Employees</td>
<td>547</td>
<td>126</td>
</tr>
<tr>
<td>50-99 Employees</td>
<td>173</td>
<td>34</td>
</tr>
<tr>
<td>100-249 Employees</td>
<td>92</td>
<td>10</td>
</tr>
<tr>
<td>812</td>
<td>170</td>
<td>1361</td>
</tr>
<tr>
<td>34.66%</td>
<td>7.26%</td>
<td>58.09%</td>
</tr>
</tbody>
</table>

As it is shown in figure 2, this research work has been divided into four phases. The first phase of this research work has been to transmit, share and discuss the need for changes in organizations to maintain or create competitive advantages, and raise awareness about the Extended Enterprise concept among the participating companies in their two basic dimensions: the technological and the business strategic dimensions.

Fig. 2. Phases for determining the position of the SMEs in the Extended Enterprise model.
To be able to analyze acting different alternatives and come for a decision based on those alternatives, in the second phase for each participating companies a Diagnosis was made. The companies are classified in five levels according to the ICTs implementation in their processes as it is shown in the figure 3.

![Digitalization levels diagram](image)

**Fig.3. Digitalization levels**

On the one hand we collected information about the company regarding the use of ICT and the digitization level the organization was, which was joined the Technological Diagnosis. On the other hand, the business positioning diagnostics was made, in which the company priorities for the medium to long term were studied and discussed.

With all this information the situation of each company and where they wanted to move was known. Finally, for each company a report was made about the potential of the Extended Enterprise concept for their organization, providing them not only information on the strategic and the technological dimension, but also information relating to the potential for Extended Enterprise project that could develop each organization based on the previous Business Analysis.

It is essential to know the status of the organization in terms of ICT available, what it is being used and preparing the company for tackle a project of implementation of new ICTs, if it necessary.

### 4. DATA ANALYSIS

Once the research systematic is explained, in this section the results obtained are presented.

In general, the analyzed organizations the basic tools of information management such as computers, internet and management programs, email, etc. were available (the 97% of companies had desktop computers (82.1% laptops), the 98.5% had Internet access and the 97% used e-mail). However, it can be said that the more degree of specialization the companies involves the less use of ICTs, as it is shown in the figure 4.
Fig. 4. ICTs are being used in the company

However, it should be noted that the use of ICT for communication and exchange of information with external agents was very low (in all of them less than 36%), except for the Electronic Funds Transfer that it was obtained a better result.

The use of ICTs is also different depending on the sector which the companies belong to as it can be observed in figure 5.

Fig. 5. ICTs are being used in the company vs. Companies activities

With regard to the preparation for the use of ICT, which was studied indirectly through the evaluation of aspects on executives, organization managers, employees and some characteristics of the relationship with customers and suppliers, the results in this research indicated that, in general, every company manager was prepared to advance the use of existing ICTs and the incorporation of new tools. This was a strong point that should be used in the evolution towards higher stages of digitization.

Furthermore, the employees of the companies had the technology to do their work optimally and were adequately trained to use technology efficiently, but they were not able to solve basic technical problems that occur in the use of the ICT.

Also, it should be noted that the percentage of companies that use Internet to obtain information about customers and competitors was very high (95.5%). The 94% of them in
used the email communication with clients and the 89.6% of them had an Internet website with information about the company and the products or services offered.

However, only the 61.2% of companies were collecting information about customers in the databases used throughout the organization.

So it can be concluded that the use of ICT in customer relations was limited to information communication and storage. That is, in most cases the phone had been replaced by electronic mail and paper by database electronic files.

As ICTs that could be used in relationships between companies and customers became more sophisticated and specialized, the use of them decreased.

In all cases the company web was used simply as a source of information or companies’ presentation card of their products and/or services.

The use of ICTs in relations with suppliers was still lower than with customers. They just use email for communication and internet to search for information about suppliers already working with the company. It also is important to say that only in the 71.6% of the considered companies was available an information system for the warehouse management.

The 100% of the companies were had computers with word processors and spreadsheets, and in the 91% of the companies analyzed their employees used the e-mail for the working relationships within the company.

The use of ICTs in relationship with banks, Administration and other agents is more significant with respect to customers, suppliers and employees.

Companies in general used the ICTs to exchange information with those external agents, although the dump of generated information in these operations was not automatically.

From the collected data it can be conclude that all the participating companies were above the second level. As is it is shown in the figure 6 the 59.7% of them were at the Level III (Interaction), the 37.3% at Level IV (Transaction) and another 3% in the Level V (Digitization). Therefore, it is important to note that companies are well placed toward greater stages of digitization or to the Extended Enterprise.

![Fig.6. Digitalization level of the companies](image)

About business competitive strategy (figure 7), the 39% of participating companies in the research work were looking for the growth of the organization; the 37% were looking for a strategy of cost reduction, and the 24% were looking for the differentiation.

![Fig.7. Competitive strategy](image)
Talking about the relationships with ICTs in the figure 8 is shown that the 65% companies were looking for be nearer to their customers, the 20% with providers, the 12% for the relationship with their own employees and the 3% were interested in other agents.

And regarding to the position occupied by the participating companies in the value chain, the 48% of them played a dependent role, the 36% had a dominant role, and the 16%, represented a regulatory-prescriber role (figure 9).

5. CONCLUSIONS

Most supply chains environment is very dynamic so the collaboration is a fundamental strategy for sustainable and profitable growth. In this work the small and medium enterprises situation in the province of Gipuzkoa (Spain) versus Extended Enterprise model concept is presented.

Extended Enterprise model concept enables a better company’s management system because it allows to manage not only their own value chain but also to take into account the other agents that make up a company, from raw material suppliers to the end customer.

The market form SMEs is more and more technologically advanced and ever more competitive. So, the incorporation and use of Information and Communication Technologies (ICT) in small and medium enterprises business environment is not an option but it is a necessity.

Through the field work it has been concluded that the SMEs know the possibilities that ICTs offer as management tools and the companies are well placed to adopt a scalable, higher value-added applications on relationships with customers and suppliers.

But the research work universe does not apply the potential of the studied management tools that they had implemented. They do not apply due to financial or technical considerations. So the IS’ development is limited.
The main objective of SMEs in its Value Chain extension is the customer, but in practice the application that is made of the tools does not match with that objective. That is because there are not enough resources to tackle projects to promote the use of tools for managing customer relationships.

So it can be conclude that although the starting point is good, there is a long way to go towards the Extended Enterprise.

REFERENCES

Reference Model of Manufacturing Resources

Richard Hedman, Robin Sundkvist, Peter Almström, Anders Kinnander

Department of Materials and Manufacturing Technology, Chalmers University of Technology, Gothenburg SE 412 96 Sweden
richard.hedman@chalmers.se

ABSTRACT

This article presents a resource reference model that can be incorporated to cost efficiently explain and quantify a manufacturing firm’s hidden productivity potential in currently used human and equipment resources. That information can be used as input to improve the firm’s financial parameters based on shop floor improvements. The representations of manufacturing resources are to some extent founded on the international standard for manufacturing data management (ISO 15531).

KEYWORDS: manufacturing resources, productivity, capacity, shop floor improvements

1. INTRODUCTION

Manufacturing systems are complex socio-technical and multi-scale systems consisting of objects, such as activities and resources, and their interdependencies. Manufacturing companies should always strive to utilize available resources in the most efficient way. That requires a deep understanding of resources’ constitution and behaviour on all levels of a manufacturing system hierarchy, starting bottom-up from the shop floor level.

This paper presents one part of an on-going research project aiming at the development of a decision-support tool that explain the financial effects of shop floor productivity improvements, giving decision makers better information for doing shop floor productivity investments. The focus lies on a reference model of manufacturing resources where the representations of resources are based on the international standard for manufacturing data management (ISO 15531) [1]. Alexander et al. [2] defines three characteristic attributes for reference models: universal applicability, reusability meaning that generic conceptual patterns can be used repeatedly, and that reference models contains best practices in order to provide recommendations for conducting business. The proposed reference model is expressed using the Unified Modeling Language (UML) which is the industrial standard for object-oriented notations [3].

In simple terms, a manufacturing system can be described as a set of processes with the central objective of transforming raw material into finished products or components. System improvement actions are typically focused towards maximizing throughput, while at the same time minimizing inventory and operating expenses, i.e. the ongoing costs for running a
Several approaches for calculating the costs of manufacturing can be found in the literature, for instance:

- Activity Based Costing (ABC) [5] and Time-Driven Activity Based Costing (TDABC) [6].
- Throughput accounting [7].

The cost accounting models ABC and TDABC and throughput accounting are used to capture costs that occur during the manufacturing phase [8]. Throughput accounting is regarded as a short-term cost behavior model based on given constraints while ABC and TDABC are referred to as a long-term cost behavior models based on resource usage [9].

Cost models influence a firm’s strategic objectives and subsequently a firm’s investment decisions and how operations are managed. To find ways to free capacity and to improve flexibility in order to increase sales is often more interesting to manufacturing firms than cutting direct salary costs. The intended area of application for the proposed reference model is not however, to assess strategic management thinking, nor advocate a certain strategic mind set. The goal is instead to cost efficiently describe and understand a firm’s hidden productivity potential with the current human and equipment resources and how they can be used in the future to improve financial parameters. The TDABC approach is suitable for this purpose and is partly integrated and further described in the resource reference model.

Following sections will describe what kind of data to measure and how the information can be organized in order to link manufacturing resource information to the effects of shop floor productivity improvements.

### 2. PRODUCTIVITY MEASUREMENT AND IMPROVEMENT

The information incorporated in the resource reference model is based on data measured directly at the shop floor level. However, manufacturing systems are subject to endless variability. Therefore it is not credible to assume that a manufacturing system can reach an ideal state, but it is possible, and indeed necessary, to measure and model capabilities against an ideal state [10]. Productivity measurement and improvement is described for activities performed by humans and equipment.

Productivity at an activity level can be improved through better methods (M), increased performance (P), and increased utilization (U). This is conceptually expressed in the following equation [11]:

\[ \Delta \text{Productivity} = M \times P \times U \]  

The method factor (M) is defined as the ideal or intended productivity rate. It is the inverse of the ideal processing time for the specific work task. In order to determine the ideal processing time for manual work tasks it is necessary to use a predetermined time system. There are a number of available systems and most of them are based on MTM [12]. The time for the work task can then be timed with stop watch, but the resulting time will not be the ideal processing time; it will be affected by the P and the U factor in equation 1.

The performance factor (P) corresponds to the speed the work is carried out at in relation to the ideal processing time. For manual work the performance factor can be both below and above 100%. The normal speed in MTM is set to be valid for a “normal” person working at this speed for 8h a day and for the whole working life without getting exhausted or injured. The performance rate is lower for not fully trained workers and for people with disabilities.

The utilization factor (U) represents the time that is spent on performing the intended work in relation the total planned time. Utilization can never go beyond 100%. The planned
production time is usually defined as the paid working time minus planned stops, such as weekly meetings or planned maintenance stops. The U-factor for manual work can be measured through a work sampling study [12].

To be able to use P and U as input to modelling of manufacturing resources it is necessary to specify the different P and U losses and divide them into several separate variables defined by Almström [13] and shown in table 1.

Table 1 Performance and utilization definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal performance rate (Pₚ)</td>
<td>The personal performance rate is affected by the individual’s physical ability and his or her motivation to work at a high speed (relative the MTM norm), independent of work task.</td>
</tr>
<tr>
<td>Skill based performance rate (Pₛ)</td>
<td>The skill based performance rate is the individual’s speed at performing a specific work task depending on the training and the experience the individual has for the task.</td>
</tr>
<tr>
<td>Need based utilization rate (Uₙ)</td>
<td>The need based utilization rate depends on the need for relaxation and personal time. It is often regulated by agreements at the work place. It includes paid breaks and losses before and after a break.</td>
</tr>
<tr>
<td>System designed utilization rate (Uₛ)</td>
<td>The system designed utilization rate is defined as the balance losses designed into the system. It can be balance losses on an assembly line as well as losses in a semi-automated work station.</td>
</tr>
<tr>
<td>Disturbance affected utilization rate (U₀)</td>
<td>Disturbance affected utilization rate corresponds to the losses caused by different random disturbances. It includes the lost time from discovery of the disturbance until the work is performed at full speed again.</td>
</tr>
</tbody>
</table>

The effectiveness of automatic activities can be measured using overall equipment effectiveness (OEE) [14]. The basic definition of OEE is the ratio between the time spent on producing goods of approved quality to the scheduled time (loading time). OEE is not the same thing as the U factor in equation (1). It is U multiplied by P (performance) and quality (i.e. yield expressed as 100% - scrap rate). It does not include the M-factor, instead the ideal time duration for activities performed by equipment is derived from computer aided manufacturing (CAM) software, equipment specifications, or similar. The OEE measure is affected by the surrounding system, e.g. the manning of the machine [15], it is therefore of special interest to analyse the combination of utilization of machine and operator.

It is important to differ between utilization and capacity. Utilization is always in relation to the planned, intended, paid, or manned time. It is always measured as a percentage. Capacity is measured as products per time unit. Two different capacities are used in this paper: Planned and Real (table 2).

Table 2 Capacity definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned capacity</td>
<td>Throughout a work shift, the company’s MPS (master production schedule) system is allocating specific resources for specific activities. This time is usually referred to as planned process or production time.</td>
</tr>
<tr>
<td>Real capacity</td>
<td>The real or &quot;practical capacity&quot; takes into consideration the losses based on the P and U factors.</td>
</tr>
</tbody>
</table>
3. PRODUCTION SYSTEM MODEL

In this section manufacturing resources are defined within the context of a production system, as shown in figure 1. Since the interrelations between manufacturing resources and other production system entities are vital for resource modelling, the resource reference model represents a fundamental part of the production system model.

![Figure 1. Production system model](image)

3.1. Facility

A factory represents the actual manufacturing facility and is the top system level of the model. It can be broken down into subsystems which correspond to defined areas of the manufacturing facility, e.g. the storage area, the painting area, or the assembly area, etc. A subsystem consists of one or several workstations which are defined areas within the subsystem. The different system levels in this hierarchy are subclasses to the entity Facility.

3.2. Manufacturing processes

In a Facility one or several Manufacturing processes are executed. The hierarchal composition of the production system definition enables a manufacturing process to be described from the views: Factory, Subsystem, or Workstation. Hence, a Manufacturing process can be seen as the entire process of converting raw material into finished products (Factory view) or as a specific set of activities performed in a Subsystem or at a Workstation. The entities Manufacturing resource (with subclasses Equipment and Human) and Manufacturing process are defined as in ISO 15531. The decomposition of the entities Facility and Activity (in Figure 1) are not part of the standard.

3.3. Activities

An activity consists of sub-activities that constitute a specific part of an activity, expressed as a sequence of elements. For example: count components, put components in box, deliver box from position A to B. Activities are organized in state categories defined by Czumanski and Lödding [16] where activities occurs with the processing of every part (cycle
bound), once for each batch (batch bound), or periodically (planned), depicted in figure 2. The elements in manual activities are standard movements, such as get, put, use etc. defined in a predetermined time system such as MTM [12]. It is possible to define elements and sub-activities using rougher time studies or even estimates dependent on the desired level of detail.

![Figure 2. Definition of activity categories](image)

In the production system definition are activities formulated as time equations, which are directly elaborated from Time-Driven Activity Based Costing by Kaplan and Anderson [7]. In a time equation each sub-activity can be assigned a time driver and the time consumption per sub-activity is the sum of element times for that sub-activity. If a product family requires a completely different set of activities to be produced, it should be defined in a separate manufacturing process.

4. RESOURCE REFERENCE MODEL

The main class modelled in the resource reference model is naturally Manufacturing resource and it consists of the subclasses Human and Equipment. As stated in the ISO 15531 definition [1] a resource is “any device, tool and means, excepted raw material and final product components, at the disposal of the enterprise to produce goods or services”. That means that everything from manufacturing personnel and machines, to pallets and fixtures can be considered as manufacturing resources. The purpose of the resource reference model is to visualize and quantify the improvement potential of manufacturing resources.

4.1. Resource roles

This role modelling approach uses generic properties, meaning that resources can assume new roles and resign roles dynamically. Roles can also be shared by different types of resources (equipment and human). Figure 3 shows the associations between the entities resource, role, and activity. A manufacturing resource can be assigned to one or several roles, which in turn is required by one or several activities.

![Figure 3. Manufacturing resource role associations](image)
Roles for humans can be described in relation to the manual activities that the human manufacturing resources are performing, for instance assembler, operator, or technician. Equipment roles can analogously be described in roles such as machine, transport conveyor or associated to other types of automatic activities. For the purpose of this reference model that categorization of roles is, however, irrelevant. Instead, the proposed role description is associated with the performance measurement dimensions; internal efficiency and flow efficiency, as defined by Jonsson and Lesshammar [17]. The internal efficiency is referred to the actual resource usage where the objective is to identify the performances of a function and is of interest for all resources, and thereby all resource roles. The flow efficiency dimension is incorporated in the role definition by distinguishing how the resource contributes to the product flow, i.e. the throughput time.

As stated earlier, a manufacturing system can conceptually be described as a set of processes transforming raw material into finished products or components. Each process consists of activities and these activities drive the transformation of material forward (downstream) either direct or indirect.

A resource has a *direct role* if it is assigned to perform an activity that is positioned direct in the product flow and thereby has a direct effect on the throughput time. That includes, for instance, activities where material:

- undergoes transformation (i.e. machining or joining)
- is being transported downstream
- undergoes inspection activities (i.e. quality control or testing)

A resource has an *indirect role* if it is assigned to perform an activity that is decoupled from the direct product flow and thereby only has an indirect effect on the throughput time.

It is important to note that activities can be identical, independent of if the resource is assigned to a direct role or an indirect role. Consider a material supply activity as an example; if a resource (human or equipment) is transporting material to a transforming activity (i.e. machining) while the transformation is being performed, then the transporting resource has an indirect role and the transforming resource a direct role. However, if the material supply activity is placed in series with the processing activity or if it is identified as a capacity constraint, i.e. starving the transformation activity, then both the transporting resource and the transforming resource have direct roles. The distinguishing of direct and indirect roles depends on the design of the production system and the planning of its operations. Therefore, when modelling a production system on site the direct or indirect roles of resources are assigned on a case-to-case basis. The distinction of roles is important in order to be able to analyse the combination of machine utilization and operator utilization, and to evaluate possible improvement scenarios with respect to flow efficiency and internal efficiency.

### 4.2. Resource characteristics

Each resource, equipment and human, is described using the resource characteristics as defined in ISO 15531 [1]. The resource characteristics are classified in a resource characteristic group, depicted in figure 4.
The entity *resource_administration* specifies the resource’s cost per time unit including potential depreciation (for equipment). Individual resources are identified based on name or reference number, not by role.

The entity *resource_capability* has a list or a reference to a list, to what activities the resource can perform and consequently that comprises a specification of the activities the resource can execute. The capability of the resource can be further classified using performance related attributes ($P_P$ or $P_S$) for each activity. For equipment $P_P$ corresponds to the performance of individual equipment in relation to the specified equipment design speed, $P_S$ is not applicable for equipment.

The capacity of the resource, described in the entity *resource_capacity*, comprises information about the potential workload of manufacturing resources. It is expressed as planned capacity ($\text{CAP}_{PL}$) which corresponds to the resource’s working schedule and consequently its availability. For equipment, it corresponds to the planned production time.

The entity *resource_constitution* is not applicable for human resources since it primarily concerns equipment related attributes such as functions, tolerances, and technical specifications [1].

### 5. SHOP FLOOR IMPROVEMENT MODELLING USING RESOURCE INFORMATION

Resource information can be combined in many ways dependent on the purpose for which it is needed. The reference model is intended to be implemented in software to provide decision support for assessing the improvement potential of production systems, for instance during rapid scenario analyses, or to provide input data to more advanced simulation tools.

As stated, shop floor improvement initiatives are typically concentrated on maximizing throughput while at the same time minimizing inventory and operating expenses [4]. This can be put in relation to the performance measurement areas internal efficiency and flow efficiency presented by Jonsson and Lesshammar [17]. According to Modig and Åhlström [20] there is an efficiency paradox stating that over-focusing on resource efficiency will have negative effects on flow efficiency. The paradox thereby says that while believing that resources are utilized efficiently the real system performance, i.e. throughput time, could in fact be inefficient. By incorporating the direct and indirect roles of resources it is possible to distinguish between internal efficiency and flow efficiency when modelling the manufacturing system.

#### 5.1. Modelling approach

The system boundary is a firm’s factory walls. Initially, the system is decomposed into small components, all the way down to individual activities. Each activity is modelled according to the defined production system structure (figure 1) consisting of sub-activities and
elements. Each activity has as stated an ideal time consumption in relation to the work standard. The inverse of the ideal time consumption thus constitutes the ideal productivity rate of the activity (the M-factor). The term ideal is not to be confused with optimized. It shows the current standard of how the work is performed, i.e. according to a work instruction. There are always hidden improvement potentials on an activity level [11]. Since the activities are built up by elements and sub-activities the potential on activity level is no longer hidden but instead visualized and quantified.

Modelled activities are reintegrated into the general system by arranging them into manufacturing processes and according to product routings. The composition of individual activities becomes a manufacturing process when resources are assigned to perform the activities. The manufacturing processes are executed on the hierarchal levels workstation, subsystem and factory. It is then possible to measure the resource utilization (U) and performance (P). The internal efficiency for resources can consequently be measured for both direct and indirect roles.

Performance losses in activities are derived from the resources’ capability to perform the activity based on skill and motivation (P_S and P_P) for humans or reduced speed losses for equipment, defined in the characteristics attribute resource_capability. These performance losses do not include variability, which is shown when measuring utilization. A utilization of 100% means that the resource spends 100% of the planned time on performing the activity. Naturally, this is never the case. There will always be utilization losses due to system design, disturbances and personal needs (for humans) as defined in table 1.

The modelling and analysis is constraint-based. The real capacity (CAP_R) on a workstation level is determined by the real capacity of the constraining activity which is calculated by multiplying the M-factor of the activity with the P- and U-factors. Subsequently the real capacity of a manufacturing process on a subsystem level is determined by the real capacity of the constraining workstation and lastly, the real capacity of a manufacturing process on a factory level is determined by the constraining subsystem. Since the resource’s time consumption, for both direct and indirect roles, is modelled it is possible to calculate the cost of capacity and to distinguish how that capacity is distributed between internal efficiency and flow efficiency.

The higher the level of abstraction, the more complex it gets and more unexpected synergies will occur. It is however the authors’ strong believe that trustworthy system descriptions, in the context of shop floor productivity improvements, can only be built using a bottom-up approach where the composition of the smallest system components are based on facts and not assumptions.

5.2. Focusing improvement efforts

After a current state of a system (factory, subsystem or workstation) has been established its manufacturing processes are expressed both as the capacity in the current ideal state (norm times and no losses), and as the real, or practical, capacity, (CAP_R) taken into consideration the losses. The definition of an ideal state is vital since only then are improvements towards it possible. A current ideal state can also be developed into a future ideal state by doing method improvements.

Two typical shop floor improvement scenarios are presented to show how improvement efforts can be organized based on the parameters and relations of the reference model.

Increasing throughput

On a workstation level throughput is increased by improving the ideal productivity rate of the constraining activity. An example of improving the method, i.e. the ideal productivity
rate, of an activity is to change from a manual screwdriver to an electrical screwdriver. The activity may be several times more efficient because of the change in method. A method improvement has the largest effect on \( \text{CAP}_R \). However, the higher up in the system hierarchy the larger impact of utilization (\( \text{U}_N \), \( \text{U}_S \), and \( \text{U}_D \)) and performance (\( \text{P}_S \) and \( \text{P}_P \)) rates. A throughput improvement on a subsystem or factory level can also be achieved by improving \( \text{CAP}_R \) of non-bottleneck workstations. Improvements of \( \text{CAP}_R \) both upstream and downstream could reduce starving or blocking of the bottleneck and consequently reduce the system design based utilization losses (\( \text{U}_S \)). Throughput improvements logically imply a focus on flow efficiency and consequently a focus on resources in direct roles.

**Reducing inventory**

Inventory can be reduced by decreasing batch sizes and buffer sizes. Lower inventory levels will typically make a production system more sensitive towards disturbances. That will in turn have negative utilization effects on resources in both direct and indirect roles.

In order to enable positive effects of smaller batches, such as increased flexibility, efficient setups are required. Improving a setup activity follows the same approach as when improving a constraining activity, i.e. by improving the \( \text{CAP}_R \) of the activity.

Smaller batch sizes and reduced buffer sizes imply a reduced amount of work-in-process (WIP). Little’s law [19] shown in equation 2 shows the relationship between throughput rate, cycle time and work in process (WIP) and it is valid for both individual workstations and systems as a whole [18].

\[
\text{WIP} = \text{throughput rate} \times \text{cycle time} \tag{2}
\]

WIP can consequently be reduced by decreasing the cycle time, where cycle time is defined as the time a unit spends within a system (workstation, subsystem or factory). The cycle time at a workstation consists of processing time and waiting time (also referred to as queuing time). In the reference model, the processing time corresponds to the activity’s ideal time consumption and is thus composed by sub-activities and elements. Waiting time is caused by utilization losses and variability. Method improvements shall then be concentrated on activities performed by direct resources in order to reduce the processing time. Waiting times can be reduced by looking at the subsystem or factory as a whole and improve system design aspects (\( \text{U}_S \)) such as improved synchronisation and increased station overlap time. Waiting times can also be reduced by improving the handling of disturbances (\( \text{U}_D \)), for instance by reducing repair times i.e. improving the internal efficiency of indirect resources.

It is not suggested that certain improvement actions are to prefer. There are multiple interrelations and it is therefore not desirable, or even possible, to focus solely on one of them. If the effects of an improvement action are beneficial or not must of course be evaluated from case-to-case.

### 6. CONCLUSION

This article have presented a generic reference model and proposed a modelling approach that constitutes one step towards a decision support for assessing the improvement potential of production systems. Improvement potentials are visualized and quantified all the way down to the lowest level of shop floor activities. Thereby, trustworthy bottom-up descriptions of a production system’s improvement potential can be created.
ACKNOWLEDGEMENTS

This research is carried out within the Sustainable Production Initiative and the Production Area of Advance at Chalmers University of Technology. The support is gratefully acknowledged by the authors.

REFERENCES


Towards a Downtime Cost Function to Optimise Machine Tool Calibration Schedules

Abubaker Shagluf*, Andrew Longstaff, Simon Fletcher, Paul Denton

Centre for Precision Technologies, University of Huddersfield, UK
abubaker.shagluf@hud.ac.uk*

ABSTRACT
A high degree of accuracy of production machines, especially machine tools, is often highly desirable. Understanding capability, maintaining or improving the performance of the machine requires a periodic full calibration. The idea of maintaining machine accuracy by regular calibration is analogous to using predictive maintenance methods to reduce the risk of machine breakdown. Studies in the field of predictive maintenance have resulted in cost calculations for the downtime associated with machine failure.
Models have been presented to determine optimal intervals between repairs by minimising global maintenance costs. However, very little work has concentrated on optimising the frequency of machine tool calibration by assessing the downtime cost considering the contribution of both technical and commercial factors.
This paper presents a cost function that forms the basis of a strategy for scheduling machine tool calibration which takes into account these influences on part tolerance.

KEYWORDS: Machine tools; Predictive maintenance; Predictive calibration.

1. INTRODUCTION

Machine tool failures in industrial organisations interrupt production operations and cause production loss, which has a direct cost-to-business and potentially a significant detrimental impact to future production. Predictive maintenance is one approach that has been successfully applied to mitigate the effects of unexpected failure by scheduling controlled production stoppages [1], rather than reacting to a breakdown. Predictive maintenance is a tool that has been adopted in some industries to improve operational efficiency and reduce maintenance cost [2]. As a result, monitoring equipment that provides information about the condition of manufacturing systems has evolved rapidly over recent years.
Calibration is a fundamental accepted process required to maintain the quality of measuring machines [3]. It can also be applied to the production process to help control output quality and maintain the credibility of the machine tool for measurement, such as in-process probing [4]. Full machine tool calibration requires measurement of a significant number of
error sources, taking up to two weeks on large machines [5]. The reason for repeatedly calibrating an instrument, machine tool or any other machine is that their performance can drift over time and usage in both their mechanical and electrical response. When considering machine tool accuracy, bedding in, wear of components and collision are some reasons for this change. At present, the prescribed interval between calibrations tends to be subjective; a fixed “annual” calibration is sometimes adopted as part of a quality paper-trail, but more likely calibration is undertaken as a reaction to change in the consistency of the machine’s output. Building a database of inspection history by measuring the machine on a regular basis with relatively non-invasive methods will make the decision of scheduling the more extensive calibration a better-informed process.

The frequency of calibration is a compromise between the desire for a high production rate and the need to maintain quality. This situation is sometimes exacerbated when production and quality departments do not operate in tandem. This leads to a potential conflict of interests where machine tool calibration is perceived as lost production time (a reduction in short-term production rate) rather than an investment in time to allow consistent, long-term production quality and enhance the overall equipment effectiveness (OEE).

A strong case is required to justify the period of downtime for machine tool calibration if the machine is producing the quantity of parts desired within their nominal tolerance. This paper proposes a new method for maintaining machine tool accuracy that is complimentary to the predictive maintenance paradigm. This strategy, called predictive calibration, is a methodology that depends on the prediction of the degradation in machine tool accuracy based upon regular data capture.

This piece of work presents a new method of identifying new boundaries of machine tool working tolerance. These boundaries of tolerance reflect the degradation level corresponding to production capacities and the quality of the part produced. This work derives a downtime cost function that can be used to optimise the frequency of calibration to reduce unnecessary downtime while maintaining the machine at the required tolerance.

2. PREDICTIVE CALIBRATION AND DECISION MAKING

In some production, machine tool repairs are conducted only when the machine breaks down or otherwise ceases to meet its key performance indices (KPI’s). The strategy is more “do not fix it until it breaks” than maintenance and is known as Corrective Maintenance (CM). Reducing unexpected machine tool downtime and assuring quality have become increasingly important as the demand for higher volume of production and “just in time” manufacturing has increased. Consequently, the adoption of periodic maintenance has evolved.

One solution is Preventive Maintenance (PM), which replaces parts before they fail, often using a probability model. However, this technique can be wasteful since healthy parts will be discarded and downtime for replacement may not be optimised. Predictive Maintenance (PdM) or Condition Based Maintenance (CBM) is a strategy that includes feedback of the instantaneous condition of the machine and detects degradation before a fault becomes critical.

Predictive Calibration (PdC) is a new methodology proposed to be analogous with, or indeed a subset of, a PdM strategy. It is intended to be a formalised approach applied to machine tools to measure and monitor any degradation in the mechanical parts to assist with maintaining the KPI of positioning accuracy, while having the added-value of revealing other maintenance issues such as wear in ball-screws, guide-ways, impending bearing failure, etc. Although inspired by PdM, accuracy is difficult to monitor “live” with available technology so a periodic approach is required. It is therefore necessary to apply the necessary technical knowledge along with management strategies and decision making skills [6].
2.1. **Metrology Equipment and benchmarking**

As discussed, PdC can be used as part of a hybrid maintenance strategy. However, the negative factors are the cost of the metrology equipment needed and the necessary skilled labour and training costs required to use them effectively. Additionally, such measurements can only be taken when the machine is not producing parts, thus the opportunity cost must be considered. Establishing an optimised PdC strategy is a non-trivial task that must be rolled out as a controlled process programme, taking into account the available technology and their relative merits. Table 1 provides brief comparison between different calibration and measurement approaches.

**Table 1. Comparison between calibration approaches [7]**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Quick check tests</th>
<th>Full calibration</th>
<th>On-machine artefact probing</th>
<th>Post-process measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Typical Duration</strong></td>
<td>30 minutes [8]</td>
<td>2 to 5 days [9]</td>
<td>5 to 10 minutes</td>
<td>A few hours [10]</td>
</tr>
<tr>
<td><strong>Target</strong></td>
<td>Measure and monitor</td>
<td>Measure and compensate</td>
<td>Check, analyse and rework to increase part quality</td>
<td>Inspect the work piece</td>
</tr>
<tr>
<td><strong>Environment</strong></td>
<td>Workshop environment</td>
<td>Workshop environment</td>
<td>Workshop environment</td>
<td>Controlled environment</td>
</tr>
<tr>
<td><strong>Data suitable for comparison</strong></td>
<td>Statistics and process control</td>
<td>Statistics and process control, More skilled interpretation</td>
<td>Statistics and process control</td>
<td>Statistics and process control</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>Occurs while the machine is running but machining process interrupted</td>
<td>Occurs while the machine is out of production</td>
<td>Performed as part of machining procedures</td>
<td>Occurs after machining and off the machine</td>
</tr>
<tr>
<td><strong>Access</strong></td>
<td>Operator</td>
<td>Skilled</td>
<td>Operator</td>
<td>Skilled</td>
</tr>
<tr>
<td><strong>Risk of missing important data</strong></td>
<td>High risk due to low coverage</td>
<td>Low risk due to high coverage</td>
<td>Low to Medium risk depending upon relevance of artefact to part</td>
<td>Low risk</td>
</tr>
</tbody>
</table>
2.2. Predicting machine tool accuracy non-conformance

Degradation of the machine condition with time can be extracted by analysis of machine history data. Such an inspection history may be provided by artefact probing, double Ballbar or any other appropriate measurement technique. Downtime cost must be traded-off against richness of data. Establishing relevant KPI’s with appropriate tolerance, then monitor against them can help to give a clear image of how many good parts of a production machine could be produced before its accuracy drifted. When machine tool degradation level crosses the accepted accuracy margin, then the machine needs intervention. Analysis of data is the basis of decision making for the machine to be corrected.

Some metrology equipment provides such measurement flexibility. For instance, Renishaw’s QC-20W Ballbar system can display a history graph of some parameters’ variations along the time of testing. It is important to know which parameters are key performance indices KPI’s for the machine such as; reversal, squareness, etc. and how to interpret the values will reference to external factors and measurement uncertainties.

![Graph showing performance indices over time](image)

**Fig. 1. XY-Plane (360° 1000 mm/min feed rate) Reversal Spikes X vs. Temperature**

Fig. 1 shows machine tool reversal spikes $X$, and Fig. 2 shows the same machine squareness test recorded for a period of seven weeks using a ball bar. Each test required 30 minutes of production downtime. Predictive maintenance process relies on the data analysis of the rate of change of the measured errors.

Here, the environment should be taken into account when deciding how frequently to measure. Consider if only two tests had been conducted at the start and end of this period. Without an understanding of the thermal effect, only a sudden change of squareness of ~25 μm /m would be seen. This also emphasises the need for temperature measurement as so many factors can be reflected to it.
3. COST OF FAILURE MODEL DEVELOPMENT

The failure of machine tools is the termination of the ability of the machine tools to perform or function within specification. The machine gradually wears over time. Consequently, there will be an unpredicted drop in the quality and quantity of production. In this case, the machine tool needs to be mechanically adjusted or compensated to perform accurately again. The cost of downtime here is unpredicted.

Downtime costs are critical to estimate because they depend on production rate and commodity and even alternative production approaches. On the other hand, estimating it will benefit maintenance decision making in a way of measuring the impact of the equipment failure on system efficiency, and also test any alternatives. Existing studies in the field of predictive maintenance have resulted in cost calculations for the downtime associated with machine failure [11]. It could be said that calibration and predictive maintenance are different applications, but that they can follow the same downtime cost calculation process to decide their applicability for a given asset.

The main objective of this work is to propose a mathematical model for downtime cost taking into account maintenance cost, loss of production and scrap/rework due to parts produced out of tolerance. This model can lead to better calibration decision-making on the relevance, or otherwise, of a PdC strategy and optimising the cycle of calibration process.

Breaking down the factors that contribute to determining the downtime cost is necessary to cover a broad range of machine tool assets, production types and scales. The more factors taken into account, the more global the model would be.

In order to create a clear, simple model based on realistic collection of data, various related costs are investigated.
The technical approach shown in Fig. 3 is not the only approach that contributes in finding the related costs to downtime. The business approach is as important as the technical approach for more accurate cost estimation. This is shown in Fig. 4.
3.1. Model description

A machine tool fails or goes out of tolerance occasionally and requires calibration. From Table 1, this downtime can be in the order of a few days. While the machine is being calibrated, there is a loss in production output. In order to maintain the machine to work within an accepted performing tolerance, which in turn, reduces the number of non-conforming parts, we could periodically perform preventive calibration actions.

The purpose of this model is to optimise intervals at between calibration actions in order to reduce the downtime cost yet maintain control over the accuracy of the machine.

3.2. Model assumptions

Let us consider the following conditions:

- A single unit system condition varies at time, \( t \).
- The unit is subjected to fail or goes out of tolerance, therefore, it requires calibration.
- Preventative calibration and adjustment applied at a fixed rate

The proposed model to obtain the total cost of downtime will ultimately lead to a full costing approach, with the differentiation between direct and indirect costs and considering the contribution of both technical and commercial factors.

3.3. Total cost estimation

The cost of downtime is a challenging task to measure accurately. This is because of the range of factors that should be taken into account [11, 12]. Losses in the areas of labour, revenue and service all contribute to the total cost of downtime [13]. Fig. 3 and Fig. 4 can be combined and shown in a concise and clearer image as in Fig. 5 in order to obtain a model that include the most effective criteria in calculating a cost function.

Fig. 5. Breakdown of factors that contributes in determining the downtime cost

Cost of downtime could include, Time to measure, resetting time, warm up and start up time, production time during steady state production and production time lost due to speed loss. This could be described as time not producing and/or producing scrapped parts. This is shown in Fig. 6.
The cost of measurement of unit within the planning time.

\[ C_m = T_{\text{measure}} \times (C_{hs} + C_{\text{Prod-loss}}) \]  \hspace{1cm} (1)

- **Cost of hire of service.**
- **Cost of lost production during measuring and fixing the machine.**

- The maintenance service hire cost \( C_{hs} \) per unit time is the rate for measurement and repair. This may include the daily rate expenses of labour travel, fuel and accommodation. The way this is calculated differs at different industries. For instance, this cost could not be included in the final downtime cost calculation where a company has its own facilities and does not need to hire this service. In this case it is considered as a fixed cost. On the other hand, other companies need to hire this service, where it is probably being measured as a variable cost. Discussion of the relative merits of each approach is outside the scope of this paper, but is a fundamental management decision that must be made with a large number of other factors taken into account.

\[ C_{hs} = C_{\text{equip.}} + C_l \]  \hspace{1cm} (2)

- **Hire cost of equipment**
- **Labour cost**

- The cost due to resetting and warm up period. It may include the cost of all scrap, rejects and adjustments until the machine settles down and reaches the steady state condition.

\[ C_{\text{warm-up}} = T_{\text{warm-up}} \times C_{\text{Resetting}} \]  \hspace{1cm} (3)

- **Cost of warm-up cycle**
- **Warm-up time**
- **Cost of resetting including all the adjustments.**

- The cost of start-up rejects. This may include all the parts rejected during the start-up period until the machine reaches steady state condition.
$C_{\text{start-up}} = T_{\text{start-up}} \ast (C_{\text{Scrap}} + C_{\text{Rework}}) \quad (4)$

$T_{\text{start-up}}$, is the start-up time.

$C_{\text{start-up}}$, Start-up rejects cost

$C_{\text{Scrap}}$, Cost of scrap part

$C_{\text{Rework}}$, Cost of rework

- Cost of uncontrolled production. It has a direct relationship with the time spent to detect that the machine is producing non-conforming part due to the machine going out of accepted performing tolerance. The cost of a scrapped part is directly affected by variable manufacturing parameters such as energy costs, raw material costs, time to manufacture, etc. For this reason, the equation produced in this study must be considered a “live” tool which must be reanalysed as these costs change.

$C_{\text{Uncontrolled}} = T_{\text{Detection}} \ast (C_{\text{Scrap}} + C_{\text{Rework}}) \quad (5)$

$C_{\text{Uncontrolled}}$, Cost of uncontrolled production.

$T_{\text{Detection}}$, Detection time required at which the machine is producing parts out of tolerance.

- Quality control time consumed is represented as $C_{QCT}$ is very important to consider at all time, where it introduces the cost of checking samples, and as important the time spent identifying faults or what is so called investigation or trouble-shooting time. Cost of material and replacement must be included for an accurate calculation of downtime cost. This usually involves senior/skilled personnel to interpret data to find fault.

$C_{QCT} = C_{T_{\text{Inv}}} + C_{M,R} \quad (6)$

$C_{QCT}$, Quality control cost

$C_{T_{\text{Inv}}}$, Cost of time spent for identifying faults or trouble-shooting time.

$C_{M,R}$, is the cost of material and replacements

- The cost due to producing non-conforming part. This may include losing contracts due to reputational harm because of customer dissatisfaction. It also includes cost of shipping, fines and penalties.

$C_{N/C \text{parts}} = C_{\text{Shipping}} + C_{\text{fines}} + C_{\text{penalties}} \quad (7)$

$C_{N/C \text{parts}}$, Non-conformance parts cost

$C_{\text{Shipping}}$, Shipping cost

$C_{\text{fines}}$, Cost of fines

$C_{\text{penalties}}$, Penalties cost

- Combine equations 1, 3, 4, 5, 6 and 7. The expected cost downtime at all-time can be written as:

$C_{\text{down-time}} = C_m + C_{\text{warm-up}} + C_{\text{start-up}} + C_{\text{Uncontrolled}} + C_{QCT} + C_{N/C \text{parts}} \quad (8)$
4. CONCLUSION

Machine tool accuracy is a key performance index for many high value machining companies. Downtime for calibration is often seen as a non-value-added cost. This paper introduces a mathematical cost function that forms the basis of a strategy for scheduling machine tool calibration which takes into account both the commercial impact and the technical influences on part tolerance.

This model is not a once-only calculation, it will have to be repeated as variations in manufacturing costs such as energy prices, cost of raw materials, etc. influence the model parameters. The next stage of work is to validate the model using case study data for both different scales of production and different part values. This will be achieved by obtaining representative data from project partners. This work will ultimately lead to a technical-driven management tool that can optimise the frequency of calibration to reduce unnecessary downtime while maintaining the machine at the required tolerance.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the UK’s Engineering and Physical Sciences Research Council (EPSRC) funding of the EPSRC Centre for Innovative Manufacturing in Advanced Metrology (Grant Ref: EP/I033424/1).

REFERENCES

A case study of long-haulage transport in China for a sustainable transport and manufacturing business

Danfang Chen, Fredrik Tengdelius

KTH Royal Institute of Technology, Department of Production Engineering, Sweden
Danfang@kth.se

ABSTRACT

The purpose of this article is to describe what the Chinese long-haulage transport industry wants and needs from truck and semi-trailer manufacturers in order to be sustainable in the long run, both in terms of road safety as well as from an environment perspective. Findings include; how companies in Chinese long-haulage transport industry manage their businesses and the problems they face regarding how they can conduct their long-haulage transport business as efficiently as possible. Three key transport segments have been identified and analyzed to identify one segment where an optimized vehicle combination i.e. truck could improve the current segment. This article presents also a proposed design of a transport vehicle that would enable transport companies to haul goods all over China in a legal manner. This would ease logistics for transport companies and reduce the environmental impact.

KEYWORDS: Semi-trailer, long-haulage transport, manufacturing business strategy, Chinese market

1. INTRODUCTION

China is a huge, growing market for many companies in a wide diversity of fields with a rapid growing economy, [1]. This paper focuses on the long-haulage transport industry in China; what Chinese long-haulage transport companies want and need from trucks i.e. tractor and semi-trailer manufacture, see figure 1.

In February 2012, highway transportation stood for 78% (2 232 million tons) of the transport demand in China, which can be compared to railway transport that only stood for 11% (313 million tons).[2]

According to Chong-En Bai and Yingyi Qian [3], only 30-40% of the demand for railway transportation is satisfied with the current railway network, making long-haulage highway transports the only natural alternative. Apart from that, highways in China expand faster than railways which also open up a bigger market for highway transports in the future. Long-haulage transport in China today is mainly conducted with a tractor + semi-trailer combination.
Currently, around one million tractor units are sold in China each year. Of these, 99% are manufactured by Chinese truck manufacturers and 1% by foreign truck manufacturers [4]. To be a part of this 1% or to expand the Chinese market, foreign truck manufacturers need to understand the current Chinese long-haulage situation and optimize their transport solutions accordingly.

Automobile manufacturing research today mainly focuses on how to improve manufacturing and its networks e.g. factory development [5], sustainability e.g. sustainability aspects [6] and remanufacturing [7] [8], digital factory [9], manufacturing management [10], logistic networks [11] [12] and production networks [13]. According to our study little focus is in what the market and customers need and their preferences which trigger manufacturing business, especially in the truck manufacturing domain. This study focuses on how companies in Chinese long-haulage transport industry manage their businesses and the problems they face, regarding how they can conduct their long-haulage transport business as efficiently as possible. To understand the Chinese long-haulage transport market the following research questions are the focus in this study:

- How does the Chinese long-haulage transport industry operate, related to trucks?
- Which is the key segment in the long-haulage transport sector in China?
- Which type of truck would be best suited to fit the key segment?

2. METHOD

The research results are based on a case study performed in China and Europe. Within this study 13 Chinese transport companies and 4 European transport companies have been interviewed and analyzed. Furthermore 5 interviews have been conducted to deepen insights in the industry from a manufacturing and legal perspective where manufacturers have been interviewed in China as well as Europe. The legal perspective was based on interviews with experts in Chinese transport laws to understand how laws are upheld in China and to understand how laws are constructed. The details of the interviewed company’s type, location and position of interviewee are provided in Table 1. The detail information of company and interviewee are not included to ensure the anonymity and confidentiality of the interviewed companies. All the interviewed transport companies are user of high-end trucks and the manufacturing companies are high-end truck manufacturers, to ensure the case study is performed in the right market sector.

To be able to answer the research questions posed earlier, this study maps Chinese transport companies into different segments based on how they use and utilize their trucks for long haulage transport, how they load and unload their semitrailers, what they transport, how goods are packed etc. After that, an understanding of how they use their trucks compared to best practice was created. In this case this entailed benchmarking against well-developed transport companies in Europe to be able to identify similarities and differences. By
benchmarking the segments against European transport companies, improvement possibilities for the Chinese transport market was identified.

Once that analysis was done and segments were in place, a key segment for high-end truck manufacture was identified by evaluating which segment had the biggest potential, market share and most suitable transport needs.

Based on these facts the study then proceeds to give a recommendation of a truck (complete vehicle) design, mainly focused on the semi-trailer that suits the chosen segment and at the same time complies with Chinese laws and standards. Suggested improvements and implications from those changes were then validated with transport companies. The validation was conducted by revisiting 3 of the initial stage interviewed companies as well as one interview with another company in order to get as good results as possible.

<table>
<thead>
<tr>
<th>No.</th>
<th>Industry</th>
<th>Region (Country)</th>
<th>Position of interviewee</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Transport logistics</td>
<td>Huzhou (CN)</td>
<td>CEO</td>
</tr>
<tr>
<td>2</td>
<td>Transport logistics</td>
<td>Guangzhou (CN)</td>
<td>Fleet Manager</td>
</tr>
<tr>
<td>3</td>
<td>Transport logistics</td>
<td>Shenzhen (CN)</td>
<td>Fleet Manager</td>
</tr>
<tr>
<td>4</td>
<td>Transport logistics</td>
<td>Fuzhou (CN)</td>
<td>Vice President</td>
</tr>
<tr>
<td>5</td>
<td>Transport logistics</td>
<td>Shenzhen (CN)</td>
<td>CEO</td>
</tr>
<tr>
<td>6</td>
<td>Transport logistics</td>
<td>Guangzhou (CN)</td>
<td>Fleet Manager</td>
</tr>
<tr>
<td>7</td>
<td>Transport logistics</td>
<td>Huzhou (CN)</td>
<td>Fleet Manager</td>
</tr>
<tr>
<td>8</td>
<td>Transport logistics</td>
<td>Shanghai (CN)</td>
<td>CEO</td>
</tr>
<tr>
<td>9</td>
<td>Transport logistics</td>
<td>Fuzhou (CN)</td>
<td>Vice General Manager</td>
</tr>
<tr>
<td>10</td>
<td>Transport logistics</td>
<td>Guangzhou (CN)</td>
<td>Senior Manager Operations Department</td>
</tr>
<tr>
<td>11</td>
<td>Transport logistics</td>
<td>Shenzhen (CN)</td>
<td>Vice President</td>
</tr>
<tr>
<td>12</td>
<td>Transport logistics</td>
<td>Helsingborg (SWE)</td>
<td>CEO</td>
</tr>
<tr>
<td>13</td>
<td>Transport logistics</td>
<td>Tegelen (NL)</td>
<td>Fleet Manager</td>
</tr>
<tr>
<td>14</td>
<td>Transport logistics</td>
<td>Huissen (NL)</td>
<td>CEO</td>
</tr>
<tr>
<td>15</td>
<td>Transport logistics</td>
<td>Södertälje (SWE)</td>
<td>CEO</td>
</tr>
<tr>
<td>16</td>
<td>Transport logistics</td>
<td>Shanghai (CN)</td>
<td>Senior Manager Operations Department</td>
</tr>
<tr>
<td>17</td>
<td>Transport logistics</td>
<td>Beijing (CN)</td>
<td>Manager</td>
</tr>
<tr>
<td>18</td>
<td>Manufacturing</td>
<td>Ätran (SWE)</td>
<td>Sales Director</td>
</tr>
<tr>
<td>19</td>
<td>Manufacturing</td>
<td>Delft (NL)</td>
<td>MD &amp; TD</td>
</tr>
<tr>
<td>20</td>
<td>Manufacturing</td>
<td>Longkou (CN)</td>
<td>VP, GM, Director &amp; Deputy GM</td>
</tr>
<tr>
<td>21</td>
<td>Manufacturing</td>
<td>Laxå (SWE)</td>
<td>Head of Sales and Development</td>
</tr>
<tr>
<td>22</td>
<td>Legal</td>
<td>Beijing (CN)</td>
<td>Working group manager</td>
</tr>
</tbody>
</table>

Some limitations have been made due to time, resources and specific focus of the product:

- The study only concerns tractors with semi-trailers intended for long-haulage transport, i.e. 1000 km / day or above.
- In this study the truck only refers to a tractor with semi-trailer, other vehicles such as construction trucks within the truck category are not taken into account.
- The study focuses only on transport companies with high-end trucks, i.e. western trucks.
- The study does not concern detailed manufacturing costs because the studied transport companies already use quality trucks with a high purchase price, and detailed manufacturing cost is an issue later on the strategy process.
- The study does not elaborate on detailed technical aspects of the trucks.
3. CHINESE LAW GB1589-2004

GB1589-2004 is the Chinese law document containing laws for both trucks as well as personal vehicles. Below is a summary of the most important information for this study regarding trucks, see also Figure 2:

- Total length – 18 100 mm.
- Total width – 2 550 mm. This does not include the rear view mirrors which are allowed to protrude 200 mm if the lower edge of the rear view mirror is less than 1 800 mm above ground and 250 mm, if it is above 1 800 mm.
- Total height – 4 000 mm.
- Semi-trailer length – 14 600 mm.
- Semi-trailer width – 2 550 mm.
- Semi-trailer height – 4 000 mm.
- Axle load tractor – driven axle – 11 500 kg provided that the tractor has three axles or two axles, air-suspension and dual tires on each side. Other combinations may reduce the permitted axle load.
- Weight on the driven axle cannot be less than 25 % of the gross vehicle weight (GVW). This rule indirectly limits the total gross weight to 46 000 kg.
- Axle load tractor – non driven axle – 10 000 kg. If the tractor is equipped with air suspension, the limit will be 11 500 kg. Other combinations may reduce the permitted axle load.
- Axle load semi-trailer – 10 000 kg for a three axle semi-trailer. If the semi-trailer is equipped with air suspension the limit is 11 500 kg. Other combinations may reduce the permitted axle load.
- Tandem load tractor – 11 500 kg if the distance between the axles is less than or equal to 1000 mm, 16 000 kg if the distance is between 1 000 – 1 300 mm and 19 000 kg if the distance is above 1 300 mm.
- Tandem load semi-trailer – 21 000 kg if the distance between the axles is less than or equal to 1 300 mm and 24 000 kg if the distance is between 1 300 – 1 400 mm.

Figure 2. GB1589-2004 measures of a truck.
4. SITUATION IN CHINA TODAY

This section describes how the Chinese long-haulage transport industry operates today based on the conducted interviews as well as observations. Analyzes are performed based on data gathered during the interviews. The section is divided into two parts. The first part focuses on the market itself and the second part focuses on the equipment used by transport companies.

4.1. The situation – Chinese long-haulage transport market

Today a majority of long-haulage transports conducted in China are characterized by the large amount of manual labour used compared to European transport companies. This is an effect of diverse goods, several different standards of pallets, poor IT utilization and tough competition in the industry, which demands a high volume utilization to maximize profits. Whereas European long-haulage transports consist of goods packed on pallets, the Chinese counterpart uses the cheaper labour to pack semi-trailers in a piece-by-piece manner manually, thereby enabling transport companies to utilize the available load volume more efficiently.

As a consequence of this loading method, loading and unloading a semi-trailer takes in average 18 work hours, which can be compared to 2 work hours with pallets and forklifts. This increases lead time for goods and affects how transport services are priced. Companies charge their customers either based on weight (average of CNY 0.3124/ton/km) or on volume (average of CNY 0.0869/m³/km) and have only recently started to include a time-guarantee for prioritized deliveries.

Another significant characteristic of Chinese long-haulage transport is the vast distances covered on routes. Due to the size of the country and insufficient railroad capacity, goods are often transported with trucks over distances up to 4 000 km. These are often covered non-stop with two or three drivers that alternates between driving and resting with stops only for fuel and/or a driver shift. An imported truck covers between 300 000 – 400 000 km / year, whilst a locally produced counterpart struggles to obtain the same coverage due to quality problems. Such long distances as 300 000 km or above are extremely high even compared to most European companies.

Benchmarking against European markets also show another major difference between the two markets in terms of equipment used. European semi-trailer manufacturers have focused over a long period of time on weight reduction in order to make transport companies able to transport as much goods as possible in terms of weight. In China however, focus has been on security to protect transported goods against damage in traffic accidents and from theft. Therefore, the semi-trailers used in China today are robust and heavy compared to the European counterpart. The most common method for long-haulage transport in China today is custom made sea containers, adapted to fit on a container semi-trailer. A comparison between a standard semi-trailer in Europe and China shows that a Chinese semi-trailer in average weighs 118 kg/m more than the European counterpart.

European trucks have also been adapted to comply with legal demands set in Europe where a truck transporting goods through several countries has to obey the laws of each country it passes from pick up to drop off. Penalty fees in Europe are high and consistent,
which has led transport companies to strictly follow these laws. In China however, laws are more ambiguous and are interpreted differently in different provinces and in dependence on the transport company. The central government issues laws with guidelines for penalty fees and each province individually interpret and uphold them. As a consequence of this, transport companies compete with different conditions. A transport company in one province can get away from penalty fees in one situation where a company from another province will not. To further complicate the situation, 53 feet semi-trailers are today allowed to be used but not produced during a phase out time where 53 feet semi-trailers are to be replaced with a maximum length of 48 feet. As a result, all studied transport companies in China exceeds at least one law from GB1589 2004, most commonly length of the truck or length of the semi-trailer.

One last major difference between Europe and China discussed in this article is the specifications of the trucks. Many Chinese transport companies invest in new equipment on an economic short term basis, which makes the purchase price of new trucks an important factor. As opposed to that, European transport companies tend to view their investments from a long term perspective, which makes other factors more important than the purchase price alone. Chinese transport companies, as purchase price is a key factor, specify their trucks with lower standards than their European counterparts to get a better price.

4.2. The situation – Truck design

The most common setup is a 6x2 tractor (6x4 for Chinese markets) with leaf spring suspension, drum brakes and a small engine paired with a 3-axle container semi-trailer. Further on, tractors are often equipped with a bare minimum of security and comfort specifications for drivers, which decreases the purchase price even more. These trucks are then used for as long as possible with little regard to depreciation time or second hand value.

In 2010, 315 955 semi-trailers were registered in China. 91 % of those had 3 axles, 97 % had double tires and 95 % were designed for a GVW of 34 000 kg or more [4]. All interviewed transport companies used mainly container semi-trailers even though the most common type in the industry as a whole is a stake body. This is likely a consequence of the sample of interviewed companies as they belong to a high-end segment of the industry. Semi-trailers are used for fifteen years which is the legal limit. This puts high demands on the durability of semi-trailers in China.

Locally produced tractors are most common and used for all types of goods. Imported tractors are however more rare and are usually used for prioritized goods and on prioritized routes. Uptime is calculated for both types of tractors. Due to the higher purchase price of imported tractors the uptime of those is prioritized higher. A high uptime is also ensured by having more semi-trailers than tractors so that when a tractor unloads a semi-trailer at its destination, it can couple with another fully loaded semi-trailer going to another destination without having to wait for the loading/unloading procedure.

By using custom made semi-trailers and not having a set semi-trailer to a tractor, the position of the kingpin and turntable are placed in a way that ensures no collision between cab and semi-trailer while driving regardless of which tractor is combined with which semi-trailer. This however has created an issue with a huge gap between the two that increases drag and thereby also fuel consumption.

5. KEY SEGMENT IN THE CHINESE LONG-HAULAGE TRANSPORT SECTOR
Interviewed and analyzed companies where segmented based on their filling rate and gross vehicle weight (GVW) into three different segments:

- **Weight** – Companies in this segment reach the maximum allowed GVW before the truck is fully loaded in terms of volume.
- **Volume** – Companies in this segment fill the trucks volume before maximum GVW is reached.
- **Mixed** – Companies in this segment reach both a maximum filling rate and a GVW maximum at the same time.

The basis for the segmentation was as follows:

- **Weight** as main bottleneck if payload ≥ 52 500kg and filling rate < 90%.
- **Volume** as main bottleneck if payload < 52 500kg and filling rate ≥ 90%.
- A **Mixed bottleneck** if payload ≥ 52 500kg and filling rate ≥ 90%.

The resulting segment split, which can be seen in Table 2, was cross-referenced with perceived bottleneck by the company itself in order to increase segmentation validity.

Table 2. Summary of properties for Chinese transport companies.

<table>
<thead>
<tr>
<th>No.</th>
<th>Perceived bottleneck (stated by company)</th>
<th>Filling rate (stated by company)</th>
<th>GVW (kg)</th>
<th>Bottleneck</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Weight</td>
<td>80 %</td>
<td>55 000</td>
<td>Weight</td>
</tr>
<tr>
<td>2</td>
<td>Volume</td>
<td>95 %</td>
<td>N/A</td>
<td>Volume</td>
</tr>
<tr>
<td>3</td>
<td>Volume</td>
<td>100 %</td>
<td>36 826</td>
<td>Volume</td>
</tr>
<tr>
<td>4</td>
<td>Volume</td>
<td>95 %</td>
<td>50 426</td>
<td>Volume</td>
</tr>
<tr>
<td>5</td>
<td>Volume</td>
<td>100 %</td>
<td>52 500</td>
<td>Mixed</td>
</tr>
<tr>
<td>6</td>
<td>Weight &amp; Volume</td>
<td>95 %</td>
<td>55 974</td>
<td>Mixed</td>
</tr>
<tr>
<td>7</td>
<td>Weight &amp; Volume</td>
<td>100 %</td>
<td>53 391</td>
<td>Mixed</td>
</tr>
<tr>
<td>8</td>
<td>Weight &amp; Volume</td>
<td>90 %</td>
<td>54 000</td>
<td>Mixed</td>
</tr>
<tr>
<td>9</td>
<td>Weight &amp; Volume</td>
<td>100 %</td>
<td>54 200</td>
<td>Mixed</td>
</tr>
<tr>
<td>10</td>
<td>Weight &amp; Volume</td>
<td>90 %</td>
<td>51 574</td>
<td>Volume</td>
</tr>
</tbody>
</table>

The segment sizes based on the analyzed companies indicates that a mix between weight and volume is the biggest segment in the Chinese transport industry (50%), volume as main bottleneck is the second biggest (40%) and weight is the smallest (10%).

A general business case analysis for each segment was then conducted. The analysis was based on current legislations and assuming a European standard of the trucks. By comparing each segment an indication of the segment with the biggest improvement possibilities was identified. To conduct the analysis, the following figures where used, which in turn was based on interviews.

- Average distance covered / year / vehicle: 365 000 km / year
- Income / ton / km: CNY 0.3124 / ton / km
- Income / m³ / km: CNY 0.0869 / m³ / km
- Average weight of current semi-trailer / m: 743 kg / m
- Average volume of current semi-trailer: 120 m³
- Average filling rate: 95 %
- Weight of European box semi-trailer / m: 625 kg / m
- Weight of European curtainsider semi-trailer / m: 478 kg / m
• Volume of legal semi-trailer: 103 m³

By combining these figures it is possible to calculate differences in income between the different scenarios. For this analysis and results presented below, only legal semi-trailers in terms of measures were considered as to be able to compare the results evenly.

It is in this analysis concluded that volume sensitive companies will not be able to benefit from a better semi-trailer in any other aspect than a slight reduction in fuel consumption by a smaller curb weight and a better configuration. The reason for that is that volume sensitive customers today use the bigger, 53 feet, semi-trailers which are not allowed to be produced today.

Further on, the mixed segment will be able to increase the payload slightly, but benefits are limited by the already high volume utilization with the same reasoning as for volume sensitive companies.

The weight sensitive segment will benefit the most as companies in this segment will be able to load more goods without reaching the GVW limit and thereby benefit the most from improved standards. A more detailed result from the analysis can be seen in Fel! Hittar inte referenskälla.

Table 3. Summary of segments.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Weight segment</th>
<th>Volume segment</th>
<th>Mixed segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result of business case</td>
<td>CNY 410 950 / year</td>
<td>CNY 0 / year</td>
<td>CNY 98 233 / year</td>
</tr>
<tr>
<td>Percentage of market</td>
<td>10 %</td>
<td>40 %</td>
<td>50 %</td>
</tr>
</tbody>
</table>

Market size suggests that the mixed segment is the biggest while the business case suggests that it can benefit from adapting European standards. This is further on supported by statements during interviews and the fact that the average filling rate is 94,5% and the average GVW is 51 543 kg, which is above the legislated maximum limit of 49 000 kg.

6. PROPOSED TRUCK COMBINATION TO BEST FIT THE CHINESE TRANSPORT INDUSTRY

A truck adapted towards the mixed segment must be as light as possible while at the same time also maximize load volume in order for it to appeal to companies with both weight and volume as bottlenecks. Proposed specifications for a truck that would suit those needs are presented in Table 4 and Figure 3.

Table 4. Current and proposed specifications.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Current</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body type</td>
<td>Custom made sea container</td>
<td>Box body of a composite material</td>
</tr>
<tr>
<td>Suspension</td>
<td>Leaf spring suspension</td>
<td>Air-suspension</td>
</tr>
<tr>
<td>Brakes</td>
<td>Drum brakes</td>
<td>Disc brakes</td>
</tr>
<tr>
<td>Axles</td>
<td>6 (3 axle tractor, 3 axle semi-trailer)</td>
<td>6 (3 axle tractor, 3 axle semi-trailer)</td>
</tr>
<tr>
<td>Kingpin</td>
<td>Kingpin #50, 2 040 mm front overhang</td>
<td>Kingpin #50, 2 040 mm front overhang</td>
</tr>
<tr>
<td>Tires</td>
<td>Dual tires</td>
<td>Super single tires</td>
</tr>
<tr>
<td>Support legs</td>
<td>Flat</td>
<td>Flat</td>
</tr>
<tr>
<td>Overall design (l<em>w</em>h)</td>
<td>(16 154mm<em>2600mm</em>4200mm)</td>
<td>(14 600mm<em>2 550mm</em>4 000mm)</td>
</tr>
</tbody>
</table>
Weight of semi-trailer | 12 500 kg | 9 500 kg

Suggested changes mainly aim at making sure that both GVW is as low as possible while at the same time maximizing load volume towards the legal limits. Another important change in specification is that tractor and semi-trailer are equipped with the same, modern, suspension and brakes in order for the complete truck to act safely and consistently on the road. This is essential as many trucks are overloaded and companies often faces problem with overheated brakes.

It is apart from these specifications also suggested that, in the long run, specifications for the safety and comfort of drivers should be improved in order for companies to be able to keep their employees rested and safe while driving. This will be essential as many companies have recognized a tough environment amongst companies where good drivers are hard to find and salaries increase. One way to diversify a transport company in China could be to offer employees a comfortable and safe work environment.

The proposed changes will undoubtedly increase the purchase price of truck (vehicle combination), but will also enable transport companies to transport more goods due to a lighter setup. They will thanks to a better and uniform solution, also be able to decrease fuel and repair costs. For instance fuel consumption will decrease as a switch from dual tires to super single tires will result in less friction against the tarmac. Repair costs will be reduced as a uniform solution works better together and an uneven wear of components could be reduced. One last advantage of the proposed setup is that trucks will comply with GB1589-2004 and thereby trucks will eliminate all costs associated with penalty fees and bribes which is a problem in China. [15]

7. CONCLUSION AND DISCUSSION

According to this case study, the trucks used for long haulage transport in China do not match the current Chinese law, GB1589-2004. There is a gap between current trucks and the GB1589-2004 law. This gives truck manufacturers a possibility to provide an optimized truck solution with regards to both the GB1589-2004 law and the needs of transport companies.
It is in this case study discussed that the most important segment is a segment that faces both weight and volume as bottlenecks. Therefore an optimal solution is a truck that is as big and light as possible. As current semi-trailers are bigger than the allowed limit, the weight is the area in which biggest improvements are possible.

Further on it is also discussed in this article that optimizing weight and volume alone is not enough. Improvements could also be made for fuel economy and safety of the trucks, if aspects such as brakes, suspension and tires are considered and matched between the tractor unit and the semi-trailer.

An improvement of the equipment used by Chinese transport companies will increase the purchase price but will in the long run pay off as penalty fees, bribes and fuel costs will decrease while at the same time more goods can be transported for which companies will make more money.

REFERENCES


**Theme 8**

**Inspection and Quality Assurance**

- Barkhausen Noise Analysis of Surface Integrity in Grinding of Large Bearing Rings
  **Miroslav Neslušan, Anna Mičietová, Mária Čilliková, Zuzana Durstová**
  Faculty of Mechanical Engineering, University of Žilina, Slovak

- New Technologies for Individual Joint Implants
  **Miroslav Piška, Josef Sedláček, Eva Pekárková**
  Faculty of Mechanical Engineering, Brno University of Technology, Czech Republic

- A Roadmap to Model Based Manufacturing
  **Dr. Martin Hardwick**
  Department of Computer Science, Rensselaer Polytechnic Institute, USA

- A Study of the Surface Integrity after Machining by means of Non-Destructive Testing Methods
  **P.L. Tam¹, S.B. Hosseini², J. Holmberg³, P. Lundin⁴, L. Mattsson⁵, L. Nyborg¹**
  ¹ Chalmers University of Technology, Gothenburg, Sweden.
  ² AB SKF, Gothenburg, Sweden.
  ³ Swerea IVF AB, Mölndal, Sweden.
  ⁴ Stresstech Oy, Rönninge, Sweden.
  ⁵ KTH Royal Institute of Technology, Department of Production Engineering, Sweden.

- Impact Acoustic Testing as NDT Method for Classification of Compacted Graphite Iron
  **Lorenzo Daghini¹, Anders Berglund²**
  ¹ KTH Royal Institute of Technology, Sweden
  ² Scania CV AB, Sweden
Barkhausen Noise Analysis of Surface Integrity in Grinding of Large Bearing Rings

Miroslav Neslušan, Anna Mičietová, Mária Čilliková, Zuzana Durstová

Faculty of Mechanical Engineering, University of Žilina, Slovakia
miroslav.neslus@gmail.com

ABSTRACT
This paper deals with detection of surface burn after grinding operations on larger bearing rings made of case-hardened steels. The paper reports about implementation of Barkhausen noise technique for non-destructive monitoring of grinding operations for rings of diameter in the range of 600 to 4000 mm. Further investigations found out that grinding burn can be reliably monitored. The paper also verifies theoretical knowledge about influence of surface integrity expressed in such terms as microstructure, hardness as well as stress state on magnetoelastic responses of surfaces. Therefore, analyses of surface hardness and microhardness measurement as well as corresponding X-ray diffraction are reported in the paper to verify the information about surface obtained via micromagnetic technique.

KEYWORDS: grinding burn, Barkhausen noise, bearings

1. INTRODUCTION

Magnetization in a ferromagnetic material is caused by the nucleation and reconfiguration of domains which result from the motion of magnetic domains and corresponding Bloch Walls (BW) with increasing external magnetic field. The motion of BW and domains is usually pinned by precipitates and other lattice imperfections and result in their discontinuous movement. Then pulsating magnetization can be obtained and corresponding discontinuous jumps of the BW occur due to the rapid magnetic flux. This phenomenon is called as Barkhausen noise [1]. Micromagnetic investigation of machined surface based on Barkhausen noise (BN) has found high industrial relevance. These techniques are mostly adopted for inspection of ground surface since strong correlation between the heat generated in the grinding wheel – workpiece contact (its dissipation by workpiece) associated surface burn and corresponding magnetoelastic responses expressed in BN values. The capability of BN technique, considering the reliable and highly sensitive detection of grinding burn, results from alteration of microstructural features as well as residual stress state (compressive stresses beneath the surface are shifted towards tensile stresses). It is well known that surface of high hardness give the poor BN values, whereas softer structures emit higher BN values [2, 3]. Moreover, magnetoelastic responses are suppressed when compressive stresses occur while more pronounced BN values can be
detected for surface containing mainly tensile stresses due to increasing domains parallel with the direction of magnetic field (while the domains perpendicular to the load are decreasing) [2 - 4].

The main advantages of Barkhausen noise (BN) method are associated with very fast surface response (in seconds), portability of BN systems and ability to be easily integrated into automatic cycles and robotic cells. BN techniques are mainly applied for monitoring surface integrity of parts loaded near their physical limits. Then surface integrity expressed in terms of residual stresses, hardness alterations or structure transformation correlates with BN values obtained from the surface as well as associated functionality of produced parts. Bearings are usually critical components in machine structures. The role of BN technique is usually connected with variable surface integrity despite the constant cutting and other conditions are kept. Acceptable cutting conditions, grinding wheel, its dressing, coolant supply and other conditions are usually properly suggested to obtain the acceptable precision of produced parts and required surface integrity. On the other hand, it should be understood that grinding process itself is affected by many parameters (about 150), these parameters can randomly vary with more or less pronounced influence on grinding process. Therefore, quality of production can vary and surface integrity of certain parts can be indicated as unacceptable. Being so, industrial relevance for micromagnetic evaluation of surface integrity is connected with detection of thermally damaged surfaces after grinding [5 - 7] as a phenomenon randomly occurring in production. The information about surface integrity then can be used for detection and understanding of critical aspects of grinding as well as possible optimizing concepts.

A specific segment of bearing industry is production of large bearings (of diameter from 600 to 4000 mm) for specific applications such as wind power industry. Producers should guarantee at least 20 years failure - free bearing operation. Otherwise extra costs (up to 300 000 Eur) should be expected for the ring replacement. Therefore, the investigation of surface integrity in production of large rings takes a significant role and micromagnetic evaluation is considered as a suitable technique to obtain fast and reliable information about surface state after grinding.

This paper discusses experience with implementation of BN technique in the production of large bearing for detection of surface integrity after finishing grinding. The detailed concept can not be fully reported due to agreement to keep a certain information secret. However, the specific aspects can be discussed.

2. RESULTS OF EXPERIMENTS

The experimental study was carried out on case hardened steel of hardness varying between 60 ÷ 62 HRC and variety of diameters. Micromagnetic testing was performed by use of Rollscan 300 device and software package Microscan in the frequency range of 70 to 200 kHz (mag. frequency 125 Hz, mag. voltage 10 V). Residual stresses were measured via X-ray diffraction technique (α-Fe, CrKa, X’Pert PRO). Vickers microhardness measurement was conducted by Hanneman micro-hardness tester by applying force 300 N for 10 seconds. Surface hardness was carried out as a dynamic test by applying Equotip 2 device.

Grinding wheel and conditions: 350x50x127, A98 80 J9V, \( v_c = 25 \text{ m.s}^{-1}, v_f = 25 \text{ m.min}^{-1} \), raceways width = 150 mm, \( Q_w = 5 \text{ mm}^3 . \text{mm}^{-1} . \text{s}^{-1} \), \( a_e = 4 \mu m, a_{e \text{ tot.}} = 200 \mu m \), external cylindrical plunge grinding, dressing conditions: single crystal diamond dresser, \( a_{ed} = 20 \mu m, v_{ed} = 25 \text{ m.s}^{-1}, v_{fd} = 90 \text{ mm.min}^{-1} \), shoe nozzle: Ecocool 3%, \( Q_{cl} = 25 l.\text{min}^{-1} \).

To attain industrial relevance of BN technique, specific aspects of production (in the scope of large ring production) have to be investigated such as:
- verification of BN technique for detection of grinding burn,
- correlation between thermal softening of the ground surface and corresponding magnetoelastic responses,
- investigation of correlation between grinding burn, associated surface integrity (expressed in such terms as structure transformations, microhardness alteration and stress state) and corresponding magnetoelastic responses expressed in RMS values of BN,
- identification of critical aspects affecting grinding burn (see final comments).

Fig. 1. Section of inner ring with indicated grinding burn
Fig. 2. Part of scanned inner ring
Fig. 3. Raw BN signals obtained from the untouched and thermally softened zone

The correlation between the surface integrity state and corresponding BN values can be investigated when surface of the variable quality and corresponding BN values are obtained. Figure 1 illustrates the section of the ring made of case-hardened steel after grinding where variable BN values are obtained through the raceway width. Figure 3 illustrates the raw BN signal obtained from the zone where significant thermal burn occurs as well as the BN signal from the untouched zone. (Figure 3 is a compilation of two BN signals. The left side indicates the example of the high BN signal and the y-axis shows the received BN voltage, the right side illustrates the low BN signal and the y-axis shows the course of applied magnetization current). Figure 4 illustrates the frequency spectrum of both signals illustrated in Fig. 3. Figure 4 shows that except the frequency range, from which the BN values are calculated, higher amplitudes of BN magnetoelastic responses can be also found at higher frequencies and inspected surfaces emit the strong low frequency as well as valuable high frequency pulses. The shape of frequency spectrum, as shown in Fig. 4, depicts that the low frequency also the high frequency sources considerably contribute to the received BN voltage. Furthermore, Fig.
4 also shows that despite the different BN voltage the shape of frequency spectrums is not remarkably altered what is indicated by the absence of structure transformation usually associated with re-hardening grinding burn.

Fig.4. Frequency spectrum of BN signals

Fig.5. BN values measured in the different sections on the ring raceway indicated in Fig. 2

Fig.6. BN values and hardness measured across the ring width where progressive increase of BN values occur

Fig.7. BN values and hardness measured across the ring width where high BN values through the whole width occur

Except for the investigation of a small section of the ring, more detailed scanning illustrated in Fig. 2 was carried out. The measurement of BN values illustrated in Fig. 5 indicates that except sections where the progressive BN increase across the raceway width
occurs (see Fig. 6) certain zones exhibit high BN values through the whole measured width (see Fig. 7). The measurement of conventional surface hardness indicates that all surfaces emitting high BN values give lower hardness as a result of thermal softening due to high heat flux during grinding. Figure 8 shows that progressive increase of BN values correlates with a certain decrease of surface hardness. The minimum surface hardness can be found near the left side of the ring raceway where the highest BN values are found. On the other hand, more pronounced thermal softening can be found when the high BN values occur through the whole raceway width. The minimum hardness is located near the raceway centre. This investigation indicates that a strong correlation between BN values and surface hardness can be found when limited thermal softening is induced by grinding process. As soon as the thermal softening exceeds a certain degree, BN values saturate and the correlation between BN values and surface hardness is missing (see Fig. 7 and Fig. 8). Saturation indicated in Fig. 7 and Fig. 9 is due to limited skin depth as a result of the high frequency BN technique. As reported Moorthy and Shaw [8, 9], the skin depth of BN signal should be considered as the depth of material from where the magnetization process contributes to the detected BN signal on the surface. BN signal is a subject of electromagnetic attenuation. Hence, BN signal generated deep below the surface can not be fully detected on the surface. Therefore, despite extending thermal softening beneath surface BN values saturate in correspondence with a certain skinned depth.

Surface hardness does not represent the proper and usual concept for evaluation of thermal softening extending beneath a surface. Therefore, the conventional microhardness technique (where hardness of structure in different layers beneath the surface is investigated) was carried out to examine microhardness profile in the different positions through the raceways width. Figure 9 illustrates that progressive increase of BN values through the raceway width and associated decreasing surface hardness corresponds with indicated microhardness profiles. A thermally almost untouched zone (Fig. 9, position 15 mm) exhibits near surface microhardness about 800 HVm and limited thickness where microhardness decrease extends. Increasing thermal softening expressed in higher BN values and surface hardness correspond with more pronounced microhardness decrease near the surface as well as extending zone of softened structure.

Fig. 8. Surface hardness measured in two different sections
The microhardness profiles investigated on the section where the saturated surface hardness is found proves the information obtained by BN technique and surface hardness measurements. The centre of the raceway width, where the highest surface hardness descent is found, exhibits the most pronounced microhardness fall (Fig. 10). All remaining measured profiles through the raceway width indicate softened structure as well. However, microhardness decrease as well as thickness of softened structure is less pronounced. Thermal softening expressed in term of microhardness profiles corresponds with FWHM parameter of diffraction peak obtained via X-ray diffraction technique as Fig. 11 illustrates. The low FWHM values corresponding with the high BN values are due to thermally induced relaxation.

It is well known that BN values depend on BW motion (average of BW motion paths) as well as BW arrangement. BW interferes with microstructural features such as dislocation, precipitates, grain boundaries, other phases and lattice imperfections as well as magnitude and nature of residual stresses state [8 - 10]. Moorthy [8] reported that microstructural features affect the pinning strength and the mean free path of the BW displacement while stresses affect mainly domain alignment with respect to the stress direction. The explanation of higher BN values with progressive thermal load during grinding is connected with superimposing of more pronounced thermal softening and a residual stress state. When thermal load of ground surface is limited, the microstructure dominates by the carbide precipitates.
and the paramagnetic retained austenite phase, which is considered as strong obstacles to the BW motion. BW also interferes with the high dislocation density. As soon as more pronounced thermal softening takes place, BW motion is enhanced due to reduced dislocation density, coarsening of carbide precipitates and transformation of paramagnetic austenite to martensite. Being so, thermal softening induced by grinding strongly correlates with BN values also with more extended HAZ thickness shown in Fig. 12. Furthermore, residual stresses shifted towards the tensile zone also contribute to the higher level of generated BN signal level. Figure 11 indicates that BN values strongly correlate with conventional measurement of residual stresses. Since the thickness of the layers associated with the different measuring technique is different (an X-ray diffraction technique is associated with the surface thickness of several micrometer; thickness of layer associated with BN technique is several times higher) strong correlation between BN values and corresponding stresses can be found because the near surface contributions to the received BN signal most of all while deeper zones (not inspected via an X-ray diffraction technique) are progressively attenuated.

The variable surface integrity expressed in term of BN values can be used for investigation of correlation of BN values and surface integrity expressed in term of heat affected zone (mainly thickness of heat affected zone). The structure transformation beneath the surface was observed via a conventional metallographic observation when surfaces of the different magnetoelastic responses were prepared for analysis. More than 40 pieces were cut from both raceway sections (the section where progressive BN values increase occurs as well as the section where high BN values are obtained through the whole raceway width). Except for raceway surface, additional 3 ground surfaces were cut from the same ring surface where very low BN values were measured (from the ring face and inner surface). The examples of structures and corresponding BN values are illustrated in Fig. 12.

The obtained micrographs show that re-hardening burn does not occur and the structure alterations are connected with pure thermal softening. Figure 13 also shows that no modification of grain size usually attributed to the re-hardening takes place. Although, burst curves obtained for thermally softened surface exhibit higher amplitude, the position of the burst amplitude (peak position) stays nearly the same. It was previously reported that the peak position (indicated as a position of the burst curve maximum) corresponds to the average grain size and the profile corresponds to the grain size distribution [8, 11]. Re – hardened surfaces (indicated as white layers) induced by abusive grinding usually emit magnetoelastic pulses of low magnitude and high frequency. Therefore, peak position is shifted towards higher values due to frequency spectrum alteration [8]. Neither micrographs, burst profiles, nor microhardness profiles evoke surface re-hardening.

BN value can be indicated as a result of corresponding structure beneath the surface. Figures 12 and 14 illustrate that BN values and structure transformations expressed in the thickness of heat affected the (thermally softened) zone. Non affected zone corresponds with very low BN values (about 22 mp, see Fig. 12a), a very thin layer corresponds with limited thermal softening of surface (about 45 mp, see Fig. 12b). On the other hand, progressive BN increase is due to corresponding increase in thickness of heat affected zone as Fig. 12c, d indicates. Figure 14 shows the correlation between BN values measured on the surfaces and the corresponding thickness of heat affected zone (HAZ). This figure indicates that progressive increase on BN values strongly correlate with increasing thickness of heat affected zone. Further, to meet requirements considering the functionality of produced parts, this correlation also enables us to suggest the suitable concept for monitoring of grinding burn via BN technique since critical BN value can be established.
Fig. 11. Residual stresses in two sections and corresponding FWHM feature

a) BN = 22 mp – face of the ring
b) BN = 45 mp - raceway
c) BN = 90 mp - raceway
d) BN = 127 mp - raceway

Fig. 12. Microstructures of ground surfaces
The stocks needed for removal of HAZ zone via additional grinding can be exactly determined since BN value corresponds with a certain thickness of a heat affected zone. Thereafter, the ring raceways can be reground to reach the required surface structure (a thermally affected layer is removed via additional grinding operation) and bearing rings can be repaired. When a raceway diameter of an inner ring is reduced because of additional grinding operation corresponding raceways diameter of an outer ring should be adapted.

3. CONCLUSIONS

The experimental measurements lead to conclusion that:
- grinding burn can be successfully monitored through the BN technique,
- high BN values are associated with thermal softening of the ground surface while low BN values correspond with untouched structure or limited and isolated areas where grinding burn can be found,
- increasing BN values correspond with increasing thickness of the HAZ,
- different character of structure transformations can be found beneath the surface.

FINAL COMMNETS

The additional experiments were conducted to analyze the specific aspects of grinding such as influence of grinding wheel wear, short time absence of coolant of surface integrity and corresponding BN. In the case of coolant supply thermal softening is clearly located and easily recognized (despite its variable thickness, see Fig. 12c). On the other hand, thermally
softened layers induced by dry grinding have a thick zone where thermally softened and untouched structures are mixed. While a thermally softened structure is dominant on the near surface, its occurrence becomes less significant in the subsurface layers. Dry grinding does not avoid free penetration of heat into the surface. However, the coolant absorbs most of the heat generated in the wheel – workpiece interface; it causes rapid cooling of ground surface and therefore hinders deep penetration of the heat and corresponding temperatures into the subsurface layers.

It was clearly investigated that the section of raceways where a progressive increase of BN values occur is associated with wet grinding and thermal softening mainly due to insufficient coolant supply. It was found that due to unexpected decrease in coolant flow cooling effect is reduced. Higher values obtained outside of the ring centre are due to unexpected decrease in coolant flow. The position of the ring during grinding is the same as it is indicated in Fig. 2. Then, in correspondence with gravitation force, during the unexpected coolant flow decrease bottom side of the raceway (near the ring centre) is well overflowed. The reduced coolant supply in the upper side of the raceway causes grinding burn associated with the high BN values. The thermally softened layer can be clearly distinguished form the raw structure with the typical whitening character. On the other hand, the structure in the sections where the high BN values occur across the whole raceway width exhibit higher thickness of HAZ with the typical dark structure as well as the thick zone where the mixture of thermally softened and untouched structure occurs as a characteristic sign of dry grinding.

ACKNOWLEDGEMENT

This work was supported by the VEGA project n.1/0223/11.

REFERENCES

New Technologies for Individual Joint Implants

Miroslav Píška, Josef Sedlák, Eva Pekářková

Faculty of Mechanical Engineering, Brno University of Technology, Czech Republic
piska@fme.vutbr.cz, sedlak@fme.vutbr.cz, pekarkova@fme.vutbr.cz

ABSTRACT

The paper deals with a technology of processing CT data of the knee joint and production of implants that substitutes only the defected bone part, minimizes the removal origin bone and guarantees a better fitting and function of the implant, especially the knee one.

The new high-technology is focused on distal part of femur and production of the most important part - the femoral component. The technology is split in three main steps – a production of the optimal implant and an insertion of the implant onto the origin bone.

Firstly, the modern methods of 3D scanning and data acquisition (CT, MRI, X-ray) and their processing by CAx applications are widely studied in the paper. Secondly, some methods of data postprocessing for CAM or Rapid Prototyping Technology - FDM and DMLS method are explained. This operation technique has been researched and developed for young patients basically. This innovation can provide significant benefits for patients a reduction of surgical trauma, reduction of time surgery and better social adaptation of patients.

KEYWORDS: Joint, Implant, Technology, CAD software, CT data

1. INTRODUCTION

The human development and standard of living include many factors to assess, but one of the most important is the level of medical care. This field covers dozens criteria again, but one of the most significant is the orthopaedics care and the sort of implants available for patients. This progress is mostly based on engineering methods and advanced technology.

The present total knee joint replacement is a set of three main components used to recover natural function of the knee joint. In particular, the femoral part, the insert and the tibial part. In detail, standard used femoral components are today mainly made of biocompatible Co-Cr-Mo or Ti-alloys by methods of casting and cutting operation. Femoral components are made in standardized sizes to cover a wide spectrum of the patients. The internal shape of the femoral component is made up of five planes that ensure suitable imposition of replacement on the distal part of the femur be the surgeon.
Physiological movement in a knee joint [2] is a complex combination of rolling, rotational and sliding movement. A knee joint loading during walking on the plain surface is a several multiple of body weight. During walking the stairs, up the hill and during carrying of loads, the joint loading increases. These facts determine the level of requirements on implants design, surgery methods and on the postoperative physiotherapy.

There are a lot of causes which may lead to knee joint diseases [3–8]. Most often it is degeneration (aging and wear) of an articular cartilage – primary and secondary osteoarthritis. The causes of primary osteoarthritis are not exactly known. Secondary osteoarthritis is caused by another disease (traumatic disease), metabolic disease (podagra) rheumatic arthritis, etc.

Fig. 1. 3D model of patient’s knee joints.

2. ACQUISITION OF PATIENT’S KNEE JOINT CT DATA

As the research object, the left patient’s knee joint (or acră femoral joint) was taken. Physical parameters of the patient: man, date of birth 1969, height 180 cm, weight 85 kg.

Computed Tomography (CT) data as output from the magnetic resonance imaging (MRI) were obtained in cooperation with St. Anne’s University Hospital Brno.

2.1. Technological Analysis of CT Data Transformation into the Stl Format

The output CT data in the form of the standard DICOM [1] format were used for generation of a 3D knee joint model. By the data decoding, i.e. by finding of a corresponding intensity values in pixels, the information about material properties of scanned objects were obtained. Complete information about the location of required knee joint section was obtained by subsequent filtration of a soft tissue (biceps, sinews, ligaments, vessel, and nerves) different from bone tissue. For this purpose, was used appropriate software (external cooperation with FIT BUT) to obtain the final model of a knee joint in Stl format Fig. 1, which is sufficient for further data processing. 3D model of a knee joint prepared in this manner is smoothed by means of a special mathematical computational procedure, eventually can be fast edited.
3. SCANNING OF A KNEE JOINT FEMORAL COMPONENT

For proper modelling process of a new type of a femoral knee joint component by means of software Catia V5 R17 was necessary to obtain a 3D model of an existing implant. By prior arrangement with orthopaedist, scanning of commonly used surgical implant of a size 5 produced by Beznoska s.r.o. Tab. 1, Fig. 3 was realized by means of reverse engineering technology (optical scanner Atos [9] from the company GOM GmbH – see Fig. 2, which is situated at the graphical working place at FME BUT). Further, the scanned data were verified within the CAD software (real solid models were not available).

Table 1. Row of dimensions – femoral component (alloy Co-Cr-Mo) [10].

<table>
<thead>
<tr>
<th>Size</th>
<th>Design</th>
<th>Order number</th>
<th>Dimensions [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A – cross</td>
</tr>
<tr>
<td>1</td>
<td>L</td>
<td>351000</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>352000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>351002</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>352002</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>351004</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>352004</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>351006</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>352006</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>L</td>
<td>351008</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>352008</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>L</td>
<td>351010</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>352010</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Optical scanner Atos [9].
3.1. Processing of Patient’s Knee Joint Scanned Data

Collection of scanned knee joint data in Stl format was trimmed to maintain only the required part of our interest – the part of patient’s left knee joint (distal femur part, see Fig. 4). Subsequently, it was necessary to make further improvements to this trimmed model.

Low-quality border created by automatic trimming was re-treated by cutting with a flat surface, incurred hole was automatically closed. Quality of the modified Stl model was checked by virtue of possible presence of inverted normal lines. In the next step, a surface covering (or copying) the polygonal Stl mesh was created. Created surfaces are covering the knee joint outside surfaces (Stl mesh in this case) with a certain deviation, which depends on the scanning accuracy and on the accuracy settings of created surface. In our case the maximum deviation between the polygonal mesh and created surface (at the location where the femoral component is positioned) is +0.219 mm – this is the sufficient value for this purpose. Finally the solid knee joint model was created from the surface knee joint model. In Fig. 5 there is a simplified scheme of a knee joint model creation from the source STL data. Furthermore, the extent of damaged places to remove was made with another software (accordingly to the data, surgical assessments and stress/strength limitations). The same software was used for fitting of the 3D model implant the modified bone.


Fig.5. Polygonal knee joint mesh – Surfaces on the polygonal knee joint mesh – Knee joint closed surface model – Knee joint solid model.

3.2. Processing of Femoral Component Scanned Data

Similarly to the scanned knee joint, the knee joint implant was modified as well. In the first step it was necessary to modify the scanned Stl model and to create a surface implant model, eventually the solid implant model. The triangles on the inner implant side were removed. In this way the Stl model was simplified and the inner side of implant was prepared for a new surface allowing the knee joint imprinting. A boundary of modified Stl model was of poor quality again and thus it was trimmed by the surface created for this purpose. Resulting and modified Stl implant model (point cloud) without internal surfaces and with trimmed boundary is in Fig. 6. Also in this case the quality control of Stl model and correction of erroneous elements was performed.

Fig.6. Stl model of trimmed implant – point cloud.

The necessary next step was creating of surfaces on the modified triangular mesh. This operation was accomplished at once – all 99% of Stl implant model points ranged in the tolerance given when creating the surface. Average distance deviation of created surface from the source triangular mesh was set to a value of 0.1 mm. Distance tolerance of free surface edges from the Stl implant model edges was set to a value of 0.08 mm. The maximum surface deviation from the polygonal mesh was 0.167 mm and maximum deviation of free edges was 0.05376 mm. All values are sufficient for this purpose. In Fig. 7 there are shown locations with maximum distances between the triangular mesh and created surface.
By reason of imprinting the knee joint model it was necessary to close all the implant surfaces. For this purpose, the rim was created round the boundaries of the implant surface model. The rim was also created because of the need of sufficient quantity of the implant material for a knee joint imprinting, because trimming of the original Stl model considerably reduced its thickness especially in the area of both condyles. Afterwards, the upper rim border was closed by another surface and thus the closed surface implant model was created. As a final step the closed surface implant model was converted into the solid model, see Fig. 8.

4. DATA DISTRIBUTION AND THE SMOOTHING METHODS

The technological process of reverse engineering is realized in several steps. First of all the data are scanned Fig. 9. After that there is a data pre-processing performed to eliminate the errors incurred by measurement inaccuracy. Finally the individual scanned data is aligned, reduced, smoothed and integrated. Subsequently the polygonal mesh is obtained by the software processing of a point cloud, see Fig. 10. For processing and visualization of a point cloud the B – spline surfaces were used.
Fig. 9. Scanning the data.

Fig. 10. Polygonal mesh.

4.1. B–spline Surfaces

NURBS curves (mathematically described free curves) provide the basis for NURBS surfaces (Non Uniform Rational B–Spline). They are a rational specialization of B–spline curves. The principle is in the use of corresponding equation for B–spline curve in the space of homogeneous coordinates. Nowadays, NURBS represent the industrial standard in geometrical modelling and they are used in many design tools, because they make possible not only to define the free form surfaces, but also the classical geometrical objects like ball, cylinder or cube.

The basis of B–spline curves [11–15] is the B–spline function which is defined as recurrence relation. It means there is a knot vector given, which is non decreasing sequence of positive real numbers, and which represents the time shape of the curve. The curve elements are called knots. When the difference of the two neighbour knots is the same for the entire vector, there is an equidistant knot vector. In the opposite case there is inequidistant knot vector and it will be used for interpolation. Non decreasing sequence of positive real numbers \( t = (t_0, t_1, \ldots, t_s) \) is called knot vector. B–spline function \( N_i^p(t) \) of the grade \( p \) is then defined as:
\[ N_i^p(t) = \frac{t-t_i}{t_{i+p}-t_i} \cdot N_i^{p-1}(t) + \frac{t_{i+p+1}-t}{t_{i+p+1}-t_{i+1}} \cdot N_{i+1}^{p-1}(t), \] (1)

where \( 0 \leq i \leq s-p-1, 1 \leq p \leq s-1, \frac{0}{0} := 0 \).

B–spline curve grade \( p \) for control points \( P_i \) and knot vector \( t = (t_0, t_1, ..., t_{n+p+1}) \) is defined as:

\[ C(t) = \sum_{i=0}^{n} P_i \cdot N_i^p(t), \] (2)

where \( N_i^p \) are base B–spline functions. If the weight \( w_i \) is added to each point, then the NURBS curve is created. Because all the weights in algorithm are set implicitly to the value of 1, formula (2) in this form corresponds to the given NURBS curve, because the denominator equals one.

Advantage of proposed method is in a simple mathematical representation of complicated surfaces without dividing and following consolidation – thus it is used in present methods.

5. IMPRINT OF PATIENT’S KNEE JOINT INTO THE MODIFIED FEMORAL COMPONENT

Modified patient’s knee joint with femoral component (implant) were constrained to each other by means of suitably selected planes to obtain subtraction of patient’s knee joint model from femoral component model. Result is shown in Fig. 11.

![Fig. 11. Final model of the femoral component.](image)

5.1. Final Verification of Femoral Component Model

Final model of modified femoral component was exported back into the Stl format, which was subsequently used for verification by means of rapid prototyping technology (Fused Deposition Modelling method - FDM). Printing of the femoral component physical model Fig. 12 was realized on Dimension 3D printer Fig. 13. Thermoplastic material from which the model of femoral component is printed is winded in the form of a wire on the coil. Rollers push the wire into the heated nozzle and the thermoplastic material is melted. In this...
case the rollers act as the piston – they push the melted material. As the melt is being released from the nozzle, it’s cooling down and solidifying and coupling with previous layer of material is established. After a layer creation the base is lowered of the layer thickness value. For supporting of overhanging component planes it is necessary to use the additional supports.

![Fig.12. Final printed model of the femoral component.](image)

![Fig.13. 3D printer Dimension.](image)

6. **CONCLUSION**

3D model of a new type of a femoral component was created from the 3D model of patient’s left knee joint (the distal femoral part) and from the scan of a knee joint femoral component of size 5 produced by Beznoska, s.r.o. company using software Catia V5 R17. The aim was to obtain (through the suitable modification of existing implant) model of a knee joint replacement tailored exactly for a patient. Outer shape of a knee joint replacement is the same as on the implants used today and the inner shape is an accurate copy of a knee joint surfaces – it is a negative surface of a knee joint. Due to the shaping of inner implant surfaces it is supposed the loading of patient will be reduced and a knee material will be preserved for possible checking operation. The main advantage of this research lies in modification of only internal shape of existing tested implant according to the patient data. External implant shape remains unchanged. New implant type will be the subject of next research and testing in given area as other important steps like precise robotic machining of the bone.
ACKNOWLEDGEMENTS

The research was supported and co-financed by the projects CZ.1.07/2.3.00/30.0039 of Brno University of Technology and Specific Research of BUT FME FSI-S-13-2138 Research of advanced technologies for competitive machinery.

REFERENCES


A Roadmap to Model Based Manufacturing

Dr. Martin Hardwick

Department of Computer Science, Rensselaer Polytechnic Institute, USA
hardwick@steptools.com

ABSTRACT
Model based manufacturing replaces manufacturing codes with integrated data models. The new models reduce costs, increase accuracy and improve performance, but by definition they are new and some aspects of code based manufacturing have been in place for hundreds of years. We describe a five step roadmap that deploys model based manufacturing by implementing services. At the end of the map, the new models have replaced the traditional codes, and shop floors are controlled as a web of machines and sensors operating within a high level programming framework.

KEYWORDS: Computer Aided Manufacturing, Data Exchange, Product Data Standards, Interoperable Manufacturing

1. INTRODUCTION

The benefits of model based manufacturing are dramatic. With a model a job shop can send a description of a manufacturing issue to an outside expert, and the expert can return a solution and send it back (see Fig.1). Implementing model based manufacturing is difficult, however, because the model must be able to carry all the necessary information. Today, the only model generally available is the manufacturing codes used to control machines (see Fig.2). These codes cannot be used for reasoning algorithms. Therefore, they cannot be used to describe problems and solutions.

In this paper we describe a plan that will deploy models that can carry the necessary information. The plan starts with a basic service that allows tooling vendors to deliver process recommendations to customers. It continues with an integration service that allows metrology vendors to deliver as-built dimensions for fitting parts into assemblies. The next stage is a construction service to build very large product models across a supply chain. The fourth stage is a simulation service for all manufacturing operations. The final stage is a prediction service to show the manufacturing results of a given set of inputs.

Each stage relies on the deployment of new technology and the successful completion of the previous stage. For example, the second service becomes possible when the first service makes industrial strength CAM translators available. The third service becomes possible when the second service makes the first commercial-off-the-shelf CAM translators available. The fourth service then begins the transition away from the codes and the last service completes the deployment.
The next section describes the concept of model based manufacturing in greater detail. The third section describes a model for model based manufacturing that has been built and tested over a ten year period. The fourth section describes compares this model with other models. The fifth section describes the five step roadmap. The last section contains some concluding remarks.

2. MODEL BASED MANUFACTURING

A team called STEP-Manufacturing has been testing an integrated model for manufacturing for 10 years [1]. The research has shown that a rich, integrated data model for manufacturing can deliver many enhancements to advanced industries. For example, the average machining process can be made 15% faster. However to get these optimizations interfaces must exist to read and write the data into optimization systems. Three barriers have been identified to making this happen:

1. Development of a model. The manufacturing model must be able to carry enough information for experts to understand the issue, and it must return enough information for a shop to deploy the recommended solution.

Given the nature and complexity of manufacturing it is inevitable that the model will be large. Fig. 3 shows some of the contents of a typical manufacturing model and how they relate to design, planning and manufacturing. STEP-Manufacturing has tested a model called STEP-NC that is now trusted as a solution for model based manufacturing [2].

2. Translator implementation. A model can only be used if programs exist to translate its data to and from systems.
System vendors are reluctant to implement translators because they prefer solutions to be implemented within their systems. Plus if the model is large, then the cost of developing the translators is significant. However, if a model has sufficient momentum then implementation can be required in purchase contracts. STEP-NC has been made into ISO standard 10303-238 to get this momentum, and the CAD/CAM system vendors have announced their intention to implement interfaces that read and write this model.

3. Piloting. The requirement to implement translators must be made credible by pilot projects. This is the current stage for STEP-NC. The roadmap is the required series of pilots.
3. A MODEL BASED MANUFACTURING MODEL

Many systems make design and manufacturing data. Each system has its own data formats so when multiple systems are used the same data has to be entered many times leading to redundancy and errors. Although repeated data entry is not unique to manufacturing, it is more significant because product data is complex and three-dimensional (3D). The National Institute of Standards and Technology has estimated this costs the United States 90 billion dollars annually [3].

The redundant data entry problem is solved by models to carry design and manufacturing data between systems. The most successful have been data exchange standards. The first ones were national and focused on geometric data exchange. They include the Standard d’Exchange et de Transfer (SET) in France, the Verband des Automobilindustrie FlächenSchnittstelle (VDAFS) in Germany and the Initial Graphics Exchange Specification (IGES) in the USA. Later a unifying effort was started under the International Organization for Standardization (ISO) to produce one standard for all aspects of technical product data and named STEP for the Standard for Product Model Data. Today nearly every major CAD/CAM system has a STEP interface.

The goal of STEP is to define models for the entire product life cycle. The initial focus was on the exchange of three dimensional models of parts and assemblies. STEP was the first neutral data standard to enable solid model data exchange. While STEP was the first to achieve this goal, after its lead other standards such as XVL, 3DXML, 3DPDF and JT also began to support the exchange of solid models, but in a more limited easier to process form.

STEP is divided into Application Protocols. The ability to support many protocols is a key strength but also a cause for complexity. Manufacturing needs many such protocols so it may not be possible to make simpler specifications. The most important is a description of the required geometry, next is a description of the required tolerances, but others such as the material properties and surface finishes are also required. This information is known as PMI for Product Manufacturing Information.

Furthermore if manufacturing solutions are to be defined then manufacturing process and tooling data must be included. STEP-NC extends STEP to include this data [4]. A STEP-NC process describes a series of operations that add or remove material. The material volumes are defined as features, generic shapes, or as the implied result of tool movements. The operations are sequenced into workingsteps belonging to a workplan that has a geometric setup. Different types of workplans can be used to make the workingsteps conditional or concurrent.

Each workingstep performs an operation such as rough milling a pocket [5]. The operation is applied to a feature that can be described parametrically or implicitly. The operation itself is also described as a set of parameters, or as a tool path. The former make the machining program more resource independent. The latter are fixed to specific resources but allow for more error checking.

A STEP-NC program can also include a description of the machine and its associated tooling and fixtures. At the users discretion they are modelled as assemblies with kinematic movement and positional accuracy data. The result is a complete model of the product, process and tooling that can be shared with vendors, users and operators as shown in Fig. 4.

Testing of STEP-NC has been performed by a team from industry, government and academia. Table 1 lists the history divided into five periods. In the first period, the emphasis was on faster art to part using feature recognition. In the second period, CAM to CNC interoperability was shown by sending STEP-NC data from four different CAM systems to
two CNC controls, where it was used verbatim to drive both AC tool tilt and BC table tilt machine configurations. The third period tested closed loop machining by including data to both machine and measure a part in a STEP-NC file. The fourth period worked on automated feed-speed optimization with extensions proposed for modelling the cross section of the material volume removed by a tool path. The fifth period extended the fourth to include tool wear modelling and did the first machine tool accuracy modelling.

Fig. 4. Machining using STEP-NC: (a) workingsteps add or remove data volumes, (b) product models define geometric dimension and tolerance constraints, (c) manufacturing features with associated cutting tools, and (d) fixtures in setups on machine tool models.

<table>
<thead>
<tr>
<th>Period</th>
<th>Capabilities shown</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 2000</td>
<td>Tool path generation from manufacturing features</td>
<td>Faster art-to-part</td>
</tr>
<tr>
<td>Feb 2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 2005</td>
<td>CAM to CNC data exchange without post processors</td>
<td>CNC interoperability</td>
</tr>
<tr>
<td>May 2005</td>
<td>Integration of STEP CAD GD&amp;T data with CAM process data</td>
<td>Integrated machining and measurement</td>
</tr>
<tr>
<td>Jun 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul 2007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec 2007</td>
<td>Cutting tool modelling per ISO 13399; cutting cross-section modelling</td>
<td>Feed and speed optimization</td>
</tr>
<tr>
<td>Mar 2008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct 2008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 2009</td>
<td>Tool wear modelling; Machine modelling; Accuracy modelling</td>
<td>Closed-loop manufacturing</td>
</tr>
<tr>
<td>Sep 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 2010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 2012</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 – History of STEP-NC testing

STEP-NC was released as a standard in 2007 by the International Standards Organization (ISO) in Geneva, Switzerland. Table 1 shows how industry users such as
Boeing, Airbus, Pratt and Whitney, General Electric, Lockheed Martin, ISCAR, Sandvik and Scania as well as the National Institute of Standards and Technology (NIST) have been evaluating its functionalities. In June of 2012 they declared success and requested implementation by the CAM vendors. A meeting was organized by the Organization for Machine Automation and Controls (OMAC). The CAM vendors for Catia, NX and Mastercam responded and the first prototypes for the new translators were demonstrated at a meeting in Richmond, Virginia on June 24 and 25, 2013.

4. ALTERNATIVE MODELS

Developing a model for all of manufacturing is a difficult task so most alternatives focus on a subset of the activities.

For example, in machining several models have been developed for tool path control. The APT language was first developed in the late 1950’s and is still the primary language for much machine tool control. In the 1980’s a language called BCL was developed as a machine independent replacement for machine codes [6]. And in recent years ISO 14649 has been developed as a specification for machining operations using much of the technology developed for STEP [7].

Multiple standards have also been developed for metrology. One of the most successful has been DMIS which is a programmatic language for describing how to measure a part very similar in structure to APT. There have also been several developments of interfaces for the different stages of measurement planning. Recently these standards have been enhanced by a framework called QIF and there has been discussion of replacing them with a vendor developed format called CAPVidia [8].

Additive manufacturing is a third area where there have been recent developments. The STL data format describing facets is used as the input to additive manufacturing processes. It is simple to implement and recently there have been proposals to make it more accurate by adding surface finish descriptions and there is an open question over how complex honeycomb shapes will be modelled [9].

STEP is different because it was designed to cover the entire product life cycle. The principal rivals to STEP are the vendor formats in the major CAD systems. There have been occasions when they have not been so extensible (e.g. the transition from Catia v4 to v5 was difficult). However, in recent years the changes have been less dramatic and subsets are being made available as formal and informal standards such as 3DPDF, XVL, JT and 3DXML [10].

The range of alternatives mean the traditional STEP architecture has become too monolithic. A new edition allows the data to be split into files of different types linked by anchor and reference sections [11]. Fig. 5 shows the key qualities. As the figure shows there are millions of systems that read and write data described by the current architecture but to process the data they must read the whole file. Splitting the information into pieces linked by URL’s (Fig. 5b) increases security because sensitive data such as tolerances can be put into sub-files that are locked, and performance becomes better because rapid visualization data can be put into sub-files that are easy to locate and process.

Dividing the data into multiple sub-files can make data management more complex so the new format also allows them to be gathered into ZIP files (Fig.5c). The same approach is used in many other systems for example PowerPoint (.pptx) files and Word (.docx) files are ZIP files. Last, but not least, a JavaScript binding is included (see Fig.5d). This may be the most controversial feature because of misconceptions about security issues but it enables the new files to define behaviour. Therefore, a product data file can have methods to show the behaviour of the product, for example, how an assembly moves, or how an operation removes material.
5. ROADMAP TO DIGITAL MANUFACTURING

Table 2 summarizes the Roadmap. The goal of each stage is to deliver a service that is useful to industry (see column 2). The benefits are summarized in column 3 with quantifications when available. Column 4 discusses the necessary pre-conditions with a successful conclusion to the previous stage being the most important and unstated condition. Column 5 spells out some key consequences for each stage.

The CAM Tooling pilot gets the whole process started. The STEP-Manufacturing team has been careful to design this pilot so that it has well defined benefits. One of the key inputs is a DLL that can be used by industry to create data and simulate manufacturing results [12]. Fig. 6 describes the CAM Tooling pilot in more detail. In this pilot an airframe company will send models of its production parts to the cutting tool vendors for optimization. The tooling vendors will then select their best tooling for the given machining task and return a process description. The assumption of the pilot is that because the tooling vendors are continually making improvements, they will have better tooling available and the changes will result in more efficient manufacturing programs for the airframe company.

A key result of the CAM Tooling pilot will be applications that make the manufacturing model data. This result will make the second pilot possible. In this pilot metrology and machining systems will be integrated so that the former can measure as-built assemblies for the latter. When large flexible structures are being built there are always minor differences because of different manufacturing orders and different loads at different times. The pilot will automate the process of compensating for these differences in five axes.

The Made-to-Measure pilot will integrate machining and measuring systems so that the latter can correct the former. In the first application this means measuring the as-built structure of the airframe and correcting the geometry of the panel to optimize the fit. The project is investigating how to communicate between the two systems. Both will have their own internal models of the data. Even if they both read the same STEP data the internal structures will be different. Consequently there needs to be a way to ask questions without knowing the details of the other system. One method being investigated is to send associative facets between the systems.
The Made-to-Measure pilot is scheduled for deployment in 2015. The next pilots are further out. If all goes as planned then the Made-to-Measure pilot will result in the CAM vendors making STEP-NC translators commercially available. This in turn will make STEP and STEP-NC data much more widely available and the plan assumes universities and others will be interested in building very large manufacturing models by linking lots of smaller models.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Application</th>
<th>Benefits</th>
<th>Other pre-conditions</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tooling Services (MRL 4)</td>
<td>Job shops get tool vendors to make tooling and process recommendations</td>
<td>Reduced cycle time/less tool wear/more reliable (=15% better)</td>
<td>DLL or similar tool to make, view and simulate the machining process.</td>
<td>Organizations have translators and can make digital manufacturing data.</td>
</tr>
<tr>
<td>Made-To-Measure Services (MRL 5)</td>
<td>Contours and hole locations/axes for composite panels and other parts adjusted to meet dimensions of large flexible structures.</td>
<td>More accurate large flexible assemblies. Reduced energy consumption because of higher accuracy. Simpler repairs.</td>
<td>Product data with tolerances for the structures. Manufacturing data with contours for the panel parts.</td>
<td>Manufacturing translators available as COTS products. Understanding of how to share data between systems.</td>
</tr>
<tr>
<td>Massive-Product Model Services (MRL 6)</td>
<td>Universities and other interested organizations build massive product models.</td>
<td>Reduced supply chain costs. Increased interest in product models.</td>
<td>Web object models for linking the product models.</td>
<td>More complete product models. Unexpected synergies.</td>
</tr>
<tr>
<td>Factory-Simulation Services (MRL 7)</td>
<td>Integrated simulations show how complete products are built in digital plants.</td>
<td>Less waste. Reduced lead times. Better planning.</td>
<td>Intelligent process simulators that can apply decision making algorithms to CAD and CAM data.</td>
<td>Manufacturing systems are programmed using a modern language such as JavaScript</td>
</tr>
<tr>
<td>Digital Production Services (MRL 7)</td>
<td>Results prediction using highly accurate resource models.</td>
<td>Reduced costs for new facilities. Easier make verses buy decisions.</td>
<td>Machines with good accuracy prediction. Voxel based simulation.</td>
<td>Direct model execution by CAD, CAM and CNC systems</td>
</tr>
</tbody>
</table>

This third pilot will make the limited sharing of the Made-to-Measure pilot more general. Fig. 7 illustrates. In this figure a part program defined by STEP-NC is linked to workpiece models defined by STEP. The result is three databases. The two originals and a third that is an integration of the others.

- In perceived order of difficulty
  - Chip load per tooth
  - Material removal rate
  - Spindle speeds
  - Feedrate
  - Cutter type selection (without toolpath change)
- Machining direction
  - Radial engagement
  - Axial (vertical) engagement
  - Cutter type selection (with toolpath change)
- Different machining techniques (tool path generation)
- Entries and exits
- Different machining order?
- Different in-process results

Fig. 6. CAM Tooling Pilot

The Massive-Product-Model service is predicted to result in second tier translators for the manufacturing model data. The Factory-Simulation service is predicted to result in the
wide spread programming of machine simulations in a modern language such as JavaScript. At the same time, and in the background work will have been proceeding on developing intelligent algorithms for simulating the results of machining processes at the voxel level. The issue with voxels is that their data volume is overwhelming, but the continued expansion of storage devices and the development of algorithms that only “volxelate” when necessary should make widespread simulation available. If so then there will be a demand for the machine vendors to show what they can do by applying digital processes to digital parts. If it has not already happened, this will complete the roadmap by causing the machine vendors to start executing the manufacturing models directly on their controls.

6. CONCLUDING REMARKS

In this paper we have describe a roadmap for deploying digital manufacturing models by developing useful manufacturing services. Each service leads to wider spread usage which makes the next level of service possible and even wider deployment. Fig. 8 summarizes the descriptions in stair steps. The necessary inputs for each step are shown above each stage and how that stage uses the new technology in Part 21 Edition 3 is shown on the right.

![Fig. 7. Linking product, process and fixture models](image)

![Fig. 8. Summary of the roadmap.](image)
A roadmap this long is unlikely to be executed perfectly. The key to success is to focus on CAM to CAM data exchange as an enabler for system independent manufacturing models. If CAM data is being exchanged then the models will be available for use by other applications and the rest of the plan can be summarized as making the usage so wide that the machine vendors are able to run the models directly on their machines eliminating the requirement for the old style G-codes.

ACKNOWLEDGEMENTS

The author would like to express his appreciation toward the entire STEP-Manufacturing team for their collective effort to test and promote the ISO 10303-238 standard. To list only a few, they are: Sid Venkatesh, David Odendahl and Leon Xu from The Boeing Company; Fred Proctor and John Horst from the National Institute of Standards and Technology; Mikael Hedlind, Magnus Lundgren and Andreas Archenti from KTH Royal Institute of Technology; Bengt Olsson from Sandvik Coromant; Doron Cohen from ISCAR; Bob Erickson from Pratt and Whitney; Charles Gilman from General Electric; Alain Brail from Airbus (retired); Larry Maggiano from Mitutoyo USA; and David Loffredo and Joe Fritz from STEP Tools, Inc.

REFERENCES


A Study of the Surface Integrity after Machining by means of Non-Destructive Testing Methods

P.L. Tam¹, S.B. Hosseini², J. Holmberg³, P. Lundin⁴, L. Mattsson⁵, L. Nyborg¹

¹ Chalmers University of Technology, Department of Materials and Manufacturing Technology, 412 96 Gothenburg, Sweden.
² AB SKF, 415 50 Gothenburg, Sweden.
³ Swerea IVF AB, 431 22 Mölndal, Sweden.
⁴ Stresstech Oy, 144 21 Rönninge, Sweden.
⁵ KTH Royal Institute of Technology, Department of Production Engineering, Sweden.

E-mail address of the corresponding author: eric.tam@chalmers.se

ABSTRACT

During metal machining, depending on the cutting conditions, surface and subsurface microstructure alteration are occasionally observed. These alternations are normally referred as “white” and “dark” layers. Due to their different mechanical properties in comparison to the unaffected material, they will have an impact on the finished part. Controlling the quality of the machined parts regarding the surface microstructure alteration by means of non-destructive testing (NDT) methods would be beneficial from production point of view. In this study, the surface integrity of AISI 52100 steel machined at different cutting conditions resulting in white and dark layers with different characteristics were studied. Surface topography, microstructure and residual stresses were examined by using light scattering, optical microscopy and x-ray diffraction (XRD) techniques. Whilst surface characterization was emphasized, one NDT method – magnetic Barkhausen noise (BN) technique – is not well defined for this purpose. The correlation between all the applied techniques was therefore investigated and a preliminary model was developed for the influence of surface roughness, stress conditions and white and dark layer thicknesses on BN signal.

KEYWORDS: machining; white layer; light scattering; x-ray diffraction; Barkhausen noise.

1. INTRODUCTION

Machining is an important manufacturing process used in industry to machine the parts to the final desired tolerances. Machining by means of grinding, turning or drilling are some commonly used practices. As presented by Hosseini [1], Akcan et al. [2] and Bosheh and Mativenga [3], surface microstructural alternation that occurs during machining and is referred as “white and dark layers”. Depending on the formation mechanisms, mainly thermally or mechanically induced, the layer could have either beneficial or detrimental effects on the
machined components [1,2]. Therefore, it becomes interesting and important to study the surface integrity generated at different machining conditions. However, although the machining condition can be chosen in such a way to promote different types of white layer, in the current investigation, only thermally-induced white layers by turning were studied. To characterize the machined surfaces, several non-destructive testing (NDT) methods including light scattering, x-ray diffraction (XRD) and magnetic Barkhausen noise (BN) techniques were used for near-surface characterization. In addition, cross-sectional images captured by optical microscope were used to distinguish the characteristic features of individual sample. Basically, the current understanding about the BN signal to the near-surface measurement is not clearly defined; therefore, one of the major objectives here is to gather the characterized results from different techniques, correlate them, and explain this particular BN information. The ultimate goal would be introducing this BN technique as a novel NDT method for white layer detection on the machined surface.

2. EXPERIMENTAL AND DATA ANALYSIS DETAILS

2.1. Machining condition

A simple experimental design of different machining conditions (see Fig. 1) was applied to create different conditions of surface integrity. The length of the flank wear on the cutting inserts ($V_b$) and the cutting speed ($v_c$) were the two variable factors. The cutting inserts used were diamond-shaped polycrystalline carbon boron nitride (PCBN) of grade BNX10 with 0.12 / -25° chamfer, 7° clearance and 0.8 mm nose radius. The depth of cut was 0.08 mm and the feedrate was 0.08 mm/rev. The wear of the inserts were introduced by using the inserts on samples with a comparable diameter prior to the final cut. The water-based cutting coolant containing 5.0 % of emulsion was applied on the rake side with a pressure of 506 kPa.

![Fig.1. The test parameters varied and the corresponding sample numbering.](image)

2.2. Material

High carbon and chromium-containing AISI 52100 steel was used in this study. The nominal chemical composition is given in Table 1. The steel was treated to have a tempered martensitic structure with hardness of 61 ± 1 HRC. The microstructure of the steel prior to the final hardening treatment was characterized to have evenly distributed spheroidized carbides (M$_3$C) in a ferritic matrix [4]. All the tested samples were in ring shape with an outer and inner diameter of 180 mm and 150 mm, respectively, and with a width of 55 mm.
Table 1. Nominal chemical composition of AISI 52100 in wt. [%], Fe bal.

<table>
<thead>
<tr>
<th>Material</th>
<th>C</th>
<th>Cr</th>
<th>Si</th>
<th>Mn</th>
<th>S</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AISI 52100</td>
<td>1.04</td>
<td>1.45</td>
<td>0.25</td>
<td>0.35</td>
<td>&lt;0.025</td>
<td>&lt;0.025</td>
</tr>
</tbody>
</table>

2.3. **Light scattering and surface topography measurements**

It is well known that surface micro- and nanotopography strongly influences the reflected light distribution and this can be used as a fingerprint of height variations, depicted via the change in total scattered intensity and the lateral distribution reflected by the angular distribution of the scattered light [5,6]. In addition, small changes in optical properties of the material at the surface can cause light scattering even for the smoothest surface [5]. In this study, light scattering is aiming for detection of white layers of the four different generated surfaces based on potential topographic changes also occurring when white layer is being formed. The out of incidence plane scattering was measured rather than in plane scattering. The latter is completely dominated by the geometrical as well as diffraction scattering from the grooves, cutting edge roughness of the insert and the feed rate. The light scattering measurements were performed both as partial integrated scattering measurements, using a Si-photodetector capturing the scattered light and as an imaging technique based on a Canon 7D digital camera provided with a macro lens. A laser diode operating at 635 nm wavelength was used as the light source illuminating the surface at an oblique angle. The surface topography investigations were carried out with the Zygo 7300 white light interferometer using a 10x objective and the analysis was made with the Zygo Metropro software, by randomly picking five 800 µm line profiles from within the measured surface of 1090 \( \times \) 1090 \( \mu m^2 \) at five positions separated by several mm in the axial direction. The shortest surface wavelength captured is \( \sim 2.0 \mu m \) and no additional filtering was done.

2.4. **X-ray diffraction and residual stress measurements**

Generally, high tensile residual stress on the machined surface can be devastating to the component and thus its measurement is of high importance. The residual stresses were determined by means of x-ray diffraction (XRD) in two different directions, i.e. the axial (feed) and the circumferential (cutting) directions, as shown in Fig. 2. The residual stress measurements were carried out with a XStress X3000 G2R equipped with a CrK\( \alpha \) source (\( \lambda = 2.2897 \text{ Å} \)) and a 3 mm diameter collimator. The diffractometer was calibrated with a stress-free ferritic iron powder with the (211) plane diffraction peak positioned at 156.4° [7]. All the other settings are listed in Table 2. Prior to the stress calculation, all the measured peaks had their backgrounds removed by a constant function and fitted with a cross-correlation method [7]. The nature and level of the residual stress can be evaluated by the \( \sin^2 \psi \) technique. This method assumes that a planar bi-axial stress state as the penetration depth of the irradiated beam is only a few microns, and in this region the normal stress component is assumed to be zero.
Fig. 2. Surface stress measurement by means of XRD.

### Table 2. Experimental settings and material data in the XRD measurement

<table>
<thead>
<tr>
<th>Equipment</th>
<th>XStress X3000 G2R</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray voltage</td>
<td>30 kV</td>
</tr>
<tr>
<td>X-ray current</td>
<td>6.7 mA</td>
</tr>
<tr>
<td>X-ray source</td>
<td>CrKα</td>
</tr>
<tr>
<td>Exposure time</td>
<td>30 s</td>
</tr>
<tr>
<td>Phi angles</td>
<td>0°, 90°</td>
</tr>
<tr>
<td>Psi oscillation</td>
<td>±5°</td>
</tr>
<tr>
<td>Diffraction angle</td>
<td>156.4°</td>
</tr>
<tr>
<td>Poisson’s ratio</td>
<td>0.30</td>
</tr>
<tr>
<td>Young modulus</td>
<td>211 GPa</td>
</tr>
</tbody>
</table>

### 2.5. Barkhausen noise measurements

Barkhausen noise measurements were conducted with a Rollscan 300 Digital Barkhausen Noise analyzer connected with the BN probes of different powers. To increase the surface sensitivity in order to correlate the BN signal to the uppermost surface, high magnetizing frequency at 1000 Hz was applied [8]. Also, to enhance the reliability, each of the presented data was measured at least 3 times at 4 different positions. The signal acquisitions were done in the same manner as for the stress measurements. The magnetizing frequencies and voltages for individual probe are listed in Table 3. Other parameters for BN analysis including the analyzing filter range, number of bursts and sampling frequency were set at 25-600 kHz, 20 and 2.5 MHz, respectively.

### Table 3. The settings of Barkhausen noise (BN) probes in different measurements

<table>
<thead>
<tr>
<th>BN probe</th>
<th>Magnetizing voltage (V)</th>
<th>Magnetizing frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>1000</td>
</tr>
<tr>
<td>B</td>
<td>14</td>
<td>1000</td>
</tr>
<tr>
<td>C</td>
<td>16</td>
<td>1000</td>
</tr>
</tbody>
</table>

### 2.6. Microstructural examination

To investigate the condition of the microstructure in the surface after machining, optical microscopy was used. The samples for this were prepared by cutting out small segments from the machined samples. The microstructure investigation was performed after the topographic, residual stress and BN measurements. After mounting in hot-resin, the samples were prepared by grinding, polishing and etching. Extra care was taken during grinding since there was a risk
of transforming certain amount of retained austenite to martensite. The samples were first
ground with MD-Piano 120, 220 and SiC papers with grit ranging from 320 to 800.
Thereafter, the samples were polished in several steps using cloth and diamond-paste
suspensions of 9, 6 and 1 µm. To study the microstructures, the samples were etched with
1.5% nital solution (98.5 ml ethanol and 1.5 ml HNO₃). The microscope used was a Leica Letz
DMRX equipped with an AxioCam MRC5 camera. During the microstructural analyses, the
surfaces along the axial direction were investigated.

2.7. Data Analysis

To explore the practicability of the BN technique for near-surface characterization,
numerical data analysis was conducted by using the JMP® 10.0.0 software to investigate the
relationship between all the measured physical properties and the BN signal. Characteristic
features from the topographic measurement, microscopic study and stress analysis were taken
into account for this evaluation. The results were analyzed by means of scattering plots.

3. RESULTS AND DISCUSSION

3.1. Microstructural examination

The microstructural study shows that no or discontinuous white layer appears on the
samples that were machined by fresh cutting inserts, i.e. Samples 1 and 2. Meanwhile, uniform
and featureless white layers were observed on those surfaces machined by inserts with
excessive tool flank wear, i.e. Samples 3 and 4. The cross-sectional images of Samples 1 and 4
are shown in Fig. 3. By increasing the cutting speed, the average value of the white layer
thickness increases from 1.5 µm in Sample 3 to 2.5 µm in Sample 4. These measured
thicknesses have large scattering among different locations (i.e. around 1.0 µm) and the white
layer could have peeled off or damaged fully/partially during sample preparation. Still, a
certain tool flank wear appears to be needed to generate continuous white layers, and a thicker
layer was also observed for higher cutting speed.

Independent of the tool flank wear and cutting speed, the near-surface dark layers are
observed in all the samples. They can be seen at the uppermost part in Fig. 3a and beneath the
white layer in Fig. 3b. According to Hosseini [1], these dark layers can be considered as an
over-tempered structure (over-tempered martensite) and consequently they show lower
hardness as compared to both the white layer and the tempered martensite in the bulk [1-3].
The dark layer is 5.0 ± 2.0 µm, which is thicker than that of the upper white layer.

Fig.3. Cross-section of (a) Sample 1 shows no white layer and (b) Sample 4 shows the
uniform white layer on the surface after machining.
3.2. Light scattering results and surface roughness

As mentioned previously, the light scattering is measured in a spatial position off from the main light scattering plane as defined by the laser beam and the normal to the mean surface. The laser beam is incident at an angle of 60° to the surface normal and was pointing in the axial direction. Figure 4 shows the qualitative results from the four differently generated surfaces, where e.g. for Sample 1, there was almost no out of plane scattering, while for Sample 4 the surface created a considerable amount of scattering. In a next step, the changes in the surface roughness will be measured and correlated to the light scattering and formation of white layers.

![Light scattering results and surface roughness](image)

Fig.4. Images of light scattering of the four samples. Illumination is obtained from a red laser diode (635 nm) illuminating approximately 5 mm of the surface in the axial direction.

Table 4 lists the surface roughness measured in the axial direction across the grooves. The measurements are based on the average of 5 x 5 800-µm profiles on each sample. No clear relationship can be found in the measured average Ra values among the samples. The standard deviation is 0.02-0.03 µm for Samples 1 and 2, while it is larger for Samples 3 and 4. Ra-range is the difference between the largest and smallest average Ra between the five spots measured. The variation is so large that the average Ra of Sample 4 is practically the same as Samples 1 and 2. The only one sticking out is Sample 3 that in some areas had height undulations of a couple of micrometers over 700-800 µm surface wavelengths.

The variation from place to place in the axial direction is huge. For Sample 4, it went for a single profile from 0.177 nm to 0.524 µm Ra. With such a large spread, it ought to be difficult to make any correlation. Still, it is concluded that white layers lead to larger standard deviation and wider Ra-range in accordance with the present results.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Ra (µm)</th>
<th>Standard deviation (µm)</th>
<th>Ra-range (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.29</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>0.32</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.48</td>
<td>0.11</td>
<td>0.19</td>
</tr>
<tr>
<td>4</td>
<td>0.30</td>
<td>0.13</td>
<td>0.30</td>
</tr>
</tbody>
</table>
3.3. XRD results and residual stress at the surface

The near-surface residual stress levels on the machined samples determined by means of XRD measurements are plotted in Fig. 5. As can be seen in the figure, there is one major feature for distinguishing the fresh and worn tool cuttings. For the fresh insert, independent of the cutting speed and the direction of the XRD measurement, compressive residual stresses were always measured for the machined surface. On the contrary, surface stresses were tensile for the samples machined with worn tools. Then, according to Fig. 5, the circumferential directions (cutting) are associated with higher tensile stress levels as compared to those in the axial directions (feed). Hence, the formation of the white layers with worn tools leads to high tensile stresses. For the condition employed when having worn tool, white layer formation is expected to be thermally induced [1]. As can be seen, Sample 2 shows surface compressive stresses in both directions, even though it has discontinuous white layers as can be seen in the microscopic study [1]. The important message here is that Samples 3 and 4 with continuous thermally induced white layers are associated with high tensile surface stresses, which should have a significant impact on the BN signal measurement.

![Surface residual stress levels of Samples 1-4 determined by means of XRD.](image)

Fig.5. Surface residual stress levels of Samples 1-4 determined by means of XRD.

3.4. BN measurements

With high frequency BN measurements, it is expected that greater fraction of the signal comes from a near-surface region [8] and thus the result would be more relevant to indicate the white layer formation. As shown in Fig. 6, the results from different probes generally show comparable trends, even though the different amplitudes of the RMS values were measured in the magnetic BN signal. This was due to the fact that different magnetizing voltages were used by individual probe, and consequently, only the general trends could be compared. The RMS values decreases in the order of Sample 1, Sample 3, Sample 2 and Sample 4. Since the bulk material comes from the same charge, it is expected that the contribution from the bulk is the same for all specimens. The differences in data are thus mainly induced by varying surface integrity, i.e. the thicknesses of microstructural alterations and near-surface residual stress profiles. Moreover, as can be seen in Fig. 6, the minor difference in the RMS values observed between Samples 1 and 3 implies that their surface conditions are similar and the RMS value is clearly dependent on other factors than the presence of white layer. Still, when comparing
Samples 3 and 4, both having high tensile stresses and white layers thicknesses of between 1.5-2.5 μm, there is much lower RMS value for the latter sample. Clearly, there are several factors that contribute to the RMS value amplitude and therefore have to be considered in the analysis of the BN signal. For this reason, an extended data analysis has been initiated as outlined below in next section. In this part, it can be supposed that cutting speed has higher impact on the RMS value than the flank wear length of the cutting tools.

3.5. Data analysis

From the present study as well as the referenced literature [8,9], the RMS value of the BN signal is related to the microstructure (phases present and grain size) and surface stress distribution. Figure 7 outlines the relationship of the parameters addressed and the BN signal. With reference to Hosseini [1], microstructure alteration at the machined surface depends largely on the machining conditions. Higher cutting speed induces thermally activated white layer formation, which in turn could have significant impact on retained austenite within the white layer formed. In addition, when machining with worn tool, this leads not only to white layer formation, but also to high tensile residual stresses and greater surface waviness. By this means, all the above changes will trigger the BN signal level and each factor contributes to this measure in different ways. Figure 8 demonstrates how each of the factors influences the acquired BN signal. The data are analyzed via scattering plots by using JMP® software. A preliminary model for evaluating the RMS value of the BN signal is established in Equation (1). This equation indicates that the value should increase when having lower surface waviness, higher tensile stresses and lower compressive stresses as well as when white layer and dark layer thicknesses decrease. Still, much further effort will be needed to establish a more refined model to understand how these different aspects of surface integrity correlates with the measured BN signal.
International Conference on Advanced Manufacturing Engineering and Technologies

Fig. 7. Illustration showing factors that are expected to influence the acquired BN signals.

\[
\text{BN signal} = \frac{\text{Tensile stress level}}{\text{Compressive stress level} \times \text{Layer thickness} \times \text{Ra-range}} \tag{1}
\]

Fig. 8. Scattering plot showing how some characteristic measures related to surface integrity of hard machined parts correlate with the RMS value of BN signal in the Barkhausen noise measurement.

4. CONCLUSIONS

Thermally-induced white layers generated during machining were characterized and studied by means of different NDT methods. Depending on the progress of the tool flank wear, white layers of different thickness and uniformity were created. By increasing the cutting speed, the thickness of this featureless layer was further increased. The formation of white layer is shown to co-exist with a change in Ra-range in surface topography as a consequence of utilizing a worn cutting tool. Samples without and with discontinuous white layer were characterized to have compressive residual stresses in the surface, while surfaces with continuous white layers showed high tensile surface residual stresses. Among the techniques employed, the BN technique is the only one that is not well defined for surface characterization. Hence, an attempt to compare the results from light scattering and XRD with the RMS value of the BN signal was employed and the preliminary analysis based on empirical correlation indicates that this value is proportional to the tensile residual stress level, while it is inversely proportional to the white and dark layer thicknesses, Ra-range and the compressive residual stress level. A preliminary empirical model is thereby established, whereby a foundation is laid for further investigation to understand the relationships among all these factors.
ACKNOWLEDGEMENTS

Financial support from the Swedish Innovation Agency (VINNOVA) within the framework of the Strategic Vehicle Research and Innovation (FFI) is gratefully acknowledged. The authors also would like to thank for the supportive environment among all the working parties in the FFI project – Non-Destructive Testing Methods: Development of Innovative Solutions for In-line Applications.

REFERENCES


Impact Acoustic Testing as NDT Method for Classification of Compacted Graphite Iron

Lorenzo Daghini¹, Anders Berglund²

¹KTH-Royal Institute of Technology, Dept. of Production Engineering, Sweden
²Scania CV AB, Sweden
m2daglor@kth.se

ABSTRACT
Automotive industry is always struggling to comply with more and more restrictive emission regulations and the trend has been to employ engine materials that can allow higher combustion pressure and lighter design. Compacted Graphite Iron (CGI) is a class of materials that can allow this. However, the material and mechanical properties may largely vary within the given specification for CGI due to different solidification and cooling rates. To ensure that the component has the right mechanical properties it is therefore a need for a fast and reliable method for classify CGI according to its properties. Impact Acoustic Testing (IAT) method of inspection measures the structural response of a part. Its volumetric approach tests the whole part providing objective and quantitative results. The aim of this paper is to demonstrate that it is possible to distinguish the test objects’ mechanical and material properties provided that geometrical dimensions are not varying. The results show the potential and limitations of this method in the given application.

KEYWORDS: Impact Acoustic testing, Compacted Graphite Iron, Non-destructive testing

1. INTRODUCTION
Automotive industry is always struggling to comply with more and more restrictive emission regulations and the trend has been to employ engine materials that can allow higher combustion pressure and lighter design. Compacted Graphite Iron (CGI) is a class of materials that can allow this.

The basic procedure of producing CGI is to carefully monitor and control the amount of Magnesium (Mg) in the melt, as well as the solidification and cooling rate. The amount of Mg affects the graphite form, and therefore also the material physical properties of the cast component. If there is not sufficient magnesium, the graphite begins to grow with a flake morphology (resulting in grey iron) during solidification, which reduces the strength of the material drastically. Too high concentration of Mg, on the other hand, leads to nodular graphite which results in undesirable properties for cylinder block applications (low thermal conductivity and poor machinability). The magnesium content must be controlled simultaneously with the inoculation level in order to produce high quality CGI.
International Conference on Advanced Manufacturing Engineering and Technologies

microstructures. Postinoculation can suppress carbide formation, practically in thin walls but it provides more sites for graphite precipitation which favours the growth of spheroidal rather than compacted graphite particles. The process window for casting CGI is therefore narrow.

The quality of the produced part needs therefore to be controlled, to ensure the robustness of the process. One way to achieve this is by employing destructive testing methods, e.g., image analysis of microstructure. However, the current destructive testing methodology is time consuming and cannot be applied to all parts as the lead time would increase excessively. Therefore, fast and reliable Non-Destructive Testing (NDT) method, which is able to categorize the parts according to the microstructural properties, is needed.

This paper investigates the feasibility of implementing impact acoustic for classifying Compacted Graphite Iron (CGI) parts according to their material and mechanical properties. The aim is to give practical guidelines for possible applications of this NDT method. The study presented here is a preliminary investigation.

1.1. **Compacted Graphite Iron (CGI)**

The strict environmental legislation is compelling the automotive industry to produce engines capable of lower emission rates. Generally this can be achieved by reducing the weight of the engine or by increasing the combustion pressure in the engine. Hence, ordinary grey cast iron, found in many of today conventional diesel engines, is being replaced, due to the inferior mechanical properties.

CGI has twice the strength and 40% higher elastic modulus compared to grey iron [1]. The reason for this is the difference in microstructure, which is usually specified by nodularity and pearlite content but also by coarseness of the pearlite, defined by the interlamellar distance in pearlite. A shorter interlamellar distance gives a stronger material [2]. The CGI component microstructure is also strongly dependent on the cooling- and solidification rate during casting which, in its turn, is controlled by the section thickness [3].

CGI is a material family where the mechanical properties span over a large range. The microstructure determines the material physical properties which affects the machinability parameters [4].

1.2. **Impact acoustic testing (IAT)**

Impact Acoustic Testing (IAT) measures the structural response of a part and evaluates it against the statistical variation from a control set of good parts to screen defects [5]. Its volumetric approach tests the whole part, both for external and internal structural flaws or deviations, providing objective and quantitative results. This structural response is a unique and measurable signature, defined by a component’s mechanical resonances. These resonances are a function of part geometry and material properties and are the basis for RI techniques. By measuring the resonances of a part, one determines the structural characteristics of that part in a single test.

Typical flaws and defects adversely affecting the structural characteristics of a part are given in Table 1 for powdered metal, cast and forged applications. Many of the traditional NDT techniques can detect these flaws as well, but often only IAT can detect all in a single test, throughout the entire part (including deep sub-surface defects), in an automated and objective fashion.

IAT is basically experimental modal analysis simplified for application to high volume production manufacturing and quality control testing. This technique performs resonant inspection by impacting a part and “listening” to its acoustic spectral signature with a microphone. The controlled impact provides broadband input energy to excite the part and the microphone allows for a non-contact measurement of the structural response. The part’s
mechanical resonances amplify the broadband input energy at its specific natural frequencies, measured by the microphone above the background noise in the test environment.

Table 1. Typical structural defects detectable by IAT [5].

<table>
<thead>
<tr>
<th>Cast</th>
<th>Forged</th>
<th>Powdered Metal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cracks</td>
<td>Cracks</td>
<td>Cracks</td>
</tr>
<tr>
<td>Cold Shunts</td>
<td>Missed or double strikes</td>
<td>Chips</td>
</tr>
<tr>
<td>Porosity</td>
<td>Porosity</td>
<td>Voids</td>
</tr>
<tr>
<td>Hardness/density</td>
<td>Hardness</td>
<td>Hardness/density</td>
</tr>
<tr>
<td>Inclusions</td>
<td>Inclusions</td>
<td>Inclusions</td>
</tr>
<tr>
<td>Heat Treat</td>
<td>Heat Treat</td>
<td>Heat Treat</td>
</tr>
<tr>
<td>Compressive and residual stress</td>
<td>Quenching problems</td>
<td>Decarburization</td>
</tr>
<tr>
<td>Nodularity</td>
<td>Laps</td>
<td>Oxides</td>
</tr>
<tr>
<td>Gross dimensions</td>
<td>Gross dimensions</td>
<td>Gross dimensions</td>
</tr>
<tr>
<td>Raw material contaminants</td>
<td>Raw material contaminants</td>
<td>Raw material contaminants</td>
</tr>
<tr>
<td>Missed processes/operations</td>
<td>Missed processes/operations</td>
<td>Missed processes/operations</td>
</tr>
</tbody>
</table>

Erauw et al [6] successfully employ such technique for identifying defective ceramic components and conclude that the appearance of defects leads to a significant modification of the vibration response of the component with respect to its original acoustic signature. In the same paper the authors state that in case of axisymmetric parts it is possible to identify geometrical flaws even without a reference part, by only analysing the vibration response symmetry.

Hertlin and Schultze [7] describe this technique and propose its use for mass produced workpieces. In the same publication it is stated that by resonant analysis within the audible spectra (0-20 kHz) one can identify defects as cracks, structural faults, shrink holes and adhesion defects in cast parts. The authors also state that acoustic measurement technology is very sensitive; even minor changes in the oscillatory behaviour of mechanical structures can be detected by looking at the position of a frequency, the shift of one or more resonant peaks, the shift of all resonances, the distance between resonances, the splitting up of a resonant frequency, the decrease of amplitude and the time signal (decay time characteristics, effective value). Each of these effects - sometimes in combination - represents a specific property of the part. To find out which of these effects can be used for characteristics to classify good and bad parts is the main task in the engineering phase. The authors state that this technique is suitable for detection of cracks and casting defects of cylinder heads.

IAT is also mentioned in literature as NDT methodology for building inspections. Ito and Umoto [8] use this method for detection of cracks in concrete and succeed in characterizing the size of the crack. Liu et al [9] [10] find this methodology fast and reliable for the evaluation of correct bonding of tile walls, i.e. presence of voids in the bonding material.
2. METHOD AND EXPERIMENTAL SETUP

This investigation has been carried out in two main stages. At first, an experimental modal analysis of the test objects has been carried out in order to thoroughly evaluate the mode shapes. These have been employed in the second stage (IAT) to select the most suitable position for the microphone.

The material types and properties of the test specimen are summarised in Table 2. Repeated tensile- and hardness tests were performed on all CGI materials and the grey iron to map the material physical properties. The test specimens were prepared from different positions in the workpieces. The positions were carefully selected and they give together significant values for the material physical properties. All specimens were machined to the same dimensions and shape.

Table 2. Material and mechanical properties of the tested specimens.

<table>
<thead>
<tr>
<th>Material type</th>
<th>Nodularity</th>
<th>Pearlite content</th>
<th>Interlamellar distance in pearlite</th>
<th>E [GPa]</th>
<th>Yield strength [MPa]</th>
<th>ε [%]</th>
<th>UTS [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray Cast Iron</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>105¹</td>
<td>182</td>
<td>1.3</td>
<td>222</td>
</tr>
<tr>
<td>CGI 11</td>
<td>26</td>
<td>85</td>
<td>274</td>
<td>131</td>
<td>330</td>
<td>2.4</td>
<td>471</td>
</tr>
<tr>
<td>CGI 14</td>
<td>6</td>
<td>95</td>
<td>248</td>
<td>132</td>
<td>334</td>
<td>2.6</td>
<td>434</td>
</tr>
<tr>
<td>CGI 17</td>
<td>29</td>
<td>95</td>
<td>261</td>
<td>143</td>
<td>353</td>
<td>1.6</td>
<td>505</td>
</tr>
</tbody>
</table>

2.1. Experimental Modal Analysis (EMA)

The EMA has been carried out with impact hammer and three accelerometers (see Table 3 for details). The measurement was executed by roving the hammer through 40 impact points, hitting and collecting the data five times per point. A fixture was specially designed to support the specimen, which was resting on polymeric foam with largely lower stiffness than the specimen (see Fig. 1). Data was collected and processed using LMS TestLab® software and its dedicated sampling hardware. The EMA was carried out only for one direction.

Table 3. Sensing equipment employed for EMA.

<table>
<thead>
<tr>
<th>Sensor type</th>
<th>Brand</th>
<th>Model</th>
<th>Serial nr.</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Hammer</td>
<td>Ziegler</td>
<td>Ixys</td>
<td>9117</td>
<td>2.24 V/N</td>
</tr>
<tr>
<td>Accelerometer</td>
<td>Dytran</td>
<td>3225F</td>
<td>6283</td>
<td>10.3 mV/g</td>
</tr>
<tr>
<td>Accelerometer</td>
<td>Dytran</td>
<td>3225F</td>
<td>6284</td>
<td>9.6 mV/g</td>
</tr>
<tr>
<td>Accelerometer</td>
<td>Dytran</td>
<td>3225F</td>
<td>6287</td>
<td>10.3 mV/g</td>
</tr>
</tbody>
</table>

¹ Typical value for Grey Iron.
2.2. **Impact Acoustic Testing (IAT)**

IAT was carried out using the previously mentioned impact hammer and a microphone (Larson Davis 900B 3112) positioned as shown in Fig. 2. The impact point was directly opposite to the microphone. The specimen were excited only once per measurement. The procedure was repeated five times per specimen in a randomised manner to ensure the repeatability of the test results.

Data was collected and processed using LMS TestLab® software and its dedicated sampling hardware.
2.3. Sensitivity assessment

It is well known that the frequency response is strongly influenced by geometry. A variation in geometrical dimensions might affect the results, therefore a FEM parametrical study varying the thickness of the specimen was carried out. For instance, Fig. 3 shows that a variation of just 0.5 mm in thickness can cause a fluctuation of the first natural frequency of about 10 Hz. This is a behaviour that should be taken into account when carrying out IAT. In this study, the specimens were prepared within a tolerance width of 0.1 mm.

Fig. 3. Parametrical study on the specimen thickness influence on compliance around the first mode.
3. RESULTS

3.1. EMA
The scope of the EMA was solely to extract the mode shapes of the specimen in order to avoid positioning the microphone on a node when carrying out the IAT. If the microphone is placed on such a node the IAT might not be able to identify some mode shapes and the FRF would be incomplete. The EMA showed that the points at the extremities of the specimen were never nodes, see for instance Fig. 4. For this reason in the IAT the microphone was placed as shown in Fig. 2. In this way all modes can be identified by the IAT as well.

Fig. 4. Mode shape for the second natural frequency (1857 Hz) of specimen CGI 17.

3.2. IAT
The IAT could easily distinguish grey iron and the different CGI specimen as expected since the mechanical properties are largely different (see Fig. 5).

The differences in mechanical properties between the given CGI specimens are more subtle (see Table 2) and this is shown by the extracted FRF as well (see Fig. 6). Nevertheless, IAT was able to distinguish CGI 17 from CGI 11 and CGI 14. The two latter resulted undistinguishable from each other.

As previously mentioned, all specimens were tested five times each in order to control the repeatability of the method. Fig. 7 illustrates, for instance, the results obtained for the CGI 17 specimen. The resulting natural frequency lies within a range of approximately 2 Hz, which is also the resolution of the measurement equipment.
Fig. 5. IAT result. Frequency response of Grey cast iron (red) and CGI 17 (green).

Fig. 6. IAT result. Frequency response of CGI 11 (red), CGI 14 (blue) and CGI 17 (green).
4. DISCUSSION AND CONCLUSIONS

Machinability of CGI is strongly dependent on its material properties (especially pearlite content [4]). Material properties also affect CGI mechanical properties. Stiffness is obviously dependent on mechanical and geometrical properties and the frequency response of a given body is dependent on the body’s stiffness. Therefore, estimating the frequency response can be a powerful method for classification of mechanical and material properties.

The results obtained were repeatable and demonstrated that IAT was capable to distinguish grey cast iron from CGI and to classify the given CGI specimen according to their mechanical properties.

As previously mentioned, CGI is a family of materials and machinability varies greatly among these. Previous studies have been showing that tool life may vary between 30 to more than 180 minutes depending on the material and mechanical properties [3]. A classification method such as IAT might be therefore employed to optimize the cutting process parameters.

The advantage of IAT is that, once the equipment is set up, the actual impact test and data extraction does not take more than few seconds. On the other hand, the classification according to the mechanical properties is reliable only if the geometrical dimensions of the test objects do not excessively vary. In this particular case, a variation of 1 mm in thickness might have been enough to misrepresent the obtained classification.

5. FUTURE WORK

The specimens available for this research did not cover the whole range of possible pearlite content for CGI, and, as previously mentioned, this is a key factor for assessing machinability. Therefore it would be desirable experiment IAT on specimens with larger difference in pearlite content.

Most workpieces do not possess such regular geometry and their material and mechanical properties vary within the geometry, therefore further investigation on the possible industrial implementation of IAT for classification of CGI is needed.
ACKNOWLEDGEMENTS

The authors would acknowledge Scania CV AB for the given support. A special thank goes to Dr. Henrik Svensson for carrying out the material characterization at Swerea Swecast, as well as to Tomas Österlind and Zegeye Eshetu Tefera at KTH for helping performing the EMA and IAT experiments. This research has been funded by KTH DMMS, XPRES and the Swedish Governmental Agency for Innovation Systems (VINNOVA) through the FFI “NDT Methods” project (Diarienr. 2010-01346).

REFERENCES


Theme 9

Industrial Session

- Tolerance chain design and analysis of in-process workpiece
  Mats Bagge, Mikael Hedling, Bengt Lindberg
  1 Scania CV AB, Södertälje, Sweden
  2 KTH Royal Institute of Technology, Stockholm, Sweden.

- Nonparametric identification of stiffness and damping in nonlinear machining systems
  Andreas Archenti, Mihai Nicolescu
  KTH Royal Institute of Technology, Stockholm.

- Using concept modelling enables improvement system development
  Jens von Axelson
  Mälardalen University, Division of Production, Design and Logistics Research, Product Realization Group, Eskilstuna, Sweden

- The IDEAS Plug & Produce System
  Mauro Onori, Antonio Maffei, Friedrich Durand
  1 KTH Royal Institute of Technology, Dept. of Production Engineering, Stockholm, Sweden
  2 ELREST GmbH, Kirchheim, Germany

- Laser assisted milling of a nickel based alloy using different insert geometries
  Michael Kuhl, Martin Dix, Giacomo Leopardi, Flaviana Tagliaferri, Carlo Rüger
  1 Chemnitz University of Technology, Germany
  2 University of Naples Federico II, Italy

- Theoretical and Experimental Studies of Internal Turning Using Damped Boring Bars
  Ove Bayard, Lorenzo Daghini, Mihai Nicolescu
  KTH Royal Institute of Technology, Dept. of Production Engineering, Sweden

- Steps Towards Holistic Modelling of HSC Machining Centre Errors
  Jerzy Jedrzejewski, Wojciech Kwasny
  Wroclaw University of Technology, Poland
Tolerance chain design and analysis of in-process workpiece

Mats Bagge$^{1,2}$, Mikael Hedlind$^1$, Bengt Lindberg$^2$

$^1$ Scania CV AB, Södertälje, Sweden
$^2$ KTH Royal Institute of Technology, Stockholm, Sweden
mats.bagge@scania.com

ABSTRACT
Process planning comprises a broad range of activities to define a complete process chain for the manufacturing of a product. Beside the definition of proper operations and operation sequence, the definition of in-process tolerances for multi-step machining is decisive for the possibilities to achieve a well running and economical production. This paper introduces the dimension dependency chart (DDC) as a methodology based on tolerance charting technique where the complete manufacturing process chain can be efficiently represented, analysed and developed. The DDC uses mapping matrices for the transformation of operation element behaviour to the process steps defined by the process planner and further on to the dimensions and tolerances for the final part. Two important comparisons can be done by using the DDC; the expected behaviour of the in-process steps versus the tolerances defined by the process planner and the allowed variations of the final part dimensions versus the designed product specification.

KEYWORDS: Process planning, tolerance chain, tolerance chart, dimension dependency chart

1. INTRODUCTION

Process planning covers a wide range of activities needed to specify the manufacturing process for a product. Alting and Zhang [1] describe ten typical activities to be performed during process planning, but the work often includes coordination of product design intentions and constraints imposed by the workshop [2]. The goal is to design a manufacturing process that is capable to produce a product that fulfils all design requirements, without driving high economical costs. Xu, Yuan and Li [3] emphasise the major impact that process planning has on factory activities and the resulting manufacturing costs.

One of the most crucial tasks for a process planner is defining geometric in-process dimensions and tolerances for machining and other shape affecting processes. Process planning is not only needed when introducing completely new products or processes, but also to manage and improve running production regarding quality and cost. In this context, in-process tolerances play an important role.
Depending on designed product specification and manufacturing process alternatives, the definition and use of in-process tolerances is more or less easily handled. If a final dimension of the part is solely produced by one process step, without contributions or influences from any other operation, the definition of a tolerance is a quite uncomplicated task because it will directly correspond to the specification of the designed product. In a manufacturing process where two or more process steps are needed to create one dimension of the final part, each process step will contribute with variations, limited by in-process tolerances, which in turn forms a tolerance chain. The effect of this tolerance chain has to be taken care of before each in-process tolerance can be established and a final conclusion of the over-all capability of the process can be drawn.

In geometric dimensioning and tolerancing (GD&T) are the tolerance limits assumed to be identical to functional limits, as defined in ISO 8015:2011 “Geometrical product specifications (GPS) – Fundamentals – Concepts, principles and rules”. In product design the required functions, specified as tolerance limits, are mainly related to the product usage. In process planning are the required functions related to manufacturing and defined during in-process workpiece design. In design for manufacturing, these two functional aspects are interrelated but still with the need to be defined separate.

As an example, relatively large tolerances for the designed functions could require tighter tolerances for the functions used in manufacturing e.g. faces used for workpiece location in sequenced operations might require tighter manufacturing tolerances compared to the design tolerances. Also, for cases with a design tolerance, referencing datum or datum system not accessible or not yet manufactured, the tolerated feature requires a reference to intermediate datum or datum system, or possibly a completely different tolerance specification.

Utilising the GPS system for in-process workpiece specifications supports comparison between the two different functional aspects. With the different functional aspects specified in separate geometrical specifications, the manufactured part has to conform to both manufacturing tolerances and design tolerances. The manufacturing tolerances are set in process planning and can change depending on process plan solution. Manufacturing tolerances are equal or tighter compared to design tolerances. In process planning is one objective with manufacturing tolerances to defined in-process workpiece specifications that are as efficient as possible for operation control.

Systematic and random shape variations in manufacturing are added in each operation element. With manufacturing tolerance limits are the allowed variation specified. For manufacturing tolerances referencing datum or datum systems, or distances between shape elements produced in different operations, there is a need to decouple the tolerance relationship to enable assignment of variations (measured or expected) for each shape element. Analysis of tolerance chains is an important issue for the understanding of how the accumulated variations of all involved manufacturing steps affects the final part and relates to the design requirements and total cost. The total effort needed to produce a part is much about the accumulated economical cost from all process steps in the manufacturing chain. One important factor affecting this cost is tolerances, both for the in-process workpiece and the final part. The reason is that more demanding tolerances require higher performing processes, which in general are more costly and vice versa [4, 5]. If the process planner lacks of appropriate methods, tools or necessary skills to manage tolerancing, improvements tend to be more of a coincident than well considered. A conclusion after many years of process planning work is that the incentives to widen a tolerance that is too tight is often very obvious, because production will not accept to produce parts out of tolerance. The opposite situation, when a tolerance can be reached with just minor achievements, tends to be put into the shade of other problems; in spite of the fact that the potential performance of, for example, machine tools and
other resources is not fully utilised. This actual condition for the process planning work has not been described in studied literature.

One way to balance the process performance against the required tolerance is to choose a less performing and not so expensive machine tool. Often workshops have a profile for manufacturing of a certain range of products, in terms of machine tool types, engineering and operational skills; therefore changing this profile is only a long term solution, for instance investments in new machine tools or manufacturing knowledge.

Another way is to make more use of the machine tool in favour of other process steps in the manufacturing chain. This is realised by tuning the distribution of tolerances between process steps and requires that tolerance chains can be identified and efficiently evaluated. The potential of balancing tolerances is not only an issue for the relation between different process steps in different machine tools, but also between single operations in one process step.

This paper proposes a methodology based on tolerance charting technique to design and analyse tolerance chains for multi-step machining processes. The methodology is applied to axial dimensions of rotational parts.

2. METHODOLOGY CHARACTERISTICS

One important characteristic of the methodology is the ability to systematically define and analyse the relations between operation elements and process plan, and between process plan and designed product specification as shown in Fig. 1.

Fig. 1. Relations between operation elements, process plan and designed product specification.

“Process plan” is here defined as the description of how operation elements are organised into a manufacturing process chain with dimensions and tolerances connected to each process step. “Operation element” is here defined as a single machining operation of which the result can be separately evaluated. Typically this is turning or milling of one surface which can be measured and evaluated independently of other operations. The operation element design and selection can though be dependent on each other and it is possible that the behaviour of one operation element is dependent on another operation element.

The behaviour of the operation elements must be appropriate represented, transferred to, and used in the process plan. The behaviour determines how the operation element contributes
to different process steps and its capability to fulfil the in-process requirements. These in-process requirements are expressed as dimensions and tolerances in the process plan.

2.1. The tolerance chart technique

The developed methodology presented in this paper is based on the tolerance chart technique, which has been described and used in different variants by many researchers [5-8]. This is an approach which combines graphical and numerical representation of part dimensions and tolerances. A conceptual view of a tolerance chart developed and used in this paper is shown in Fig. 2.

![Tolerance Chart Conceptual Description](image)

The principle for the tolerance chart is that both the specification of the designed product and the in-process workpiece definition are represented in the same manner regarding dimensions and tolerances.

Several process steps are combined to form a process chain which is intended to produce a part according to the design specification. This process chain also entails tolerance chains which have to be analysed and evaluated regarding their effect on the final dimensions of the part. The mapping matrix is introduced in this paper as a representation of the process chain and used for transformation of in-process dimensions and tolerances to process chain dimensions and tolerances. The main objective with the tolerance chart is to manage the tolerance chains and calculate the expected outcome of the process chain, which in turn can be compared to the design specification.

A tolerance chart can be used in different process planning stages depending on the circumstances for the work. Either there is an existing process plan that shall be evaluated or a new process chain design to be accomplished. To get the full potential of the methodology the complete tolerance chart must be filled, but there is a possibility to use it partially.

2.2. The dimension dependency chart

The process planner defines and combines operation elements to a complete manufacturing process chain where the outcome is the finished part. How well the finished part corresponds with the design part depends on how the operation elements are utilised and put together in the process plan. Based on that, not only the relations between process plan and specification of the designed product, but process plan and operation elements are important and must be known.
An important characteristic of the tolerance chain design methodology is the ability to handle these relations. This is realised by introducing the dimension dependency chart (DDC) where the tolerance chart technique is used twice, but with slightly different purposes. In one part, operation element behaviour are defined and transferred to the in-process steps defined in the process plan. This part is later on referred to as the “lower view” of the DDC. In the other part, the “upper view”, the dimensions and tolerance chains are calculated by integrating involved in-process steps. The result represents the allowed outcome of the process chain and is compared to the product specification.

2.3. The importance of coupled operation elements

Possible dimensions and tolerances for each machining operation are depending on constraints and behaviour of the machine tool, but also the choice of cutting tools, fixtures and technology parameters like cutting data. The usage behaviour of the machine tool, cutting tools and fixtures are essential inputs to process planning because they are important contributors to process variations, which implies the use of both in-process and design tolerances. These variations cannot be completely eliminated but depending on type, handled in different ways.

If they are random and unpredictable, without known reasons, there must in general be tolerances wide enough to overcome these variations. If the variations are systematic and predictable, they may be minimised by, for example, automatic compensation for tool wear or temperature changes in the machine tool, or by a manual control strategy. However, there is a possibility to avoid the effects of systematic variations, and in a certain degree random variations, by smart process planning.

One example of smart process planning is when the same cutting tool is chosen for producing two surfaces in the same machine set-up, and the dimension to be evaluated is defined between these two. Here, the tool wear is a systematic variation that affects each surface, but the distance in-between will not change due to their tool related interconnection [9]. These two operation elements are coupled and form a resultant dimension.

By utilising couplings between operation elements in the process plan and sidestep some variations, tolerances can be kept tighter without risking producing parts out of specification. The possibility to define more capable processes in this way has therefore been integrated in the proposed methodology for design and analysis of in-process tolerances.

3. Establish a Dimension Dependency Chart

Machining of the shaft-like part shown in Fig. 3 is used as an example to illustrate how to establish the DDC. The part is machined in three set-ups with different fixtures and machining datums. Process steps 1-4 are performed in set-up 1, process step 5 in set-up 2 and process step 6 in set-up 3.

![Fig. 3. Designed product dimensions and machining set-ups used in the DDC example.](image-url)
Establishing of the DDC is here described by referring to different fields (A–K) and two views (upper and lower) shown in Fig. 4.

**Symbols:**
- Design dimension
- Machined feature and extent of dimension
- Finally machined feature
- Machining location and datum surface

### 3.1. Establishing the upper view of the DDC

The shape of the product is represented by a figure put in field A. All features of the product that will be included in the work are indicated, numbered and given an own column in the matrix shown as field B. The position and extent of every design dimension corresponding to the view in field A is also put into the matrix. Each row in the matrix represents one specific dimension.
Beside this matrix there are two columns in field C which contain the nominal dimensions (Dim) and their corresponding tolerances (Tol).

The same semantics for representation of dimensions and tolerances is used both for the designed product and the in-process workpiece. For the in-process workpiece, the dimensions and tolerances decided by the process planner and entered in field E corresponds to the features created in each process step which are illustrated in field D. An arrow head indicates what feature is machined and the arrow line how the dimension is defined. “X” indicates the feature used as a machining datum and location surface, and “F” indicates where a finished feature is created. This approach to schematically represent the dimensions for the in-process workpiece does only indicate the position of each feature in its final state. In-process dimensions may include for example extra stock left for fine turning or grinding, but this is defined by the nominal dimension values entered in field E.

All process steps contributing to a certain design dimension can now be identified, gathered and put into the tolerance chain mapping matrix in field F. This is done by finding the path of the process chain between the features that bounds the dimension. Figure 5 illustrates how the design dimension 50 mm, between feature 2 and 5, is examined.

![Fig. 5. Examination of one process chain.](image)

Process steps 6, 5 and 2 are involved in this process chain which is entered as directional in the first column of the mapping matrix. The entered values are either 1 or -1, depending of the direction of the path in each particular process step. The direction is positive when moving from left to right and vice versa. This convention is also applied on the direction of the design dimension to be examined. In this example, the path must be walked from feature 2 to feature 5, not from 5 to 2.

The mapping matrix in field F is used to calculate dimensions and tolerance chains for the complete process by elementary algebraic operations, later on shown in section 3.3.

The results are entered in field G.

**3.2. Establishing the lower view of the DDC**

The upper of the two DDC views, where the process steps are defined, gathered and mapped against the specification of the product design, is established according to sections 3.1. The next step is to establish the lower view which aims to connect operation elements to the process plan. Two important things must now be emphasised:

311
1. The field D, which illustrates the in-process steps, and the tolerance column of field E are identical in the upper and lower views of the chart.

2. In the upper view; dimensions and tolerances are mapped from the process plan onto the designed product specification. But in the lower view is the behaviour of the operation elements mapped upon the process plan.

As stated above, the fields D and E appear in both the upper and the lower view. Field H contains the specification of each operation element with the same notation as the process steps.

The behaviour of the operation elements can in general be expressed in different ways. Typically, process behaviour is quantified by statistical measures like the standard deviation or the range of the variation stated with a certain degree of confidence. It is also possible to define mathematical functions for this purpose. In this example, and further on in the paper, is the behaviour represented as an expected range of variation for each operation element. This range of variation is divided into two parts; systematic error (SE) and random error (RE). The approach to divide the variation in two components is comparable to the definition of machining errors by Rong [10] where “deterministic” and random machining errors are used.

This split is decisive of the possibility to benefit by coupled operation element behaviours which is further explained below. These systematic and random errors are put into the columns SE and RE respectively in field I.

Similar to the upper view of the DDC there is a mapping matrix (field J) which here defines the relations between operation elements and process steps in field D.

This mapping matrix contains information about which operation elements that contribute to the in-process steps and also how they contribute. If one, single operation element corresponds to an in-process step, “1” is entered in the matrix. If two operation elements without any coupling in-between are contributing, “1” is entered in the matrix for both operation elements.

However, as explained in section 2.3, it is sometimes possible to apply “smart process planning” to get use of the fact that the behaviour of two operation elements are coupled, e.g. when the same cutting edge creates two different features. If an in-process dimension is defined between these two features, the resulting systematic error (SE) for the dimension will ideally equal to zero [10]. To get this effect in the calculations, the coupling is represented in the mapping matrix by letting one of the operation elements get “-1” instead of “1”.

The mapping matrix is applied on the operation element behaviours to calculate the expected outcome of the process in field K. The leftmost column shows the total error for each in-process step, calculated as the sum of systematic and random errors.

The following section 3.3 explains the algebraic calculations of the complete DDC by defining equations (1) to (5).

3.3. Algebraic representation and calculation of the DDC

Definition of mapping matrices:
The mapping matrix in field F is defined as: \( \mathbf{MM}_{\text{In–proc.workp} \rightarrow \text{Proc.chain}} \)

The mapping matrix in field J is defined as: \( \mathbf{MM}_{\text{Op.element} \rightarrow \text{In–proc.workp}} \).

Definition of vectors:
Columns in the DDC containing tolerances, dimensions, systematic errors, random errors or total errors are defined as column vectors. Vectors used in the DDC calculations are shown in Table 1.
Table 1 Vectors used in the DDC calculations

<table>
<thead>
<tr>
<th>Input vectors</th>
<th>Output vectors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Field E</strong></td>
<td><strong>Field G</strong></td>
</tr>
<tr>
<td><em>Dim</em>&lt;sub&gt;ln-proc.workp&lt;/sub&gt;</td>
<td><em>Dim</em>&lt;sub&gt;Proc.chain&lt;/sub&gt;</td>
</tr>
<tr>
<td><em>Tol</em>&lt;sub&gt;ln-proc.workp&lt;/sub&gt;</td>
<td><em>Tol</em>&lt;sub&gt;Proc.chain&lt;/sub&gt;</td>
</tr>
<tr>
<td><strong>Field I</strong></td>
<td><strong>Field K</strong></td>
</tr>
<tr>
<td><em>SE</em>&lt;sub&gt;Op.element&lt;/sub&gt;</td>
<td><em>SE</em>&lt;sub&gt;ln-proc.workp&lt;/sub&gt;</td>
</tr>
<tr>
<td><em>RE</em>&lt;sub&gt;Op.element&lt;/sub&gt;</td>
<td><em>RE</em>&lt;sub&gt;ln-proc.workp&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

**Definition of equations:**

\[ \text{Dim}_{\text{ln-proc.workp}}^T \cdot \text{MM}_{\text{ln-proc.workp} \rightarrow \text{Proc.chain}} = \text{Dim}_{\text{Proc.chain}}^T \] (1)

\[ \text{Tol}_{\text{ln-proc.workp}}^T \cdot \text{abs} \left( \text{MM}_{\text{ln-proc.workp} \rightarrow \text{Proc.chain}} \right) = \text{Tol}_{\text{Proc.chain}}^T \] (2)

\[ \text{SE}_{\text{Op.element}}^T \cdot \text{MM}_{\text{Op.element} \rightarrow \text{ln-proc.workp}} = \text{SE}_{\text{ln-proc.workp}}^T \] (3)

\[ \text{RE}_{\text{Op.element}}^T \cdot \text{abs} \left( \text{MM}_{\text{Op.element} \rightarrow \text{ln-proc.workp}} \right) = \text{RE}_{\text{ln-proc.workp}}^T \] (4)

\[ \text{SE}_{\text{ln-proc.workp}} + \text{RE}_{\text{ln-proc.workp}} = \text{TE}_{\text{ln-proc.workp}} \] (5)

In this paper, all tolerance chain calculations are simply made as “worst-case”. This means that the range of each in-process tolerance is added and the sum represents the tolerance-range for the complete process chain. These calculations are made according to equation (2). The same approach is used for the random errors, where the ranges of variation are summarised according to equation (4).

**3.4. Evaluation of the calculated results**

There are two comparisons to be done when the complete DDC is calculated. One is the comparison between the process chain dimensions and tolerances, representing the allowed outcome of the manufacturing process, and the specification of the designed product where all requirements must be met. Another comparison is done between the expected outcome of each in-process step and the tolerance defined by the process planner. If the variation range of the outcome exceeds the chosen tolerance; the tolerance, operation elements or the definition of the in-process step must be changed.

**4. DISCUSSION**

In this paper the worst-case approach is used for tolerances and operation element behaviour. However, many other tolerance chain calculation methods can be used, for example statistical model [4], constant factor model or estimated mean shift model [11].

The presented DDC methodology does require some model, but not particularly the worst-case or any one of the other models, to be capable to calculate tolerance chains. Instead, it provides the possibility to incorporate and evaluate different kinds of calculation methods depending on conditions, like available input data.
There are various reasons for operation elements to be defined as coupled. As described in section 2.3, one is when the same cutting edge on a tool is used for machining of two features. Another is when the strategy for measuring and controlling the process, or tool-change intervals, determines a relation and the operation elements behaviour will be coupled.

In multi-step machining processes, containing for example soft machining, heat treatment and hard machining, the aspect of material flow has to be considered regarding couplings between operation elements. Even if there is a first-in first-out flow through the first process, for example a machining line, the workpieces may be mixed during transport or in the next process. The logistic system has by this reason an impact of relations between operation elements.

By this background, the commonly used machine and process capability tests can be considered. Machine capability tests basically aim to evaluate the performance of the machine tool executing one operation. Process capability tests include not only one operation, and the measuring of this operation, but many process steps.

The tests are carried out by putting the measured variation of a number of produced parts in relation to the allowed tolerance for the operation element, the in-process step or the process chain. The variation is commonly expressed as standard deviation, and the tolerance as the range between upper and lower tolerance limits.

In parallel to the definitions in this paper, the machine capability test represents the evaluation of one operation element. The process capability test represents the evaluation of in-process steps forming a manufacturing process chain.

One desired goal for the process planner should be prediction of the outcome of the complete manufacturing chain by using information about, for example, operation element behaviour, in-process steps, process chain, process control strategies and logistics. By using only results from capability tests based on standard deviation, this is not possible.

The DDC is usable both as a tool for process planners and a workbench for further research. The presented methodology is after some additional development expected to be suitable for the handling of time- and sequence-dependent aspects of the operation element behaviour. Future work is proposed to include that development and also evaluation of the impact of process control strategies, statistical and analytical definitions of the operation elements and logistic aspects. Many examples of tolerance charting include the calculation of stock removal allowances. The DDC does not deal with this in its present form in favour of the clarity of the methodology. However, stock removal allowances will have a potential impact on the behaviour of the operation elements and needs attention in coming development stages of the DDC.

5. CONCLUSIONS

This paper introduces a methodology for tolerance chain design and analysis of an in-process workpiece. The methodology is based on tolerance chart technique, but with a novel approach to include and connect operation element behaviour to the process plan and evaluate the expected outcome of the process chain. This is realised by introducing the dimension dependency chart (DDC).

To facilitate handling of coupled operation elements, and evaluate their influences on the process chain, the behaviour of these elements are defined in two parts; Systematic error and random error. This makes it possible to quantify and integrate “smart process planning” in the process plan and enables analytical calculation and evaluation of the results.
REFERENCES

Nonparametric identification of stiffness and damping in nonlinear machining systems

Andreas Archenti\textsuperscript{1}, Mihai Nicolescu\textsuperscript{1}

\textsuperscript{1}KTH Royal Institute of Technology, Stockholm
archenti@kth.se

ABSTRACT
The demand for enhanced performance of production systems in terms of quality, cost and reliability is ever increasing while, at the same time, there is a demand for shorter design cycles, longer operating life, minimisation of inspection and maintenance needs. Experimental testing and system identification in operational conditions still represent an important technique for monitoring, control and optimization. The term identification refers in the present paper to the extraction of information from experimental data and is used to estimate operational dynamic parameters for machining system. Such approach opens up the possibility of monitoring the dynamics of machining system during operational conditions, and to be used for control and/or predictive purposes.

KEYWORDS: machining systems, self-excited vibrations, operational dynamic parameters

1. INTRODUCTION
The demand for enhanced performance of production systems in terms of quality, cost and reliability is ever increasing while, at the same time, there is a necessity for shorter design cycles, longer operating life, minimisation of inspection and maintenance needs. By help of advanced computing systems it has become less expensive both in terms of cost and time to perform numerical simulations, than to run time and material consuming experiments. The consequence has been a shift toward computer-aided design and numerical trials, where virtual models are employed to simulate experiments, and to perform accurate and reliable predictions of system behaviour.

Even if the technology of virtual prototyping is steadily growing in manufacturing environment, experimental testing and system identification still play a key role because they provide valuable information to production engineers concerning the influence of process modification on system’s performance, the prediction of failures in the system and support for the maintenance of the production systems.

In the classical machining theory and practice, a large body of research has been dedicated to study the machining system’s (MS) dynamics and it originated from the early works by Tlusty [1] and Tobias [2]. At a closer examination there are some phenomenological
and technical shortcomings in the classical methodology. One critical issue is the parameters describing dynamical behaviour of MS are extracted independently for the structure and for the process before connecting them together to study system’s stability (see Fig. 1) [3].

Fig. 1. Analysis of the dynamic behaviour of the MS: (left), extraction of dynamic parameters of the machine tool (elastic structure), (middle) stability charts, (right) chatter marks as a result of dynamically unstable machining.

MS represents the interaction between the elastic structure and the chip formation process. This interaction is controlled at interface tool – chip – workpiece by a tiny elasto-plastic material volume. Separate extraction of modal parameters using off-operational experiments, e.g. experimental modal analysis (EMA), and controlled experiments for extraction of cutting process coefficients is not an efficient way to characterize the MS in operational conditions. From theoretical point of view, using the direct analysis of the elastic structure and cutting process based on the fundamental physical laws may be insufficient for accurate description of the behaviour of the system in real operational conditions. Regarding the cutting process, estimation of e.g. cutting force coefficients can be done only with barely simplified hypotheses, for few particular operations, and only in laboratory controlled conditions.

In order to control and predict dynamic behaviour of MSs it is important to estimate the actual operation dynamic parameters (ODP) in machining of actual parts, in actual cutting conditions and configurations [4].

The MS is excited by loads which are unmeasurable and unpredictable. In addition, the dynamics of the system is changing continuously during machining due to variation of the load magnitude, orientation and position. Consequently, the dynamic behaviour of a MS is not the same in open loop i.e. in off-operational conditions, and in closed loop i.e. in operational conditions. In this paper, the excitation forces are considered as stochastic processes. Lightly damped structures of machine tools have a response in a narrow frequency range and therefore the excitation can be at least at the first approximation considered as an ideal white noise. When the dynamic response of MS is analysed, linear theory is often used. For linear MSs, a large body of theory has been developed for the identification and modelling of system parameters in various dynamic configurations [5]. One of the main concerns in these theories is the study of self-excited vibration and the related stability analysis. For non-linear MSs with broad band excitation, closed form solutions are not available. In this paper, the main objective is to investigate the response of the non-linear MS and to develop a procedure for dynamic parameter identification.

The main contributions of this paper are: (1) development of non-parametric models based on identification technique with the purpose to integrate in a single step the estimation of dynamic parameters characterizing the MS, (2) in nonparametric identification, implementing techniques for ODPs and random excitation estimation. By this, a step is taken beyond the classical way to analyse the dynamics of a MS by separately identifying the
structural and process parameters. With the process considered, the two substructures, tool/toolholder and workpiece/fixture, are coupled, in addition to the open loop (elastic structure), by a feedback loop closing the energy loop, through the thermoplastic chip formation mechanism. The MS can be completely analysed only in closed loop i.e. in operational conditions, since specially designed off-line experiments with controlled input, such as modal testing, give the response from only the open loop.

1.1. Non-linear machining systems

In the problem involving the dynamic behaviour of MSs, one major source of uncertainty is the excitation. The chip-formation process takes place through the intricate closed-loop interaction between the tool/toolholder and the workpiece/fixture with the machine tool structure in the primary loop (open-loop) and the chip formation process in the feedback. The static and dynamic behaviour of a MSs is governed by the mass, stiffness and damping. In a non-linear system, damping and stiffness characteristics depend on the energy levels of the excitation [6, 7]. The excitation of the MS is mainly created by the cutting force. On the cutting force, other loads, of thermal or mechanical nature, are superimposed to generate the complex system’s response. The nature of these forces is non-linear conservative and dissipative. Non-linear conservative forces are restoring forces that arise from sources such as gravitational field and internal stresses generated in deformed structural elements. Apart from the cutting force, during cutting process there are dynamic forces generated due to dynamic unbalance of rotating parts, inertia forces of reciprocating movable elements, bearing irregularities, and geometric imperfections in structural elements.

The machine tool’s elastic structure consists in a relatively large number of structural elements. These elements are connected to each other by fixed and movable joints. Consequently, the elastic structure can be represented by a model with lumped masses connected by damping and ‘springs-like’ elements (see Fig. 2) [8].

Fig. 2. Machine tool with represented joints (stiffness and dampers) and an elastic element representing the cutting process.

In the classical machining dynamics theory, regenerative chatter is considered the major cause for instability [1]. Thus, the dynamic behaviour of the system is characterized through the stability charts which graphically present the stable and unstable lobes in a plane described
by the depth of cut and spindle’s rotational speed. The analysis method is usually implemented in three steps: (1) the modal parameters of the elastic structure are computed. This step is performed either experimental, e.g. experimental modal analysis (EMA) or the corresponding modal parameters are obtained from FEM models. (2) the process stiffness and damping are estimated from cutting experiments and (3) stability diagrams are computed. The dynamic stability of MS depends on joint stiffness and damping of the elastic structure and process. Regarding the elastic structure, for linear systems, the stiffness is obtained without difficulties from the material and geometric properties. If the structure is non-linear the stiffness is not easily evaluated. The damping is normally not possible to analytically compute even in linear case. Concerning process stiffness and damping, their accurate estimation is difficult since they depend of the operational conditions and the estimation is possible only for very simple operations. Complications arise also due to the fact that for estimating the dynamic stiffness coefficient for one cutting variable the other variables must be maintained constant which normally is not trivial [3].

2. ESTIMATION OF MACHINING SYSTEM’S DYNAMIC PARAMETERS

The approach presented in this paper has the purpose to identify the operational dynamic parameters (ODPs) of MS i.e., the equivalent stiffness and damping existing in a particular system configuration and at particular time moment. Therefore, the approach introduces a probabilistic concept where nonparametric identification models are employed. The MS is considered inherently non-linear. The contribution of the process stiffness and damping gives a new dimension to the system’s nonlinearity when the MS is considered as an entity.

MS is subjected to complex loads and it changes its configuration continuously as the tool moves along the workpiece or discretely as the system is reconfigured for various cutting operations. As a consequence both the excitation and the parameters characterizing the system may be considered random. An adequate description of excitation, and therefore of the response of the system to such loads has to be developed within the framework of the statistical dynamics theory [9]. Embracing such a probabilistic point of view implies that some statistical characteristics of both the stochastic excitation and the systems response to this excitation have to be considered. In dynamic systems with random excitation and involving the interaction of coupled structures through a fluid or through other media, self-excited phenomena are frequently present [10].

The problem of discrimination between random forced vibration and self-excited oscillations is a key issue for the evaluation of the system stability boundary during operational conditions. The response of a dynamic system to a broadband random excitation will be a nonzero steady-state signal both in stable and in unstable state. In case of self-excited vibration, the excitation persists even in the absence of the random excitation. The discrimination between self-excited and forced vibration of a dynamical system in operational conditions is also important for selecting the type and strategy of control that may be implemented to reduce or cancel the vibration.

The field of application of statistical dynamics which is the focus in this paper is related to two concepts: (i) identification and (ii) discrimination of the response of dynamic systems. The term identification refers to the formulation of a mathematical model of the dynamic system based upon on-line signal measurements, and belongs to a class of inverse dynamic problems encountered in various technological fields [11]. The discrimination term aims to determine from the identified model a function to characterize the nature of the system’s
response, i.e., whether $x(t)$ represents a forced vibration response or it is a self-excited vibration.

The problem of interest in MS dynamics is the discrimination between forced and self-oscillations in view of the following considerations:

- Formulation of a qualitative/semi-qualitative mathematical model of the MS for subsequent quantitative analysis,
- Evaluation of the system’s stability boundary,
- Implementation of a suitable design for chatter control.

The term “qualitative” implies that the model is based on a statistical analysis of measured system’s response. Although the model-based identification approach presented in this paper leads to the estimation of key dynamic parameters (hereby the term “semi-qualitative”), they are nevertheless meaningful only within a certain confidence interval.

2.1. Nonparametric identification of stiffness and damping

The nonparametric identification method presented in this paper is implemented in three stages for generation of the response of MS using numerical simulation. At the first stage, a Gaussian white noise process is created. At second stage, the equation of motion is integrated with a suitably chosen time step, using a Runge-Kutta algorithm of $4^{th}$ order which enables accurate response histories to be obtained. Finally, the response data are processed appropriately to estimate the system’s parameter.

As already mentioned, the excitation of the MS is considered as an external zero mean white-noise process with the covariance function

$$E[W_0(t)W_0(t + \tau)] = 2\pi S_0 \delta(\tau)$$

where $E[\sigma]$ is the ensemble average, $S_0$ is the intensity of white noise and $\delta$ is the Dirac’s delta function. White noise process, labelled $W_0(t)$, consists of a train of Dirac’s delta impulses at $\Delta t$ time increment. The pulses are linearly interpolated by the algorithm during numerical integration. The power spectral density (PSD) of the white noise process is given by [6]

$$S_w(\omega) = S_0 \left[\sin\left(\frac{1}{2} \omega \Delta t\right)\right]^4, \quad S_0 = \sigma^2_w \frac{\Delta t}{2\pi}$$

The equation of motion of the MS is described as a non-linear stochastic differential equation with additive excitation

$$\ddot{X} + h(E) + u(X) = W_0(t)$$

where $hI$ and $u(X)$ are non-linear damping and restoring forces of the MS. It is assumed that the state space variables $X$ and $\dot{X}$ enter equation only through the total energy function $E$, then the damping is a function of the energy only. The total energy is given by
\[ E = \frac{1}{2} \ddot{X}^2 + U(X) \]  

(4)

where

\[ U(X) = \int_0^X u(\xi) d\xi \]  

(5)

In Eq. (4) the right hand represents the sum of the kinetic energy (first term) and the potential energy (second term). With a change of variables, \( X_1 = X \) and \( X_2 = \dot{X} \), Eq. (3) is transformed into a set of two stochastic differential equations. Under the assumption of Gaussian white noise excitation, the state space vector \((X_1, X_2)\) represents a Markov process and the probability density is the solution of Fokker-Plank equation [12]

\[ x_2 \frac{\partial p}{\partial x_1} - \frac{\partial p}{\partial x_2} \{[h(E) x_2 - u(x_1)] p\} - \pi S_0 \frac{\partial^2 p}{\partial x^2_2} = 0 \]  

(6)

\( p(x_1, x_2; t) \) is the joint probability distribution of the state space vector \((x_1, x_2)\). The initial conditions are given in the form: \( p(x_1, x_2; t_0) = \delta(x_1 - x_{10}) \delta(x_2 - x_{20}) \) for \( x_1(t_0) = 0 \) and \( x_2(t_0) = x_{20} \). Different boundary conditions are possible such as reflective boundary, absorbing boundary and periodic boundary.

The solution to the Eq. (6) was obtained in the form [13]

\[ p(x_1, x_2) = C \exp \left[ -\frac{1}{\pi S_0} \int_0^E h(\zeta) d\zeta \right] \]  

(7)

where \( C \) is a normalizing constant. The integral in the above expression

\[ H(E) = \int_0^E h(\zeta) d\zeta \]  

(8)

is named damping potential. Equation (8) shows that the damping function \( hI \) can be obtained from the derivative of the damping potential \( HI \). Damping potential is computed from Eq. (6) after the joint probability \( p(x_1, x_2) \) is evaluated. Because, the probability density distribution of the energy envelope process, \( E(t) \), can be estimated from system’s response, it is required to relate the distribution \( pI \) to the distribution \( p(x_1, x_2) \).

3. IDENTIFICATION OF MACHINING SYSTEM ODPs

The procedure developed in section 2 will be applied for the identification of the join characteristics of the MS. Steel bars (\( C < 0.20\% \)) with length of 1200 mm and initial diameter 42 mm were machined in longitudinal turning between tail and chuck at a cutting speed \( v_c = 180 \text{ m/min} \), feed rate \( f_r = 0.3 \text{ mm/rev} \) and variable depth of cut (0.5 – 3mm). Cemented carbide inserts with 1.2 mm nose radius were used in all experiments. The vibration signals were measured by a pre-polarized microphone and sample at 12.8 kHz sampling rate. In Fig. 3 an
example of stable machining is presented, while Fig. 4 illustrates an unstable machining. Coming down to diameters below 34 mm, as the tool approach centre of the bar, chatter vibration is generated.

Fig. 3. Stable machining (D = 38 mm).

Fig. 4. Unstable machining: chatter marks on the machined surface.

Fig. 5. Unstable process, as the tool approach the centre of the bar heavy vibration arises.
Figure 5 presents the time signal showing first a stable process then the chatter in the middle of the bar, then as the tool approaches the chuck, the system is recovered stability. In Fig. 6, the unstable and stable signals are represented in frequency domain.

![Figure 6](image)

Fig. 6. The stable and unstable signals in frequency domain (1024 samples).

### 3.1. Operational stiffness identification

Following the procedure described in section 2.1, the join stiffness of the MS is estimated. For a stable process, Fig. 7 illustrates the potential function $U(x)$. The function shows a soft non-linear characteristic. After fitting to a quadratic polynomial function, the stiffness function can be extracted from Eq. (9). The blue line represents the theoretical linear system at the system’s natural frequency.

![Figure 7](image)

Fig. 7. Stiffness estimation – stable machining.

The estimation of the join stiffness of the MS entering unstable behaviour is presented in Fig. 8. The potential function $U(x)$ and consequently the stiffness function show a hard non-linear behaviour which is characteristic for self-excited vibration. The deviation from a non-linear behaviour is apparent by comparing to the linear characteristic shown by the blue dotted line. As these results show, if the stiffness of a MS enters the inelastic range of the material, or...
the process becomes non-linear, or a combination of both and the degree-of-nonlinearity are large, using a linearization method for the non-linear stiffness may yield large errors in response estimation.

![Fig. 8. Stiffness estimation – unstable machining.](image)

**3.2. Operational damping identification**

Following the approach described in section 2.1, the damping potential and equivalent damping are estimated from the system’s response both in stable and unstable condition. However, the results are presented only for unstable MS.

![Fig. 9. Estimation and fitting of damping potential for unstable machining](image)

The estimation procedure starts with computation of the damping potential $H(E)$ from the experimental computation of probability density function at different levels of energy. From Eq. (7) and Eq. (8), the function $H(E)$ is computed and then fitted to a polynomial. As in the case of stiffness, the damping in unstable MS shows a non-linear behaviour and a cubic polynomial is then employed. In the second step, the covariance function is calculated for gradually increasing level of energy. Some of these functions are represented in Fig. 10. It is
worth to notice that, as the stiffness is non-linear, the natural period depends on the energy level. The same can be stated for equivalent damping (see Fig. 11). The total energy is calculated as a sum of the kinetic and potential energies. The kinetic energy is extracted from the derivative of the response while potential energy from a iterative procedure applied to the identified potential function $U(x)$.

![Covariance function calculated for various energy levels.](image)

Fig. 10. Covariance function calculated for various energy levels.

![Computation of equivalent damping function.](image)

Fig. 11. Computation of equivalent damping function.

### 4. CONCLUSIONS

The MS can be completely analysed only in closed loop i.e. in operational conditions, since specially designed off-line experiments with controlled input, such as modal testing, give the response from only the open loop.

The paper describes the application of nonparametric models for ODPs identification. The presented method is used in conditions where input excitation is unknown which is the case of MS. The nonparametric identification approach is best suited in systems with low damping. Nonparametric identification technique gives possibility to estimate the departure from linearity of system parameters. The experimental results show that close to chatter the
nonlinearity of the system increases. Nonparametric identification technique gives also possibility to estimate ODPs at different energy levels as nonlinearity of the system requires.

Identification procedure in case of nonparametric technique follows three steps: (1) potential energy estimation, and fitting to a polynomial. Then the stiffness function is calculated from the derivative of the polynomial. (2) the damping potential is calculated from the response probability density function. (3) the equivalent damping is calculated for different levels of energy and fitted to a polynomial. From the polynomial and with knowledge of potential function the operational damping can be calculated.

A longitudinal turning operation is used to demonstrate the capability to estimate ODPs both in stable and unstable systems.

ACKNOWLEDGEMENTS

This work was funded by DMMS/KTH (Centre for design and management of manufacturing systems) and VINNOVA (The Swedish governmental agency for innovation systems), and has been supported by XPRES (Swedish initiative for excellence in production research).

REFERENCES

Using concept modelling enables improvement system development
Jens von Axelson

Mälardalen University, Division of Production, Design and Logistics Research, Product Realization Group, Eskilstuna, Sweden
jens.von.axelson@gmail.com

ABSTRACT
Many enterprises adapt new approaches, methods and tools in their improvement system as a response to the increasing global competition. A problem is that terms and concepts could mean different things to different system development team members. Concept models could serve here as shared mental models. These facilitate understanding, development and later implementation of a new future state of an improvement system. The objective has been to explore the concept modelling method in an operations management context. Two different cases on how to use concept models in research and development activities are presented. It is concluded that concept modelling is a fruitful support method in an improvement system development context.

KEYWORDS: concept model, shared mental model, improvement system development

1. INTRODUCTION

Many Swedish manufacturing companies are operating today in a global market environment and are under heavy competition. In order to stay in the market there must be a continuing effort in improvement to even maintain the ability to be competitive. Thus, the severe challenge requires companies to continuously develop their production systems [1].

One key element the last decades for enhancing the competitive edge of the production systems has been to adopt different improvement initiatives. Several different approaches, concepts, methods and tools, have been developed and adopted by companies such as Total Quality Management (TQM), World Class Manufacturing (WCM), Management systems (e.g. ISO 9001, TS 16949 and ISO 14001), Lean Production, Business Process Reengineering (BPR) and Six Sigma. One thing with all these approaches have in common is that they provide a component in the company’s production system improvement initiatives.

However, with incremental development over many years the number of concepts, methods and tools could be numerous. As a metaphor there is always a risk to build a Frankenstein’s monster with good part, but and in total an un-coordinated and complex system of improvement initiatives.
The current “toolbox” in operations management and in literature is continuously increasing with new approaches or new applications of old ones. One approach currently in focus is “Kaikaku\(^1\) which has more than ten different definitions \(\text{[2]}\). The amount of definitions of one term in this area can be extensive. Hence, a problem is that terms and concepts mean different things to different people and in different contexts. Especially in contexts where people from different backgrounds work in multi-disciplinary R&D project teams the differences in interpretations of terms and concepts often are present.

In industrial development processes it is never far to an example where researchers or practitioners have experienced difficulties due to the use of concepts and terms with others. In one example it took the project group nine months to understand the concept of the word ‘improvement’ \(\text{[3]}\). The method development process was disrupted of course due to this conflict of understanding. One way to illustrate this is presented in Fig. 1.

\[\begin{array}{|c|c|c|}
\hline
\text{Terms} & \text{Same} & \text{Different} \\
\hline
\text{Concepts} & \\
\hline
\text{Same} & \text{Consensus} & \text{Correspondence} \\
\hline
\text{Different} & \text{Conflict} & \text{Contrast} \\
\hline
\end{array}\]

Fig. 1. There are different outcomes when using same or different terms and concepts in a group \(\text{[4]}\).

The project members in the above example were definitely in the conflict quadrant according to Fig. 1 while they wanted to be in the consensus quadrant. Here it was obvious that it was of great importance to have definitive control of the words and their connection to the system that was about to be changed. There is often linguistic confusion in using terms and concepts due to synonyms (different words that have same meanings) and homonyms (same words that have different meanings) \(\text{[4]}\). Therefore, the need to define terms and concepts and to clarify connections and rules between different words demands insights, visions and ability to make decisions by those who work with the concept models \(\text{[5]}\).

In the project, “Kaikaku – Innovative Production Development” the research team focused on innovative production development as a way to realize the production strategy. The project was working with different research areas such as Production Development, Spatial Design, Innovation and Creativity. Several companies and research organizations were involved in the project such as Mälardalen University, Swerea IVF, Deva Mecaneyes and Volvo Construction Equipment – a multi-disciplinary group. At a comprehensive level, the expected results could be interpreted as identifying why and when a Kaikaku is needed as well as how to realize a Kaikaku. Furthermore, the project investigated how to be more innovative and creative in the production development process as well as how to integrate Kaizen\(^2\) and Kaikaku within a lean transformation.

One objective in the Kaikaku-project was to investigate how to realize a holistic way of working with improvements, i.e. in this case – how to implement a Kaikaku method in an

\(^1\) Approximately radical improvement
\(^2\) Approximately continuous improvement
already existing system of improvement initiatives. The need for different tools of system visualization and representation was obvious.

Concept modelling has been used as a development method for some years by Production Engineering researchers at the Royal Institute of Technology in several projects, e.g. for describing a machining centre for information modelling for the manufacturing life cycle [6]. Another example was building knowledge and secure communication in information management for factory planning and design [7].

The method development process itself and its design could facilitate not only better developed methods but also give the method a jump start in the innovation diffusion process [8]. One key issue was that the scope and purpose is clear of any new concept, method or tool. This is very important for the following development process and later dissemination [9].

Where there is a clear lack of method support in the improvement system, e.g., if there is no improvement work process for reducing throughput time, Value Stream Mapping could be chosen to be implemented [10]. When there already are several programs, methods and tools it is more complex. Concept models could be one building block in order to design and later spread a new improvement system design within the enterprise.

Therefore, in this case, the desired purpose of a conceptual modelling method is to model states, concepts and terms related to an enterprise improvement work in order to improve the improvement system design process itself. Thus the objective of this explorative study has been to investigate concept models as a means for shared mental models in an operations management development context. The research question has been: Is concept modelling a fruitful tool in an improvement system design and development context?

2. ASPECTS OF CONCEPT MODELING

The context of this study has been to realize and enhance the improvement work in a brownfield production system. Hence a multidisciplinary approach and theory base was needed. Different aspects of modelling have been reviewed with respect to change management, innovation diffusion and operations management and strategy.

2.1. The role of a model

Anyone who plays a strategy game knows that you need a plan. A bad plan is better than not having one at all. During the game things happen that were not predictable in advance. The plan now helps one make decisions while moving in the direction of the goal. A more skilled player could change the plan during the game in order to find better countermeasures to the challenges than the opponent. If the game is translated into a rather complex enterprise an ideal model of the enterprise would consist of a single diagram that includes the important aspects of the enterprise. That is often impossible but an enterprise model could be composed of 1) views, i.e. an abstraction from a specific viewpoint, 2) diagrams showing specific parts of the structures, and 3) objects and processes [11]. The model could then serve as a simplified reality which one could use for getting different answers.

2.2. Concept model as a shared mental model

A concept model could be used as a representation of a mental model. The definition of a mental model is that is an explanation of someone's thought process about how something works in the real world. It is a representation of the surrounding world, the relationships between its various parts and a person's intuitive perception about his or her own acts and their
consequences. Mental models can help shape behaviour and set an approach to solving problems and doing tasks [12].

Team performance could be enhanced by providing teams with sufficient information to build a shared mental model of each other’s tasks and goals. It is believed that the shared display helps building shared mental models which boost later task performance [13]. Furthermore, organizations that use learning procedures intensively share a greater amount of knowledge, and the congruence among their members’ mental models is higher than that of organizations that utilize the same mechanisms but less intensively. Such organizations are presumed to be more effective [14]. A shared mental model among team members, allows them to utilize efficient communication strategies during high-workload conditions, and results in improved coordinated team performance [15].

It is clear that a concept model has an interesting role to play as a communication link between reality and, in many cases, complex systems. Enterprises undergo continuous changes and as a consequence the support systems and methods also need to change.

2.3. Concept modelling, different notation system for different purposes

There are several official “languages” for concept or conceptual modelling. UML – Unified Modelling Language – has been developed for software development purposes. Concept models based on UML have been used not only for software development but also for enterprise modelling [11] and for configuration of mass-customizable products [16].

The IDEF are a group of modelling languages, which cover a wide range of uses, from functional modelling to data, simulation, object-oriented analysis/design and knowledge acquisition. The IDEF methods have been defined up to IDEF14. For instance IDEF0 support Function Modelling, IDEF1 Information Modelling, IDEF2 Simulation Model Design, IDEF3 Process Description Capture IDEF4 Object-Oriented Design and IDEF5 Ontology Description Capture [17]. Out of these IDEF5: Ontology Description Capture is related with the Astrakan conceptual model language [18]. The Astrakan method which has been used in the modelling activities in the studies is described in short in the Methodology section.

Every modelling language has its specific purpose but can also be used in other different situations.

3. METHODOLOGY

3.1. Interactive research

The modelling activities have been executed in two different case studies in the same research project - the Kaikaku-project at Mälardalen University in Sweden. In both cases participants have been represented from both the research domain as well as the company domain which is illustrated in Fig. 2.
Characteristic data collection methods have mainly been used in this qualitative research: observations and documentation [20]. Collecting and analysing data is a simultaneous process in qualitative research [21]. Direct observations mixed with participation were applied consistently over the case study during meetings, presentations and modelling workshops. Reflections and ideas have been written down in a “field journal”.

Further, qualitative research applied in this type of case study always faces the threat to validity of the research [22]. The main countermeasure for avoiding these threats have been triangulation, where convergent data is collected from different sources such as research seminars, and workshops with the participating companies and through respondent validation through presentation of the results from the different case studies at different conferences [23, 24].

3.2. Concept modelling

The Concept modelling language that has been used is developed by the Astrakan education company and is similar to IDEF5 and other ontology notation systems.
The main purpose with the Astrakan method is to clarify terms and words relationship to each other in order to define specific terms. The method consists of a work process and a notation standard for creating a graph according to Fig 3. Fundamental steps in the work process have been performed according to the Astrakan modelling method description [18]:

1) *Define project*, i.e. that the concept model project needs to have an objective, a project plan and dedicated resources.
2) *Term inventory and prioritizing*. In this step the groups, in the two different modelling activities, first made a literature study from which relevant terms were extracted, and also presented to the group relevant terms from the participating companies which were added into a gross list of terms.
3) *Modelling and define terms*. The gross list were analysed here and grouped into relevant groups of synonyms, homonyms or other relevant groupings. The models consist of a graph based on the symbols in Fig. 3. and a definition that could be derived from it. These were incrementally developed in the two activities.
4) *Decision and use*. The models have subsequently been used as input to the case studies. In the first case the graph and definition were used as a blueprint for a Kaikaku-method and in the second case as description model for an improvement system.

4. RESULTS

Two different modelling activities in two work packages within the same project, the “Kaikaku project” have been used as empirics in this study. The results are described below.

4.1. *The Kaikaku modelling activity*

The Kaikaku method development process in the project has been carried out in a multidisciplinary context in several iterative steps. Quite early in the project concept modelling came up as a feasible method to define the group’s view on ‘Kaikaku’ in order to find a requirement specification for the method. According to the Astrakan concept modelling method the second step after defining the project is to make a terms inventory. This was carried out in a series of workshops [18]. The terms were found mainly through brainstorming and from literature. A gross list of terms was created. In the third step the term kaikaku and the closest terms were selected by the team in consensus based upon the current understanding of Kaikaku within the specific project context. A graph and a definition were created. Some new terms were needed and introduced and the resulting graph is presented in Fig. 4.
The definition that was derived from the graph runs as follows:

**Kaikaku is a process that requires aggressive target setting, leads to radical change and is facilitated by innovative thinking.**

This definition of Kaikaku postulates that the method should guide, support or instruct that:

- An aggressive target is set
- A production strategy is used
- A radical change in the production system is planned and performed
- The change leads to a performance increase of critical measures
- Facilitation of innovative thinking

Out of this a first draft of a Kaikaku method was developed [22]. A case study company was selected to test this first version of a method [24]. The different concepts in the model were connected to the work process both as input/output content or (sub-) processes and activities. For example the concept “innovative thinking” was manifested as a creative workshop to create a new future state of a welding manufacturing cell. Here the connection between the development work process and the concept model was clear. There was also a pointed out need for better descriptions of the content of each concept. The model was created and enhanced in three one and a half hour workshops. The discussions were very focused on the model and clearly facilitated consensus regarding the teams view on the term kaikaku.

### 4.2. The Holistic Improvement System model

The background of this modelling activity was the research group’s need for a description of a general and holistic improvement system. It was unclear however what such a system looks like and what components it consists of. Thus, the objective of the study was to investigate and describe how an improvement system from a holistic view relates to enterprise processes. A second objective was to define the improvement system in order to enable a good choice of programs/methods/tools for an enterprise’s overall strategies [23].

First an overall literature review was performed in order to investigate the terms: ‘holistic’ and ‘improvement system’ in the literature. Concept modelling was used in order to
clarify the terms relationship between an improvement system and general parts in a manufacturing enterprise.

As a second step the purpose was defined and the concept of a holistic improvement system was further investigated in a more detailed literature review. In the third step a gross list based on the literature review of terms was assembled. The fourth step was to group and prioritize the terms into “headings” and model the terms into a concept model graph. The graph has then been revised according to comments from other researchers first and then industrial experts in improvement work. The first version of the map was enhanced after the feedback. The concept model graph in Fig. 5. then represented the research group’s shared view of a “Improvement System” which was later used as a basic viewpoint for related research in the “Kaikaku-project”.

Fig. 5. The concept model graph defines an improvement system according to the group.

The definition that can be derived from this concept model graph runs as follows:

An improvement system consists of programs, methods and tools, supporting different perspectives of improvement, thus optimizing the performance of the production system. The production system, which the improvement system is a part of, is measured by the performance measurement system and is controlled by the management system in coordination with the strategy.

An interesting topic derived from the concept modelling itself, as well as from the discussion over the model with the industry representatives, is how to use the concept model in practice in the future. What has been developed so far is a concept model that theoretically explains how enterprise different parts relate to each other including with the improvement work in focus. The experiences from working with concept modelling in this case were that it led to relevant discussions and decisions in the development group. The presented model was created in four one and a half hour workshops over a two weeks period.
5. CONCLUSIONS, DISCUSSION AND FUTURE WORK

Is concept modelling a fruitful tool in an improvement system design and development context? The answer is ‘yes’. The conclusions so far using concept modelling are that the method facilitates concordance between the method development team members. This way of working will lead to well-defined methods fairly swiftly and in consensus according to Shaw and Gains [4]. Axelsson and Hidefjäll also pointed out that the modelling team should have insights and visions regarding the enterprise [5]. These studies strengthen that statement. The clarity of the terms and concepts also provides a good skeleton for description of developed methods. This will most likely ease innovation diffusion, both internally in the enterprise or to others.

To lead a concept modelling workshop demands training but with practice this way of working seems fairly easy to adopt. In the two cases one can note that the results are not quite in accordance with the notation standard. However, the concept models facilitated concordance between the team members and in a shorter time period than expected – i.e. the method is fruitful.

One problem was that there are already many different definitions and this methodology let us create another. But who are the definitions for? In this case they are for researchers and practitioners who are working in their own specific context and the concordance in the group or company is more important. The concept modelling method could include a literature review which in turn provides “traceability” between the literature and the result. This will serve as the “memory” of where the solution later comes from.

In order to develop a useful support tool that could be applicable in industry in the future, the concept model method needs to be further improved. One way to accomplish this, suggested by the research team, is to work in three steps; 1) concept model, 2) content model, and 3) process model. The concept model theoretically explains terms/concepts in relation to each other. The content model imparts a deeper understanding of the concept model’s different parts and could be used to develop e.g. check lists. The process model might support practical use of the developed method in question. Future research could validate and verify these steps and determine whether the sequence of these steps is important, or if any of them could be performed concurrently.

ACKNOWLEDGEMENTS

This work has been carried out within the XPRES program (www.xpres.se) in cooperation between, Mälardalen University in Eskilstuna (MDH), Swerea (IVF and KIMAB), and Royal Institute of Technology (KTH). XPRES is funded by the Swedish government through VINNOVA (Swedish Governmental Agency for Innovation Systems). The case studies have also been funded by VINNOVA (Dnr: 2009-03978) within the “Kaikaku project” carried out at Mälardalen University.
REFERENCES


The IDEAS Plug & Produce System
Mauro Onori¹, Antonio Maffei¹ Friedrich Durand²

¹Dept. of Production Engineering, KTH Royal Institute of Technology, Stockholm, Sweden  
²ELREST GmbH, Kirchheim, Germany  
onori@iip.kth.se

ABSTRACT

Current major roadmapping efforts have all clearly underlined that true industrial sustainability will require far higher levels of systems’ autonomy and adaptability. In accordance with these recommendations, the Evolvable Production Systems (EPS) has aimed at developing such technological solutions and support mechanisms. Since its inception in 2002 as a next generation of production systems, the concept is being further developed and tested to emerge as a production system paradigm. The essence of evolvability resides not only in the ability of system components to adapt to the changing conditions of operation, but also to assist in the evolution of these components in time. Characteristically, Evolvable systems have distributed control, and are composed of intelligent modules with embedded control. To assist the development and life cycle, a methodological framework is being developed.

After validating the process-oriented approach (EC FP6 EUPASS project), EPS now tackles its current major challenge (FP7 IDEAS project) in proving that factory responsiveness can be improved using lighter Multi-Agent technology running on EPS modules (modules with embedded control). This article will detail the particular developments within the IDEAS project, which include the first self re-configuring system demonstration and a new mechatronic architecture.

KEYWORDS: Evolvable Production Systems, Modularity, Distributed Control

1. INTRODUCTION

According to the results attained by many roadmaps [1],[2],[3] one of the most important objectives to be met by European industry is sustainability, which is multi-faceted: including economical, social and ecological aspects. The obvious conclusion to this holistic problem is that future manufacturing solutions will have to deal with very complex scenarios. The truly interesting characteristic of this conclusion resides in the word “complex”. Although often used, the basic essence of complexity is that it may not be fully understood and determined in all its ruling parameters; however, we seem to continue to build production systems based on known functionalities and predicted operational scenarios. This, to the authors, remains a rather disturbing factor.
Albeit the enormous efforts made in the 1990’s by Flexible Assembly and Manufacturing Systems, followed by Holonic [4] in the late 1990’s-early 2000s, Reconfigurable Systems [5] and other approaches, the dream of cost-effective, high-variant assembly remains elusive. One of the reasons may lie in the fact that one cannot solve unpredictable scenarios with a focus on predictable functionalities. Nature does not work with predictability. Nature does not propose an evolutionary change based on a single factor, nor does it do so by selecting a single motivating factor. Living organisms evolve by proposing a variety of solutions, but this is done in ways that are not yet fully understood. Yet the adaptation is guaranteed.

Based on this pre-conception that it may be more realistic to assume that a production environment is not fully predictable, and that we should not focus entirely on the required functionalities alone, Evolvable Assembly Systems was proposed in 2002 and has, since then, been developed and tested to emerge as a production system paradigm (see EUPASS, A3 projects [6], as given by [7])

In January 2011, the EPS approach was finally proven to work at the FESTO premises in Germany. The assembly system was re-configurable and exhibited self-organisation. Developed in the IDEAS FP7 project, the details will be given herewith.

2. BACKGROUND

First of all, the solutions proposed by EAS are not intended to be understood as a general panacea for all assembly scenarios. At present they are a potentially cost-effective approach to large variant flora production and/or short product lifecycles. Once the methodology is completed, and the technology matures, EAS could become more generally viable and cost-effective.

EAS may be viewed as a development of reconfigurability and holonic manufacturing principles. It was initially developed in 2002 from the results of a European roadmapping effort (Assembly Net), and was subsequently further developed in a series of European projects (EUPASS, A3, IDEAS). Its objectives have all been drawn from roadmapping conclusions and are well elaborated in earlier publications [8],[9].

As defined in [10]) RMS incorporates principles of modularity, integrability, flexibility, scalability, convertibility, and diagnosability. These principles impose strong requirements to the control solution. In particular, centralized approaches become completely unsuitable due to their intrinsic rigidity. Decentralised solutions must be considered that take into account the fundamental requirements of pluggability of components, which includes the aspects related to dynamic addition/removal of components, as well as adaptation in the sense that the system does not need to be reprogrammed whenever a new module is added/removed. This is a fundamental aspect behind any control solution approach to solve the defined requirements. Therefore, the major challenge in the control solution is how to guarantee proper coordination and execution in a system in which both its components and working conditions can be dynamically changed. This is a challenge that needs a completely new approach and this is why in the context of EPS a solution based on concepts inspired from the Complexity Theory and Artificial Life is being developed. The next section covers what concepts from non-traditional manufacturing research domains are being used to create truly dynamic control solutions.

Hence, the control approach to be developed in the context of EPS wants to go back to the basics, that is to say relying stringy on the original idea of considering each component as a distributed intelligent unit that may aggregate in order to create a complex system. In this context, concepts such as emergence and self-organisation become more and more important
to be applied to new generation control solutions. However, true implementations of these new concepts within shop floor are still very few.

Considering what was stated above, one may view Evolvable Production Systems (EPS) as a development of the Holonic Manufacturing Systems (HMS) approach; however, a closer looks reveals that, although there are similarities in the exploitation and implementation phases, the paradigms differ quite substantially in their perspective (or trigger issue), and that only EPS achieves fine granularity. By granularity it is considered the level of complexity of the component that compose a manufacturing system. For instance, when a line is composed of several cells and these cells are modules that can be plugged in and out, this is coarse granularity. If, on the other hand, the components that can be plugged in or out are grippers, sensors, or pneumatic cylinders, this is fine granularity. This issue is in fact a very important one in terms of distinguishing the paradigms. The target for EAS is the shop-floor control, which normally demands programming, re-programming and vast integration work.

This is where EAS plays a decisive role. The two fundamental aspects are:
1. A methodology that allows the user to define modules at fine granularity level, from a control-point-of-view.
2. The development of control boards capable of running the agent software and, simultaneously, be small enough to be embedded in the smaller modules.

The IDEAS project proved this to be viable as a shop-floor solution.

3. IDEAS-THE BASICS

IDEAS stands for Instantly Deployable Evolvable Assembly Systems. This is an FP7 project that started in 2010 and will end in 2013. It is to develop EAS systems for two industrial customers, IVECO and ELECTROLUX.

The project took advantage of several developments that were done during the EUPASS (FP6...) project, such as:
- ontological descriptions of the assembly processes [9],
- equipment modules prepared for embedded control [11],
- data exchange protocols verified, [12],[13],
- basic methodological principles set [14].

IDEAS had as a main objective to implement the agent technology on commercially available control boards. This would enable distributed control at shop-floor level. What is being considered here is not the planning or logistics level but the actual operational level of the assembly system.

To this effect the ELREST company and FESTO research division set out to specify the exact requirements, based on the needs detailed by the industrial customers Electrolux and Centro Ricerche FIAT. MASMEC, Karlsruhe Institute of Technology and FESTO supported the effort by developing system modules, TEKS provided the simulation software, and UNINOVA and KTH developed the agent technology. Finally, the methodological framework upon which the whole project would base its work, was developed by University of Nottingham.

The project’s first objective was to prove the validity of the approach by running a medical assembly system at the FESTO facilities (see Fig.1).
The system shown above ran the following processes:

- **Glueing unit:** Dispensing glue for assembly of small components
- **Pick & Place unit**
  Pick and place handling system
- **Electrical testing unit**
  Testing unit for quality/functional product test
- **Stacker unit**
  Pneumatic/Servopneumatic handling system

This assembly system, called the MiniProd, was finally demonstrated in January 2011. It ran with a multi-agent control setup, could be re-configured on-the-fly, and the modules self-configured. This was achieved thanks to the fact that the agent software could be run on commercial control boards (Combo, ELREST), which are shown in the following Fig.2.

As this could probably be viewed as the first time an assembly system actually operated with a totally distributed control system, and self-configured, it was shown again for the European Commission in November 2011. The system performed flawlessly, confirming that multi-agent control can be used for truly reconfigurable assembly.
4. THE IDEAS DRIVERS

In order to attain this success, IDEAS has relied on many years research (including the work done in RMS, etc.) and the following own developments:

- A simple and effective mechatronic architecture
- Control boards developed for multi-agent applications
- An elaborate and well-structured methodology
- Industrial commitment

The mechatronic architecture is, first of all, an architecture that considers the control demands from an embedded-system point of view. That is, each assembly system module is an entity with its own control, hence the “mechatronic”. The difficulty was in creating an architecture out of which an effective control structure could be instantiated for any assembly system layout. As the demands on assembly are extremely diversified (see conveyor system in MiniProd-free-moving pallets!), this posed challenges. The final Mechatronic Architecture is based on four basic agents:

- Machine Resource Agent
- Coalition Leader Agent
- Transportation System Agent
- Human Machine Interface Agent

In order to implement this, the project developed several tools. The actual agent development environment, called IADE (IDEAS agent devt.env.) is based on an elaboration of JADE. The Java Agent DEvelopment framework is FIPA compliant and also provides basic development tools. The IDEAS project further developed these tools and included others to support the simulation of the agent control prior to its being downloaded into the modules. Experiments made at the simulation level and real module also indicated that the simulated module and real unit actually run the exact same code, rendering the simulation extremely accurate (1:1 relation).

The second main development has been the development of commercial control boards capable of running the multi-agent setup. The ELREST company provided the project with several alternatives, out of which the Combo211 was selected for use. This required quite some developments, amongst which:

- Combo200 series runs on WinCe6
- Implemented CrEme™, a Java Virtual Machine (NSI.com)
- Fits to the needs of the Agents and supports JADE
- Implementation of 24V I/Os, Ethernet, CAN and RS232/RS485 connections

The control boards function very well and have also been thoroughly tested at the other partners labs. The project currently intends to develop three variants of these control boards, depending on the required granularity and number of agents/module (from very small, cheap, to mid-size capable of running more than one agent).

Thirdly, the project would have never succeeded if the tools that are required to engineer such solutions were not specifically designed and integrated within the IDEAS methodology. This work, led by University of Nottingham, has brought together many partners (KTH, MASMEC, KIT, TEKS, ELREST, FESTO): the synchronisation and integration are sensitive aspects. The objectives included:

- Develop Semantic Representations for Devices and Skills
- Create Requirements and Target Specification Language
- Semantic Rules for Integration & Validation of Skills
– Develop a rapid System Configuration Environment
– Develop Visualisation and Transparency Tools

Note that this includes skill definition support, Workflow definition support, simulation tools and more.

One of the most interesting outcomes of the work has been the link between simulated system and real system. Using commercial software (Visual Components) coupled to the multi-agent programs made it possible to run the exact run-time code prior to download. That means that the simulations represent exactly what will occur in reality (at control level).

All the developments, from EUPASS to IDEAS and beyond, would be quite superfluous if industry had not provided the critical mass and know-how to achieve such results. Industrial aspects are the key ingredient as the certification procedures, variation of hardware constraints, specific customer needs, market demands, etc., all play a decisive role in the effective deployment of a technology. IDEAS took this a step further as it set as an objective that one of the “missing links” had to be corrected: develop a control board for such applications. This was made possible by the industrial commitment, both at control development and requirements specification.

5. FUTURE STEPS

The project is now consolidating these results and developing them further. The next step will be to build two industrial systems, in order to verify the full-scale utilisation at customer-level. The two systems will be built at KTH (Stockholm) and MASMEC (Bari). The products to be assembled are an ECU (electronic control unit) from a commercial vehicle, and some specific washing-machine components. The figure 4 below illustrates the schematic layout.

Figure 5. The ECU Assembly System(MASMEC)

Both solutions will be thoroughly validate and Life-Cycle Analyses performed. Finally, a new Business Model has just begun to be developed in support of the more strategic decisions that will be encountered.
6. CONCLUSIONS

The article describes the first realistic developments for multi-agent control for assembly applications. The work is extremely valuable but there remains a fair amount of research and development work to be done.

First of all the human role in such automated systems needs to be studied such that people may become an integrated element in EAS solutions. This includes the development of role models, interfaces and data capture methods. Secondly, the tools mentioned earlier need substantial elaboration, such that a solid and robust development methodology (guidebook and set of tools) can be generated. This is a highly multi-disciplinary requirement as computer specialists will have to collaborate with production and system engineers at detailed level.

Industrially, these solutions seem to generate sufficient interest, especially as these first tests have clearly shown the viability. The show-stopper is, therefore, not particularly at industrial level but, rather, at academic: consensus as to which “paradigm” is chosen as the most promising is not being based on true industrial development results but on theoretical details. This attitude needs to change and closer, more practical collaboration is required in order to truly support industry. As Thomas Kuhn would possibly put it, we must abandon normal science and search for a true industrial breakthrough.

ACKNOWLEDGEMENTS

The authors wish to thank the European Commission for the funding and, in particular, their Project Officer Jan Ramboer.

The IDEAS partners have shown incredible commitment and worked well beyond the budget limits and normal working hours. Therefore our deepest thanks go to KTH, FESTO, Electrolux, UNINOVA, MASMEC, Univ.of Nottingham, ELREST, TEKS, Karlsruhe Inst.of Technology, and Centro Ricerche FIAT.

One of the final demonstrators, to be built at KTH, is also developed with the collaboration of XPRES (eXcellence in Production RESearch), a Swedish national R&D initiative.

REFERENCES


[13]. L. Ribeiro, J. Barata, J. Ferreira; “An Agent-Based interaction-oriented shop floor to support emergent diagnosis”; Proceedings of 8th International Conference on Industrial Informatics (INDIN); Osaka, Japan; July 2010.

Laser assisted milling of a nickel based alloy using different insert geometries

Michael Kuhl¹, Martin Dix¹, Giacomo Leopardi¹, Flaviana Tagliaferri¹,², Carlo Rüger¹

¹Chemnitz University of Technology, Germany
²University of Naples Federico II, Italy
carlo.rueger@mb.tu-chemnitz.de

ABSTRACT

Laser assisted milling (LAM) is a hybrid machining technology which aims on the benefits of hot cutting technology using the laser beam as local heating source. In this study nickel based Alloy 59 is cut with conventional flooding and laser assisted milling under dry conditions. Therefore, a statistical approach based on design of experiment (DoE) is carried out in order to investigate and compare the effect of LAM process parameters and their interactions. Hence, full factorial design and analysis of variance (ANOVA) are applied. The influence of the process parameters insert setup, tool feed and laser spot scanning frequency are analysed. The aim of this paper is to investigate the benefits of LAM compared with conventional milling in terms of tool wear, cutting forces, surface roughness and machinability improvements.

KEYWORDS: hybrid machining, laser assisted milling (LAM), difficult to machine materials, design of experiments (DoE), cutting forces, tool wear

1. INTRODUCTION

Nickel based superalloys are widely used in aerospace, oil and chemical industry due to their superior properties like high mechanical strength, excellent corrosion resistance and chemical stability at high temperatures [1, 2]. However, these alloys show a poor machinability, due to short tool life and high cutting forces [3]. A method to improve nickel alloys machinability, but more in general superalloys, consist of reducing material strength by means of hot machining (Fig. 1). This can be achieved using a laser beam as localized heating source before cutting edge (Fig. 2). The laser heating changes temporarily the material properties reducing the mechanical strength.

Indeed, good results, in terms of cutting forces reduction and long tool life were scored using laser assistance in turning and micromilling [4-7].
However, hybrid machining like laser assisted milling (LAM) started to be developed since a few years. Researchers aim to achieve similar results to laser assisted turning when machining superalloys (i.e. titanium, Inconel 718).

Due to the complexity of milling process interactions with laser heating (cutting speed, laser power, etc…) only a few papers [4, 8] are published in this field.

The main objective of this study is to determine effective parameters of LAM by means of Design of Experiment (DoE) comparing to conventional milling by changing feed rate, insert and laser scanning frequency.

2. EQUIPMENT AND EXPERIMENTAL PROCEDURE

The test procedure was executed on a 5 axis milling machine (DMU 125P). A continuous wave (CW) medium power laser beam is generated by diode laser Optotools 800 W ($\lambda = 915$-$980$ nm) (Fig. 3).

The laser optics is provided by using an f-theta lens and the laser beam is deflected by a mirror, placing the beam in front of cutting tool. The workpiece material (58 x 21 x 7 mm$^3$) is alloy 59 [14], provided by ThyssenKrupp (Table 1).
Table 1. Chemical composition (in weight %) of alloy 59 [14]

<table>
<thead>
<tr>
<th>element</th>
<th>Ni</th>
<th>Cr</th>
<th>Fe</th>
<th>C</th>
<th>Mn</th>
<th>Si</th>
<th>Mo</th>
<th>Co</th>
<th>Al</th>
<th>P</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>min.</td>
<td>22.0</td>
<td>15.0</td>
<td>0.1</td>
<td>15.0</td>
<td>0.1</td>
<td>16.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.015</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>max.</td>
<td>bal.</td>
<td>24.0</td>
<td>1.5</td>
<td>0.01</td>
<td>0.5</td>
<td>0.1</td>
<td>16.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.015</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The tool is a square shoulder milling cutter with a diameter of 25 mm. Two different inserts (type A and B) were used for the experiments (Table 2). Insert type A is recommended for nickel based superalloys, especially in unstable conditions. The main focus of this insert geometry and coating (chemical vapour deposition CVD) is a high mechanical toughness.

On the contrary insert type B offers a sharper cutting edge and a more positive rake angle for an easier chip flow. Additional it has a highly heat resistant coating (physical vapour deposition PVD). Therefore, it is recommended especially for milling stainless steel which offers a lower mechanical strength but strong thermal load and smearing tendencies.

The main difference in comparison with insert A is the reduced mechanical toughness but the increased heat resistance.

Table 2. Main milling inserts properties

<table>
<thead>
<tr>
<th>Insert A</th>
<th>Insert B</th>
</tr>
</thead>
<tbody>
<tr>
<td>substrate</td>
<td>tungsten carbide</td>
</tr>
<tr>
<td>coating</td>
<td>multilayer CVD</td>
</tr>
<tr>
<td>geometry</td>
<td>strong, light positive geometry</td>
</tr>
<tr>
<td>advantage</td>
<td>positive axial rake angle</td>
</tr>
<tr>
<td>recommended</td>
<td>excellent in unstable conditions</td>
</tr>
<tr>
<td>titanium, superalloys</td>
<td>stainless steel</td>
</tr>
</tbody>
</table>

Each test was performed using a new cutting edge. During the tests the resulting forces ($F_x$, $F_y$, $F_z$) were measured by a dynamometer (Kistler 9255B). The signals were amplified (QuantumX) and recorded by the PC software (Ecatman-easy). An optical microscope (Keyence 600) was used to investigate the wear development of the cutting edges. The surface roughness of the milled workpieces was measured by a confocal microscope.

All LAM tests were done in down milling mode, under dry conditions and with inner air cooling (6 bar) through the tool for a better chip removal. The tests were performed with the laser spot scanning and heating workpiece at different frequencies in front of the milling tool and using different machining parameters (Fig. 3). The Design of experiment (DoE) and all relative parameters used are described in the next section.

3. EXPERIMENTAL DESIGN AND SET-UP

Before proceeding with tests, some pre-tests were done for screening in order to find a good configuration of experiments and minimize undesired effects (noise factors influence during experiments, etc…). Thus, a full factorial design $2^3$ was adopted (with a defining relation of $I = ABC$) where no main effect or two-factor interaction is aliased with other main effects or two-factor interactions [9].

The control factors used were as follows: feed per tooth $f_z$ (A), laser scanning frequency $f_s$ (B) and insert type (C). The tested value of the control factors (i.e. control factor levels) are reported in Table 3.
Fig. 3. Laser assisted milling: scheme and machine set-up details. Laser optics and scanner (1), laser fixture (2), milling machine head (3), deflecting mirror (4), spindle (5), milling tool (6), workpiece (7), vise and thermocouples (8).

Table 3. Control factors of DoE with relative values.

<table>
<thead>
<tr>
<th>Control factor</th>
<th>Low level(-)</th>
<th>High level(+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Feed per tooth $f_z$ [mm]</td>
<td>0.1</td>
<td>0.16</td>
</tr>
<tr>
<td>B Scanning frequency $f_s$ [Hz]</td>
<td>12.5</td>
<td>25</td>
</tr>
<tr>
<td>C Insert type</td>
<td>Type A</td>
<td>Type B</td>
</tr>
</tbody>
</table>

Feed per tooth is specified from manufacturer catalogue and laser scanning frequencies were chosen after pre-test analysis.

The other milling parameters were kept constant with following values: cutting speed $v_c = 40$ m/min; depth of cut $a_p = 1$ mm; radial depth of cut $a_e = 7$ mm (28% of tool diameter).

Each test was repeated three times (24 runs in totals). The workpiece was machined in down milling strategy [10] and a machining path of 55 mm was chosen for each test.

4. STATISTICAL ANALYSIS AND RESULTS

The ANOVA method was applied in order to test the statistical significance of the main effects and the two-factor interactions for maximum value of active force ($F_a$), maximum value of passive force ($F_p$) and the maximum value of first peak of active force ($F_{a,ci}$).

Diagnostic checking was successfully performed via graphical analysis of residuals. The experimental results for maximum value of active force and the maximum value of first peak of active force are shown in Fig: 4, respectively, using Pareto charts of standardized effects ($\alpha = 0.05$).

The control factors which have significant effects ($\alpha = 0.05$) on the maximum value of active force are feed per tooth $f_z$ (A), insert type (C), and interaction between feed per tooth $f_z$ and laser scanning frequency $f_s$ (AB).

Fig. 5 plots the main effect and interaction effect of the control factors, respectively for this response variable. As expected the maximum value of active force increases when the feed per tooth $f_z$ (A) increases, but it should be noted that its mean value is lower when the insert type B (C) is used.
The analysis of the results also points out the statistical significance of the interaction between feed per tooth $f_z$ and laser scanning frequency $f_s$ (AB); when the feed per tooth $f_z$ (A) is at the low level (0.10 mm) the mean value of the maximum active force (Fig. 5 right side) is lower at the high level of laser scanning frequency $f_s$ (B).

Regarding the maximum value of the first peak of active force, the control factors that have significant effects ($\alpha = 0.05$) are feed per tooth $f_z$ (A), scanning frequency $f_s$ (B), and their interaction (AB). Fig. 6 plots the main effects and interaction effects of the control factors, respectively for this response variable. As expected the maximum value of first peak of active force increases when the feed per tooth $f_z$ (A) and the laser scanning frequency $f_s$ (B) increase, but there is no significant influence of the insert type (C). Moreover, the interaction plot (Fig. 6 right side) shows at the high level of the feed per tooth $f_z$ (A), the influence of the laser scanning frequency $f_s$ (B) is higher in terms of maximum value of first peak of active force.
5. LAM AND CONVENTIONAL MILLING: CUTTING FORCES, TOOL WEAR AND SURFACE ROUGHNESS.

5.1. Tool wear

After each test the inserts were checked under a microscope. Both types show good results in milling with conventional emulsion flooding (Fig. 7). After cutting, the tools exhibit regular shaped starting flank wear. With the feed per tooth of \( f_z = 0.1 \) mm the rake faces are hardly polluted with embedded material and the cutting edges show no critical damage or breakouts. With the higher feed of \( f_z = 0.16 \) mm for type A and B the formation of build-up edge (BUE) increased slightly. Both inserts show starting notching in the depth of cut (DOC) line. All milling lanes in conventional cutting show little burr formation in the cut-out area like seen in Fig 7.

However, in LAM the cutting process using insert A showed growing vibrations. The chips are not removed away from the insert after the cut-out, but constantly welded to each other forming a long belt of chips sticking on the cutting edge. In general, insert A suffers from strong BUE and cutting edge breakouts leading to heavy tool damage, what is assumed to be the main reason for the process vibrations.

The coating and geometry of Insert A are designed to resist high mechanical loads of strong and abrasive material in unstable cutting conditions (Table 2). Hence, it gives a very good performance in conventional cutting of nickel based alloys. However, in LAM the additional heating energy and the missing coolant lubrication moves the process properties significantly away from Insert A's intended use.

Due to the reduced material strength the main tool load in LAM is not mechanical but thermal induced beyond the optimal operating range of insert A.

Insert B shows in LAM less embedded material than insert A, but small breakouts occurred along the cutting edge under all tested conditions. Thereby lower feed of \( f_z = 0.1 \) mm shows tendency to increase the amount of BUE on insert B and the higher feed \( f_z = 0.16 \) mm tends to raise the size and number of breakouts. Insert B is designed to have high heat resistance, low friction coating and a sharp cutting edge for an easier chip flow especially in ductile materials like stainless steel.

Fig.7. Inserts wear and burr formation in conventional milling and LAM (\( f_z = 0.1 \) mm). In order: milling insert A (yellow), insert B (black) and milled surface.
The properties of the part material are changed by laser heating reducing the strength but raising the temperature in the cutting area drastically. These conditions, apart high temperature, are close to the usual working range of insert B which performed better results than insert A.

In comparison with conventional flooding in LAM both insert types show a heavy increased burr formation in the cut-out area. This can be mainly ascribed to an interaction of the changed material and cutting edge properties. In an emulsion cooled process the uncut workpiece material is cold and brittle, but in LAM the surface reaches higher temperatures becoming softer and more ductile, increasing chip flow.

On the other hand breakouts and BUE on both insert types change the cutting edge conditions generally degrading the sharp cutting edge of the tool. These negative circumstances are likely to prevent the tool from clean cutting. Hence the material is not removed as a chip, but squeezed and pushed aside by forming the burrs [12, 15]. This effect dominates in the cut-out area, where the undeformed chip thickness becomes lower than the actual effective cutting edge radius of the tool.

5.2. Cutting forces in LAM and conventional milling

In order to understand the main advantages and issues, the cutting forces of LAM (dry process) were compared with conventional machining (flooding lubricant). Actually, in down milling the undeformed chip thickness is changing during the cutting contact. The maximum thickness \( h_{\text{max}} \) is at the beginning of the edge engagement (cut-in) and ends up with \( h = 0 \) at the cut-out point [13]. The machining forces \( \left(F_{\text{cut}}, F_{\text{feed}}, F_{\text{p(assive)}}\right) \) are directly related to \( h \) value, therefore in applied shoulder milling the maximum forces and hence mechanical tool load occur close to the beginning of the edge engagement [13].

The experimental setup of the laser heating was concentrated very local in the cut-in area where the effects are expected to achieve the best results. The dynamometer records the resulting forces on three orthogonal axes as \( F_x, F_y \) and \( F_z \) from which the active and passive forces are calculated as the machining forces in LAM.

The active component of the machining force \( (F_a) \) is defined as the resulting force in the working plane. In shoulder milling it can be calculated by equation (1) [16]. The passive machining force component \( F_p \) is defined orthogonal to the active force and can be directly assigned to \( F_z \) like seen in (2).

\[
F_a = \sqrt{F_c^2 + F_N^2} = \sqrt{F_f^2 + F_{f(ass)}^2} = \sqrt{F_x^2 + F_y^2} \tag{1}
\]

\[
F_p = F_z \tag{2}
\]

When the insert engages the material, the active machining force \( F_a \) increases rapidly. In conventional milling this causes an abrupt impact load on the tool during each cut-in. On the other hand, in LAM, the cutting edge enters immediately the pre-heated and softened material area.

Due to the decreased strength a considerable smaller force peak marked as \( F_{a,ci} \) was measured (Fig. 8 left side). The subsequent cutting path then passes directly through heat softened area, requiring less force. The local minima of \( F_a \) and \( F_p \) are found close to the centre of the heating path indicating the highest reduction of the material strength. After leaving the heating zone the machining forces in LAM rise to a maximum.
Fig. 8. Machining forces in LAM and conventional milling.

In the cut-out area the forces exceed the comparable values of conventional milling forces on each parameter setup (Fig. 8 left side). As explained earlier the reason for this is arguably due to an interaction of the missing lubrication in LAM and the decreased cutting ability through strong BUE and damaged cutting edges.

In conventional milling $F_a$ shows only one maximum point for each tool contact. For every milling test the average value of the maximal force peaks $F_a$ and $F_{a,ci}$ (Fig. 8 right side) of all single edge engagements were calculated.

Despite their different geometries and coatings inserts A and B show comparable active force values on both tool feed rates in the conventional emulsion flooded milling process. In LAM the machining forces of type A shows a high fluctuation between the single tests. This can be mainly ascribed to the heavy BUE and tool damage described previously. Thereby conclusions on the influence of certain process parameters on the machining forces can hardly be made for insert A because of the perpetual changing cutting edge conditions.

Insert B, which was less afflicted by tool damage and BUE in LAM, showed a decrease of the machining forces. The highest reduction of the entire maximum of $F_a$ on the feed of 0.1 mm was 25% and 29% on the feed of 0.16 mm (Fig. 8 right side). These results are comparable with the experiments of Zäh and Wiedemann. They had found a similar behaviour of cutting forces and measured a reduction of the maximal peak of about 20% on LAM of titanium alloy TiAl6V4 [17]. But their work did not mention the decreasing of the cut-in force peak $F_{a,ci}$ in LAM seen in Fig. 8 (left side).

The maximum load occurring on tool impact could then be reduced by 29% on feed per tooth 0.1 mm up to 37% on feed per tooth 0.16 mm in LAM nickel based alloy 59. In LAM a higher feed rate can therefore achieve larger material removal rates without an increased mechanical tool load. As seen in Fig. 8, the reduction of the force peaks was also dependant on the scanning frequency. It was found that the slower spot movement effects lower maximal forces.

5.3. Milled surface roughness

After each milling test the surface roughness (Sa) was measured in 3 points of the tool path. Generally, the values of roughness are influenced by different parameters i.e. machining parameters and cutting phenomena [18]. However, due to the high tool wear and breakage of insert A, only roughness generated using insert B was analysed.

It was found that insert B has no benefits on surface roughness in LAM compared to conventional milling. In Fig. 9 are shown the values of Sa at $f_z = 0.1 - 0.16$ mm with scanning frequency $f_s = 12.5 - 25$ Hz and conventional milling.
Fig. 9. Surface roughness Sa using insert B in LAM and conventional milling.

On the left side of diagram \((f_z = 0.1 \text{ mm})\) it can be noticed that increasing \(f_s\) increases \(Sa\). On the contrary, at feed \(f_z = 0.16 \text{ mm}\) increasing \(f_s\) causes a decrease of \(Sa\).

This trend of \(Sa\) can be related to the characteristics of LAM process (feed per tooth and laser heating frequency). Indeed, tool wear of insert B is similar in both testing conditions.

6. CONCLUSIONS

An influence of the laser heating on the milling process of nickel based alloy 59 was shown. The forces in the cut-in area were lower as well as the occurring force peak are lower for LAM in comparison with the tool impact of conventional machining. Additionally, the maximal occurring active machining force \(F_a\) could be reduced. So in LAM a more consistent smooth cutting can be achieved avoiding high local force peaks.

It was also found that the local heating and the necessary absence of the cooling lubricant (dry cutting) change the requirements of the cutting tool. Due to the increased temperature the material in the cutting zone exhibits decreased strength but an increased ductility. So the occurring forces are not the main cause for tool failure in LAM.

The most critical aspect for the used tungsten carbide tools was the additional heat. Especially, the smearing and sticking of material forming strong BUE is crucial for damaging the cutting edge. Both insert types were affected, but it could be shown that a sharp and heat resistant tool configuration, like insert B, which is mainly designed for cutting stainless steel, is advantageous in handling these changed requirements. Insert A with its strong configuration normally operating on nickel based alloys was found to be not suitable for the changed tool requirements of LAM. An explanation for this behaviour can be the interaction of tool shape with the LAM process.

Another main aspect found in LAM was the increased burr formation in the cut-out area. Nakayama and Arai found that sideward burr can be decreased by a higher tool rake angle and increased cutting speed [15]. Burr decrease should be investigated further.

Finally a recommended tool setup for LAM is a sharp edge combined with a highly heat resistant tool body. Due to the decreased material strength the geometry must not be as tough as in conventional milling nickel based alloys. Furthermore a positive rake angle and a low-friction surface are expected to be advantageous for an easier chip flow and hence less BUE formation. It was shown that in LAM the operation limits of coated tungsten carbide tools are reached. So ceramics or even CBN which can both resist higher temperatures are expected to be more eligible for LAM process.
ACKNOWLEDGEMENTS

The Cluster of Excellence “Energy-Efficient Product and Process Innovations in Production Engineering” (eniPROD) is funded by the European Union (European Regional Development Fund) and the Free State of Saxony.

REFERENCES

Theoretical and Experimental Studies of Internal Turning Using Damped Boring Bars

Ove Bayard, Lorenzo Daghini, Mihai Nicolescu

KTH Royal Institute of Technology, Dept. of Production Engineering, Sweden
ove.bayard@iip.kth.se

ABSTRACT
The benefit of passive damping solutions for vibration suppression is well established in various fields of mechanical and civil engineering. This paper presents the theoretical and experimental results regarding composite metal-viscoelastic materials used to design efficient damping treatments of tool holders used in internal turning. Boring bars for single-point internal turning are primarily susceptible to chatter and have been the subject of numerous studies. The present study was motivated by the applications where diameter-to-length ratio is greater than 5 and high dynamic stability is required. The computational model is used to determine the flexural and torsional vibration modes. Machining tests reveal good dynamic behaviour in a relatively wide range of cutting speeds and feed rates while machined surfaces keep a high quality surface finish for tool lives up to 90 min.

KEYWORDS: boring bar, viscoelastic material, cutting stability, tool life.

1. INTRODUCTION

Engineered passive damping for structures is usually based on one of four damping technologies [1]: (1) viscoelastic (VE) materials, (2) viscous fluids, (3) magnetics, or (4) passive piezoelectrics. Each of these damping mechanisms must be understood in order to select the most appropriate type of damping treatment. Efforts to produce metal alloys with both high internal damping and also adequate strength and resistance to failure by fatigue have not been very successful except for a few non-ferrous alloys described by Birchon [2]. These special alloys are difficult to produce and the damping is large only at high alternating strains.

Essentially, passive damping treatments work by absorbing significant amounts of strain energy from the vibration modes in the frequency range of interest and dissipating this energy through some type of energy dissipation mechanism. One of the approaches to suppress excessive vibration and noise in structures is the viscoelastic damping treatment, unconstrained or constrained. The added viscoelastic polymer which exhibits high material loss factor provides the major damping mechanism in vibrations due to its extensional deformation [3]. However, the lack of rigidity and excessive long term creep prohibits the
fabrication of a structure exclusively from high damping polymers. One option is to design a composite structure of metal with damping layers bonded to it [4].

Although not always consciously designed in, machine tool structures benefit from passive damping that comes from friction in fixed and movable joints. As machine tool structures are over dimensioned in terms of strength, the main part of elastic deformation occurs in joints [3, 5]. This deformation mechanism can be employed in design for improving the dynamic performance of the machine tool structures through the damping treatment of the joints. Viscoelastic polymers provide high energy dissipation. Viscoelastically damped structures have been successfully applied in many engineering fields, particularly in the aerospace industry [6] generally employing the VE polymer in three different ways: as free-layer dampers (FLD), as constrained-layer dampers (CLD), in tuned viscoelastic dampers (TVD) [3]. FLDs are defined as those dampers composed by a single layer of damping material positioned on top of the base material by gluing or other bonding techniques as illustrated in Fig. 1a. The FLD structure’s loss factor increases with the thickness, with the storage modulus and the loss factor of the damping layer. CLDs consist of a sandwich structure where the damping material is constrained between two layers of elastic material as shown in Figure 16(b). Kerwin [7] has studied this configuration and came to the conclusion that this is far more effective than the FLDs. CLDs have also been studied in patched applications in order to optimize the usage of damping material and enhance damping only for particular vibration modes [8]. TVDs consist of a mass residing on a damping layer bonded to the base material as shown in Figure 16(c).

Design of structures using VE polymers requires proper methods to predict the overall damping values expected from various structural configurations. The loss factor and the response of the damped structure to the dynamic load are critical parameters that describe the effectiveness of the damping treatment. The loss factor is a measure of the inherent damping in a material when it is dynamically loaded. It is typically defined as the ratio of energy dissipated in unit volume per radian of oscillation to the maximum strain energy per unit volume. Loss factor damping is sometimes referred to as material or structural damping. Viscoelastic polymers are polymeric materials whose long-chain molecules cause them to convert mechanical energy into heat when they are deformed, as schematically illustrated in Fig. 2.
The most important advantage of VE polymers is their high loss factor. The low storage modulus of VE polymers allows for easily adjusting to the shape of the various structural elements. However the low storage modulus prevents design of damped structural elements entirely based on VE materials except those in free vibration surfaces.

As mentioned above, the loss factor is a measure of the energy dissipation capacity of the material while the storage modulus is a measure of the stiffness of the material. The storage modulus (shear modulus) is important in determining how much energy is transferred into the viscoelastic element, and the loss factor determines how much energy is dissipated. Both the shear modulus and loss factor of VE polymers are temperature and frequency dependent, though temperature has a greater effect on damping performance (see Fig. 3).

![Fig. 3. Shear modulus and loss factor for the 3M damping tape type 830 [9].](image)

2. DAMPED BORING BARS (DBB)

The boring bar is a tool holder used in internal turning operations. The static and dynamic stability of internal turning operation depends on the cutting parameters and the ratio length-to-diameter of the effective part of the boring bar.

![Fig. 4. Conventional, undamped boring bar.](image)

2.1. Model of the conventional undamped bar

The conventional, undamped boring bar illustrated in Fig. 4 is a steel bar with the length of the clamping area 65.5 mm, active area length 88 mm and diameter 16 mm. The resulting effective length-to-diameter ratio or overhang ratio is 5.5. Generally, exceeding 5 times length-to-diameter ratio strongly reduces the stiffness of the tool holder (see Fig. 5). Depending on work material and the cutting parameters, higher length-to-diameter ratio (>5) sets the tool holder system in heavy vibration.
Fig. 5. Stiffness dependence on overhang ratio (Sandvik Coromant).

Fig. 6. FEM simulation of the first natural frequency in Z-direction.

FEM modelling and simulation of the undamped boring bar reveals the mode shapes and natural frequencies shown in Fig. 6. The first mode is a bending mode at 1741.62 Hz. Damping mechanisms are provided by the material damping and friction damping in the clamping area.

### 2.2. Model of the DBB

The model is based on the design of the boring bar consisting in the initial, conventional bar to which rings of composite aluminium and VE damping material are attached. The relatively soft VE layers are constrained by aluminium facings and the entire assembly, pre-stressed by an axial force between 1 and 4 kN (see Fig. 7), is mounted on the clamping area of the toolholder.
The damping mechanism is enforced partly through the deformation of the individual damping rings in the bending mode of the toolholder and partly through shear deformation of the damping assembly in the torsion mode of the toolholder. Owing this design, the damping assembly has higher strength in radial direction to sustain the clamping force while damping forces act in torsion and bending. Table 1 shows the computed displacement values at the toolholder’s tip for an axial pre-stress force of 1500 N, for length-to-diameter ratios, L/D, 2 to 6, and various number of damping rings, N, in the damping assembly. The cutting force, 1000N was applied at the tool tip. The encircled values correspond to minimum displacement for each N and L/D pair.

Table 1. The displacement values calculated at the tool tip

<table>
<thead>
<tr>
<th>N</th>
<th>L/D</th>
<th>2x</th>
<th>3x</th>
<th>4x</th>
<th>6x</th>
</tr>
</thead>
<tbody>
<tr>
<td>195</td>
<td>0.024</td>
<td>0.18</td>
<td>0.33</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>0.020</td>
<td>0.14</td>
<td>0.32</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>0.022</td>
<td>0.13</td>
<td>0.18</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>0.023</td>
<td>0.24</td>
<td>0.36</td>
<td>1.23</td>
<td></td>
</tr>
</tbody>
</table>

The computation model of the damped boring bar (DBB) is accomplished in two steps.

**Step1: Computation of the motion equation of the constrained VE ring based on the analysis of an annular disk**

The equation of motion in terms of transverse deflection $w$ is as it follows (according to [3]):

$$
\left[ D \frac{\partial^4 w}{\partial r^4} + 2 A_1 \frac{\partial^3 w}{\partial r^3} + A_2 \frac{\partial^2 w}{\partial r^2} + A_3 \frac{\partial w}{\partial r} \right] = -\rho h \frac{\partial^2 w}{\partial t^2}
$$

(1)

where

$$
D = \frac{E_r h^3}{12(1-\nu^2)}
$$

(1b)
E\textsubscript{rr} is elasticity modulus in radial direction, h represents the disk thickness, ν the Poisson’s ratio, ρ is the mass density and b < r < a, b and a are the inner and outer radii, respectively, of the annular plate (see Fig. 8) and

\[ A_1 = \frac{1}{r} \left( D + r \frac{\partial D}{\partial r} \right) \]
\[ A_2 = \frac{1}{r^2} \left( -D + r(2 + \nu) \frac{\partial D}{\partial r} + r^2 \frac{\partial^2 D}{\partial r^2} \right) \]
\[ A_3 = \frac{1}{r^3} \left( D - r \frac{\partial D}{\partial r} + r^2 \nu \frac{\partial^2 D}{\partial r^2} \right) \]

Following the procedure described in [3], Eq. (1) is written as a system of two equations

\[ D \frac{\partial^4 w_0}{\partial r^4} + 2A_1 \frac{\partial^3 w_0}{\partial r^3} + A_2 \frac{\partial^2 w_0}{\partial r^2} + A_3 \frac{\partial w_0}{\partial r} - \omega^2 \rho hw_0 = 0 \]

and

\[ \frac{d^2 \Psi}{dt^2} + \omega^2 \Psi = 0 \]

where \( \Psi \) is a time function. From (2) the expressions for maximum strain energy, \( V_{\text{max}} \), and maximum kinetic energy, \( K_{\text{max}} \), can be determined in the form

\[ V_{\text{max}} = \frac{1}{2} \int_0^{2\pi} \int_0^a \left[ D \left( \frac{\partial^2 w_0}{\partial r^2} \right)^2 + 2 \frac{D}{r} \frac{\partial^2 w_0}{\partial r \partial \theta} B_1 + \frac{D}{r^2} B_2 \left( \frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial w_0}{\partial \theta} \right) \right)^2 \right] r dr d\theta \]
\[ K_{\text{max}} = \frac{1}{2} \omega^2 \int_0^1 \int_a^b \rho hw_0^2 r dr d\theta \] (4)

Where

\[ B_i = \left( \frac{\partial w_0}{\partial r} + \frac{1}{r} \frac{\partial^2 w_0}{\partial \theta^2} \right) \]

and

\[ D_{r\theta} = \frac{Gh^3}{12} \]

is the shear rigidity, \( G \) is shear modulus, and \( D_1 = \nu D \).

Using the above procedure, characteristics of the VE constrained ring have been determined and the effect of the VE parameters on the dynamic behaviour has been analysed. The computation approach has been implemented in a FE-program, where the main part concerns VE-material behaviour evaluation. Viscoelastic materials have a time-dependent response even if the loading is constant. Linear viscoelasticity is adopted in all calculations and the stress is considering varying linearly with strain and strain rate. It is also assumed that the viscous part of the deformation is incompressible, so that the volume change is purely elastic. Fig. 9 shows the storage shear modulus and loss factors values as frequency functions.

![Fig. 9. Loss modulus and storage modulus of VE polymer.](image)

**Step 2: Computation of the flexural and torsional motion of the DBB**

The energy introduced in the elastic structure can be very large during internal turning where the ratio length-to-diameter is large. Therefore, it is important to design the damping assembly with an optimal number of damping rings and the corresponding pre-stress. As described above, in the model, the composite plates, formed by a VE-material constrained between two Al-plates are mounted on the boring bar. The damping effectiveness of VE constrained layers is related to strain, thus the behaviour of the strain energy has been studied as well. In Table 2 the strain energy density and the total power dissipation are presented corresponding to a damping assembly with 120 rings.
Table 2. Strain energy and power dissipation for the damping assembly with N = 120 damping rings

<table>
<thead>
<tr>
<th>Eigenfrequency</th>
<th>Strain energy density VE120 (J)</th>
<th>Strain energy density ALL (J)</th>
<th>Total power dissipation density ALL (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1349.62+0.696i</td>
<td>1.3750e7-1.27321e5i</td>
<td>1.333e8+27113.74326i</td>
<td>-5.666e8</td>
</tr>
<tr>
<td>1354.48+0.622i</td>
<td>1.3870e7-1.47194e5i</td>
<td>1.389e8+40525.9054i</td>
<td>-5.962e8</td>
</tr>
<tr>
<td>3865.45+0.112i</td>
<td>9.502e5-68399.55496i</td>
<td>9.307e8+1.08225e8i</td>
<td>-7.009e8</td>
</tr>
<tr>
<td>4255.53+153.143i</td>
<td>1.504e9-4.93704e7i</td>
<td>2.480e9+8.98078e7i</td>
<td>-2.406e12</td>
</tr>
<tr>
<td>6349.31+11.042i</td>
<td>1.945e8-8.25921e7i</td>
<td>4.843e9-1.84689e9i</td>
<td>-3.604e11</td>
</tr>
<tr>
<td>6687.58+4.688i</td>
<td>8.247e7-5.02059e6i</td>
<td>5.625e9+1.59583e8i</td>
<td>-1.658e11</td>
</tr>
</tbody>
</table>

3. EXPERIMENTAL WORK

The purpose of experimental work was to study the dynamic behaviour of the DBB in internal turning and to assess the performance of the DBB in terms of quality of the machined surfaces and the tool life extension. Machining tests were performed on a lathe type Haas ST30 with a maximum spindle speed 3000 rpm and 30 KW motor power.

![Accelerometer](image)

Fig. 10. Set-up for internal turning showing the work, DBB overhang and sensor system.

Workpiece with external diameter D = 130 mm and initial inner diameter, d = 30 mm were clamped in the chuck and machined at constant rotational speed, n = 1800 rpm. The work material was steel with 300 HB hardness. The DBB vibration during machining was measured by a three-component accelerometer (Dytran 3023A2) located on the toolholder (see Fig. 10) and recorded by an LMS TestLab® system.

The DBB geometry and design are described in section 2. Inserts type Mircona SNMG120404-NM1 were used for all experiments. The performance of DBB was assessed by comparing with machining using conventional tools in similar cutting conditions.
Fig. 11. Comparison of vibration measured by accelerometer for CT (a) and DBB (b).

Fig. 11 shows the comparison of the power spectral density of the signal measured by the accelerometer in the direction of the cutting speed when machining at a depth of cut \(a_p\) of 2 mm. The vibration amplitude generated by the conventional tool (CT) is extremely higher than the one generated by the DBB.

4. RESULTS AND DISCUSSIONS

The machining tests have revealed that, in the specific case, the DBB was able to machine in stable cutting conditions at a depth of cut three times higher than the conventional boring bar. This result is confirmed also by the stability limit diagram (SLD) computed for DBB and conventional boring bar [10] (see Fig. 12).

Fig. 12. SLD comparing the conventional boring bar (red) and the DBB (Blue). (after [10])
This performance enhancement has a direct repercussion on the surface finish of the produced parts. In fact the DBB has been capable of producing finer surfaces at all the tested cutting parameters, see for instance Fig. 13.

![Fig. 13. (a) Surface finish employing DBB in internal turning. Machining time is 28 min. (b) Surface finish using conventional boring bar.](image)

In Fig. 14 the microscopic pictures of the tool when machining with a conventional boring bar is shown. The used cutting data are $ap=2\text{mm}$, $f=0.25\text{mm/rev}$ and $vc=180\text{m/min}$. The measured flank wear (in Fig. 14(a)) is 0.11mm. Fig. 14 (b) and (c) show the corresponding rake face and the nose area of the same insert. Traces of adhering workpiece material can be seen both on the flank side of the insert (see Fig. 14(a)) and on the rake face of the tool (see Fig. 14(b)). The same observation can be made when studying the nose area of the insert (see Fig. 14(c)).

![Fig. 14. The tool after machining when using a conventional boring bar seen through a microscope with 200 times magnification. Cutting data are $ap=2\text{mm}$, $f=0.25\text{mm/rev}$ and $vc=180\text{m/min}$. (a) The clearance side of the tool. (b) The rake face of the tool. (c) The nose area.](image)

![Fig. 15. The tool after machining when using a DBB seen through a microscope with 200 times magnification. Cutting data are $ap=2\text{mm}$, $f=0.25\text{mm/rev}$ and $vc=180\text{m/min}$. (a) The clearance side of the tool. Machining time 16min. (b) The clearance side of the tool. Machining time 28min. (c) The clearance side of the tool. Machining time 84min.](image)
In Fig. 15 three pictures from the flank side of three different tools are seen. In all three cases the same set of cutting data has been used, \( a_p = 2\, \text{mm} \), \( f = 0.25\, \text{mm/rev} \) and \( v_c = 180\, \text{m/min} \). Fig. 15 (a) shows the flank wear after 16 minutes of machining, (b) the flank wear after 28 minutes of machining and (c) the flank wear after 84 minutes of machining. The measured flank wear is 0.10mm for tool (a), 0.12mm for tool (b) and 0.15mm for tool (c). The trend with increasing flank wear after longer machining time is what can be expected. In Fig. 16 the mean value of the estimated flank wear is 0.10mm. In both pictures (Fig. 16 (a) and (b)) workpiece material is sticking to the insert.

![Fig. 16. The tool after machining when using a DBB seen through a microscope with 200 times magnification. Cutting data are \( a_p = 2\, \text{mm} \), \( f = 0.25\, \text{mm/rev} \) and \( v_c = 180\, \text{m/min} \). (a) The clearance side of the tool. (b) The nose area of the same insert. Machining time 16min.](image)

Fig. 17. Results from surface finish measurements on the workpiece after machining 4 minutes using a DBB. (a) 3D-plot of test area (b) Profile Plot. (c) Table with values for \( R_z \), \( R_a \) and \( R_q \).
Evaluating the surface profile after machining 4 minutes using a DBB (see Fig. 17 (c)) gives a mean value for the ten-point height $R_z=12.089\mu m$, an arithmetic mean value $R_a=2.38\mu m$ and a quadratic mean $R_q=2.897\mu m$. The profile and 3D plot (see Fig. 17(a)(b)) show a surface with the same characteristics. The conventional boring bar was not able to perform in a stable manner at the tested depth of cut.

5. CONCLUSIONS

The paper presents the theoretical and experimental results concerning the dynamic performance of the DBBs, characterization of the machined surfaces and tool life improvements.

The DBB is able to perform stable cutting at settings of cutting parameter that would be prohibitive for a conventional tool. An important objective with this study was to assess the effect of dynamic behaviour on the insert wear and on wear mechanisms. As a result of the experimental tests the inserts mounted on the DBB showed a tool life up to 90 minutes at the tested cutting parameters.

From the industrial application point of view, the presented approach allows the end user to select the most suitable parameters in terms of productivity that extends over the range allowed for conventional tooling systems.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Mircona AB for supporting with laboratory equipment and tools for carrying out the experimental work. This paper is one of the results of the FFI Robust Machining project, financed through VINNOVA, the Swedish Governmental Agency for Innovation Systems.

REFERENCES


Steps Towards Holistic Modelling of HSC Machining Centre Errors

Jerzy Jedrzejewski, Wojciech Kwasny

Wroclaw University of Technology, Poland, jerzy.jedrzejewski@pwr.wroc.pl

ABSTRACT

This paper sets out the reasons for creating holistic models of machining units, especially the ones intended for high-speed precision machining. The structure of a holistic model is described, mainly for a 5-axis high-speed machining centre. The paper focuses on modelling the errors generated in the controllable axes. The verification of the model for the spindle axis position error in the Z axis in the central part of the workspace is presented.

KEYWORDS: holistic model, machining centre, error

1. INTRODUCTION

The development of the modelling of machine tool static, thermal and dynamic properties is closely connected with the development of the modelling of the physical phenomena which occur in the drive units of the particular controllable axes, and with advances in computers and numerical modelling [1],[9]. Modelling must keep up with technological progress which brings novel constructional materials ever more suitable for achieving an ever higher speed of motion, reducing energy consumption, extending life time, reducing the noise level as well as, which is particularly important, achieving low deformability of the particular assemblies and the load-bearing structures. This inevitably leads to high precision of motions, having a direct bearing on workpiece accuracy and machining productivity.

Modern High Speed Cutting (HSC) machine tools are characterized by very high motion dynamics. Their spindles reach speeds as high as 50 000 rpm (and higher), accelerations of 2.5 G, and bearing speed coefficients \( n x d_m = 5 \times 10^6 \). The ball screws reach feed rates as high as 50 m/min and accelerations of 2G. The direct drives reach speeds of 80 m/min (and even 120 m/min) and accelerations of 4G. In micromachining centres the spindle aerostatic bearings reach rotational speeds above 100 000 rpm. In the case of high speed cutting, the heating up of the drives (motors, bearings, nuts and ball screws) must be reduced and their temperature needs to be controlled. At such high dynamics of the drives, in order to improve the operating properties of the latter the modelling of the loads (including the thermal ones) and the deformations and displacements of the executing units requires very faithful representation of the physical phenomena involved and their interactions in real time. Highly accurate holistic modelling is needed for purpose. Some necessary model simplifications should not affect the
accuracy with which errors and error over-time variation in the machine tool operating conditions are determined.

Most of the models described in the literature are very simplified and do not represent the naturally generated power losses, deformations and interactions and consequently, the operating properties of machine tools. Analyses based on such models do not lead to any marked improvement in the machine tool operating properties. Therefore holistic modelling should be used, intensively developed and improved. Through holistic modelling one can effectively implement a machine tool improvement strategy.

This paper presents the results of research on holistic modelling, conducted under the authors’ direction in Wroclaw University of Technology [7].

2. HOLISTIC MODEL STRUCTURES OF MACHINE TOOLS

Holistic modelling is based mainly on the creation of accurate models of the behaviour of the main drive units in the machine tool controllable axes in the operating conditions (Fig. 1).

Fig.1. Computing model of machine tool.
The modelling of the units involves optimal design and continuous improvement by teams of specialists, in many cases independent of the machine tool manufacturers.

By integrating the models of the units into whole machine tool computing systems (developed by the authors and the commercial ones) one can determine the total errors which in the machining conditions find reflection in the workpiece, depending on the position of the tool in the machine tool workspace. In order to verify the integrated holistic models one needs to measure spatial errors, including the thermal ones, in both thermally stationary and nonstationary states [8]. Because of their complexity and the limited computing power of the available computers, holistic model computations are time-consuming. Therefore special procedures/functions, which take into account the correlation between the total error and the real time parameters, need to be developed in order to eliminate (through error compensation) the effect of errors on machining accuracy (Fig. 1).

Fig. 2. Holistic model of: a) – 5-axis machining centre with internal loads, b) – high-speed and high-precision machining system with internal loads.
Limiting the holistic model to the machine tool alone, one must integrate the predominant thermal errors and the errors generated by the dynamic loads (especially the spindle shift) into this model (Fig. 2a).

The total error mapped on the workpiece takes into account also the interactions between the machining tool and the machining process (Fig. 2b). The machining process is usually optimized through complex actions aimed at increasing, among other things, machining accuracy, e.g., by controlling the acceleration and the rate of acceleration of the work motions. This has a significant bearing on the power losses generated in the machine tool drives and consequently, on the workpiece error.

Workpiece position fixing and workpiece supporting and clamping, even if intelligently performed, are a source of direct machining errors. Errors generated while machining flexible aircraft engine parts are minimized through intelligent functions performed by the fixture, and simultaneously compensated by a CNC system. The authors with their team are currently working (within research projects) on modelling and minimizing such errors.

3. COMPLEXITY OF HOLISTIC MODELS

3.1. High-speed axis error modelling

The creation of holistic models is oriented towards achieving and aiding high productivity of the process of manufacturing products/parts, including achieving the required accuracy. The models are used to precisely identify and optimize the properties of an analyzed machining module with regard to all its main components and functions, illustrated in Fig. 2b and comprehensively presented in Fig. 2b. A holistic model of such a unit is highly complex and because of the nature of the unit’s structure, it is usually hybrid.

Modelling covers all the mechatronic devices and the software control of the process, including the duty cycles of the particular units, but it focuses on the precision of the machined parts. Modelling considerations can begin with the recognition of errors generated in the machine tool controllable axes, and their interactions. Then the sources of the errors generated beyond the controllable axes can be considered.

In holistic error modelling and minimization it is critically important to accurately mathematically describe:
- the generation of power losses in heat sources in machine tool operating conditions, and their variation,
- the heat transfer in joins and closed spaces and to the environment,
- the autogenous flow of heat in the sources,
- the dynamic excitation of loads, deformations and disturbances (variations in rotational speeds, feed rates and acceleration), and the corrective and self-repairing actions.

In the direct drives of tools and workpieces, the power losses in the rolling bearings and in the motors must be particularly precisely modelled. An example of such modelling of power losses in the bearings of a spindle is shown in Fig. 3. In its bottom part, the heat flow and the variation in power losses over time are presented.

A hybrid model, combining the finite difference method and the final element method, works well in the precise modelling of such friction nodes [4]. Besides the bearings, the model must also accurately represent the power losses in the motor of the electrospindle.

Since the heat generated in both the bearings and the motor must be intensively removed, an accurate cooling model is needed. If with regard to the permissible thermal deformations motor cooling is inadequate, the model must take into account the introduction of cooling of, e.g., the bodies walls, the spindles, the ball nuts and screws and the guides.
Fig. 3. Interrelations between factors affecting power losses in spindle bearing units.

Selected factors which such a model takes into account are shown in Fig. 4 [10]. In more detail they are discussed in [6].

The spindle’s thermal deformation and displacement and its deformation and displacement (shift marked with arrows in Fig. 5) due to centrifugal forces add up.

Element “Fluid1”

$$\sum_{i=1}^{\sigma} Q'_{S-C} + Q_{F} + Q_{A} + Q_{U} = 0$$

i – number of surface element with fluid layer

Element type “Air”

$$\sum_{j=h+1}^{\sigma} Q'_{S-P} + Q_{P-C} + Q_{A} + Q_{U} = 0$$

j – number of surface element without fluid layer

Fig. 4. Model of heat exchange in cooler and inside headstock.
First, shift appears and quickly increases with rotational speed over acceleration-dependent time. Then the displacements assume the character of a thermal curve. Each next change in rotational speed is followed by a rapid change in the axial displacement of the spindle by the shift value. The increment in shift over time and its dependence on the rotational speed of the spindle tested by the authors for selected speeds are shown in Fig. 6. The percentage of shift in the total spindle displacement may be significant, especially at high rotational speeds. For a known stiffness of the spindle bearing unit tension springs the percentage can be determined using a mathematical model while the spindle dimensional changes caused by centrifugal forces can be determined using the final element method [5].

Fig. 5. Rate of spindle displacement after change of rotational speed.

Fig. 6. Changes in spindle shift as function of rotational speed.
3.2. Feed axis error modelling

The positioning error in the feed axis depends on the drive design [2]. In a drive with a ball screw (Fig. 7a) this error is considerably affected by the elastic displacements occurring in the ball screw itself and in the bearings which support it. The bearings carry their own preload force, screw prestretching force $R$ and pull force $F_{\text{pull}}$.

Under the action of screw stretching force $R$, elastic displacements $\delta_{A_{\text{max}}}$ and $\delta_{B_{\text{max}}}$ appear in the cold bearings, commensurate with their axial stiffness (Fig. 7b). In operating conditions, stretching force $R$ decreases as the temperature of the ball screw increases. This results in, e.g., a reduction in losses in the preloaded bearing unit (Fig. 7c) from $Q_1=116\,\text{W}$ in the cold state ($R=7200\,\text{N}$) to $45\,\text{W}$ in the heated state after the stretching force has dwindled ($R=0$).

The ball screw heats up due to power losses in the bearings of the two supports and power losses in the moving heat source, i.e. the preloaded ball screw/nut unit. The magnitude of the heat fluxes heating up the ball screw depends on the autogenous flow of the heat generated in the ball nut and in the bearings, the heat capacity of the housings and the ball screw, and the forced cooling conditions.

![Fig. 7. Ball screw behaviours.](image-url)
The pre-tensioned and axially loaded ball screw behaves like a pair of preloaded angular bearings. Then as a result of the deformation at the ball/race contact the generated heat flux changes commensurately with the increase in the internal forces in one of the nut parts and with the reduction in the forces in the other part.

Besides the pre-tension, the value of the carried axial force, determined by the motion resistances, the shifted masses, the drive operation dynamics (acceleration) during startup, the rotational speed during steady motion (with and without load) and the deceleration during braking, has a significant influence on the size of the generated power losses. An exemplary diagram of power losses for such a duty cycle is shown in Fig. 7d.

When modelling the behaviour of the drive during machining one should also take into account the dependences between the drive unit stiffness and the direction in which the external load acts, especially when the ball screw is unilaterally fixed (Fig. 7e). Then the axial stiffness in one of the extreme table positions may be very small.

Only when all the factors determining the generation of power losses and loads, and their variation in machining conditions are taken into account, one can reliably determine positioning errors in the controllable feed axis.

### 3.3. Additional rotation axes

Additional rotation axes, such as tilting tables, increase machine tool flexibility but at the same time constitute an additional source of linear and angular errors. These errors are the sum of geometric, kinematic and thermal errors.

Thanks to the direct drives commonly used in such units very high torques can be achieved, which is the main advantage of the drives. But then large amounts of heat generated in the stator and rotor windings must be removed. Another significant source of heat in controllable axes A and C are rolling bearings, roller axial/radial bearings or cross-roller rings.

A model of the thermal behaviour of tilting tables with a direct drive should comprise:
- models of heat generation in bearings and motors,
- a bearing stiffness model,
- a cooling model.

![Algorithm for calculating power losses in torque motors and bearings](image)

Fig. 8. Algorithm for calculating power losses in torque motors and bearings, as well as modelling of heat dissipation by a cooler.
Power losses in motors should be determined taking into account the tilting table duty cycle which includes the start-up phase, the continuous running phase and the braking phase. The maximum values of power $P_c$ dissipated in the motor, the peak torque values and the torque values during continuous running are specified in the manufacturer catalogues for the motor winding temperature of $130^\circ$C. The actual value of this temperature is influenced by the motor duty cycle, the motor housing structure, the size of the heat exchanging surfaces and especially by the cooling system.

In order to determine the proper value of power losses in a motor one must assume initial winding temperature $\Theta$ during motor running for the assumed motor housing and cooling conditions and then iteratively search for the actual temperature by means of the FEM model [3]. The starting point for the iterations should be the power value calculated from the catalogue data and the assumed duty cycle. An algorithm for this iterative process, elaborated by authors, is shown in Fig. 8.

After the actual operating temperature is calculated using the FEM program, the motor power loss values and the power losses in the bearings, which are also a function of temperature, are corrected. The iteration process is continued until the temperature assumed for calculating the motor power losses comes level with the winding temperature obtained from the FEM model simulations.

### 3.4. Example of machining centre calculations by means of holistic model

The spindle unit is the largest source of errors, as confirmed by the experimentally verified calculation results shown in Fig. 9.

For the spindle rotational speed of 50 000 rpm the total error determined by means of the holistic model amounted to $113\mu$m in axis Z. The total error was made up of the spindle unit error amounting to $68\mu$m (including the load-bearing structure deformations) and a shift of $40\mu$m, being a sum of the deflection of the slidable sleeve springs and the dimensional changes of the spindle caused by the action of centrifugal forces and the thermal displacement of the tilting table, amounting to merely $5\mu$m. As the experimentally determined total error diagram shows, the measurement results very well agree with the results of the holistic model calculations.

![Fig. 9. Results of thermal deformation 5 axis machining centre and spindle shift calculations for 50.000 rpm (a), experimental verification of total spindle displacement relative to table in thermally steady states against displacements for other spindle rotational speeds (b).](image-url)
4. CONCLUSIONS

The holistic modelling presented above needs to be improved by improving its many constitutive models. At this stage the most advanced is the modelling of spindle units in idle running conditions. Because of the lack of sufficiently accurate models of the machining process, taking the loads generated by the cutting forces into account still poses difficulties.

The next stage in this research will be devoted precisely to taking the machining tool/machining conditions interactions into account.

ACKNOWLEDGEMENTS

This research is supported by the Ministry of Science and Education of Poland and the Ministry of Knowledge Economy – the Republic of Korea, and NCBiR (project: INNOTECH – K1/IN1/75/155671/NCBR/13).

REFERENCES

Theme 10

New research endeavour

- Advanced cutting materials for hard turning
  Ivan Baranek
  Slovak University of Technology in Bratislava, Faculty of Materials Science and Technologies in Trnava, Slovak Republic

- Wear and renewal of cutting tools properties
  Ivan Baranek
  Slovak University of Technology in Bratislava, Faculty of Materials Science and Technologies in Trnava, Slovak Republic

- The reverse engineering to optimize the assembly of a conical spur gear by CAD
  F. Belarifi\(^1\), E. Bayraktar\(^2\)
  \(^1\)I.M.S.I/ L.S.M.C- Oran University, Algeria,
  \(^2\)SUPMECA (Paris) -Saint- Ouen, France
Advanced cutting materials for hard turning

Ivan Baranek

Slovak University of Technology in Bratislava,
Faculty of Materials Science and Technologies in Trnava, Slovak Republic
Contact e-mail: ivan.baranek@stuba.sk

ABSTRACT
The paper is focused on the analysis of the modern metal cutting. Among the most important trends of machining development nowadays belong high speed cutting, dry machining and hard machining. The aim was to cover cutting materials for hard turning. There are described chosen results of experiments focused on the tool life and the surface integrity. Based on the experiments there was investigated the tool life of various kinds of PCBN by hard turning. Taylor equations can be used for the cutting conditions setting. The measured values of surface roughness, residual stresses and micro-hardness show possibility of hard turning usage in practice. In case of the white layer zone appearance this needs to be eliminated by superfinishing.

KEYWORDS: Cutting materials, Tool life, Surface integrity

1. INTRODUCTION

The search for new materials with improved mechanical properties led to development of materials that are harder, more abrasive and hence more difficult to machine.

These special materials call for suitable cutting tool materials and appropriate machining conditions. To machine these very hard-wearing materials, manufacturers offer a comprehensive package of new cutting tool materials and tool-holders, together with the necessary technical service.

Among the most important trends of machining development nowadays belong high speed cutting, dry machining and hard machining [1].

Turning steel with a hardness over 45 HRc (typically within the range of 55-68 HRc) is defined as hard part turning and is a cost efficient alternative to grinding. Hard part turning has been proven to reduce machining time and costs by 70% or more, and also offers improved flexibility, better lead times and higher quality [2].

Hard part turning is a well accepted method, especially in the automotive industry. Typical components include: gear box housings, brake rotors, transmission gears, steering pinions, valve seats, engine blocks, pistons, cylinder liners and clutch housings [2].
2. **CHOICE OF CUTTING MATERIALS**

The main cutting materials used in metal–cutting production are as follows:

- High speed steel
- Hard metal
- Ceramics
- Polycrystalline diamond and polycrystalline cubic boron nitride.

One of the basic requirements for a cutting material is that it should be harder than the material of the work-piece. Carbide, ceramics and polycrystalline cubic boron nitride are mainly interesting from this point of view.

![Cutting tool materials for hard turning.](image1)

Carbide is not recommended when the hardness is above 50 HRc. Ceramics can be used between approx. 50-60 HRc when surface finish demands are moderate. Cubic Boron Nitride grades (CBN) are the ultimate cutting tool materials for hard part turning. However, they should not be used on steels softer than approx. 48 HRc [2].

![Example of a part for hard turning.](image2)
Work-pieces treated by us had a nominal hardness value of 62 ± 2 HRc. Therefore we focused more on investigation of CBN. Example of this machine part is on Fig.2.

3. TOOL LIFE TESTING OF PCBN TURNING TOOLS

3.1 Polycrystalline cubic boron nitride

Cubic boron nitride (CBN) is synthesised under ultra-high pressure /5-6 GPa/ at temperatures of 1500-1600 °C, similar to the conditions for man–made diamonds. The starting material in the low–pressure formed of boron nitride, which has a hexagonal crystal structure and properties similar to graphite.

Cubic boron nitride (CBN) is nearly as hard as diamond, its toughness being superior to that of ceramic cutting tools. CBN is high thermal stable up to 1300°C and resistant to a chemical attack, particularly resistant against ferrous materials.

Main component of polycrystalline cubic boron nitride (PCBN) is cubic boron nitride (CBN), which is mixed with ceramic and metallic binder materials and sintered during a high-pressure process.

The cutting edges are sintered under ultra high pressure and are supplied on hard metal supports or as solid indexable inserts.

Solid PCBN inserts provide multiple cutting edges on two sides. PCBN inserts also come in full-face and tipped types.

The full-face type has a complete PCBN face sintered onto a carbide substrate, and provides multiple cutting edges on one side only. These inserts are less expensive than solid PCBN inserts.

The tipped style contains a small PCBN segment brazed onto one corner of a carbide insert, providing either a single cutting edge or double cutting edges.

The PCBN inserts come in industry-standard sizes and can be used in the insert pockets of standard tool-holders and milling cutters.

Most PCBN inserts used today are tipped. Tipped PCBN inserts are economical and reliable for a wide range of roughing and finishing applications, but some applications require a solid or full-face insert.

Insert edge preparation strongly influences the success of PCBN machining. A T-land about 0.2 mm by 20° is good for roughing, and light hone is appropriate for finishing.

The properties of the micro-geometry combined with an optimum ratio of CBN contents to bonding phase provide for an enormous performance potential and have a considerable influence on the results of machining [4]. Coated PCBN are also used [5, 6].

Although the prices are similar to polycrystalline diamond /about 15-20 times higher than a comparable hard metal or ceramic cutting edge/, CBN tools are increasingly used for machining hard ferrous materials with a hardness between 45 and 65 HRc, such as fully hardened cold–work tool steels or high speed steels, chilled cast iron, white iron, and martensitic cast iron with high contents of nickel or chromium.

Machining processes are:
• Turning, even with interrupted cutting
• Boring
• Milling.

Typical examples of applications of PCBN tools are:
• Machining of rolls and discs
3.2 Test conditions

Cold work steels STN 19436 (1.8 – 2.05%C, 0.2 – 0.45% Mn, 0.2 – 0.45% Si, 11 - 12% Cr) were selected as work – piece materials for the machining study. Work-piece was supplied in bar form – 90 mm diameter by 500 mm long – and heat treated to a nominal hardness value of 62 ± 2 HRc. The minimum diameter to which the bars were machined was such that the length to diameter ratio never exceeded 10.

Cutting tool materials, tool holders and cutting conditions are presented in Table 1. An 11 kW SU 50 A lathe, fitted with an infinitely variable spindle speed drive, was used for the cutting tests which were all conducted under dry cutting conditions.

Tool life data points were determined for the various cutting conditions (ISO 3685). For the purpose of these tests there was taken the flank wear land VBc with value 0.4 mm. The test was also interrupted while chipping of the cutting edge.

3.3 Experimental results

On the basics of tool life testing PCBN and cemented carbides /Table 1/ the effectiveness of the work with these materials were judged. Machining with PCBN is multiple cheaper then machining with carbide tools (Fig. 3).

Fig. 4 shows the Microstructure of PCBN „B“ and Fig. 5 shows the wear of PCBN „B“ cutting edge .

![Fig. 3. Machining costs for various cutting materials in % (K10=100%).](image-url)
Table 1. Experimental results.

<table>
<thead>
<tr>
<th>Cutting materials</th>
<th>Geometry of the active part of cutting tools</th>
<th>Cutting conditions</th>
<th>Taylor equations</th>
<th>Roughness of the surface $R_a$ [μm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool holders</td>
<td>$\gamma_0$ $\alpha_0$ $\kappa_r$ $\varepsilon_r$ $r_c$ $\lambda_s$ $b_\gamma$</td>
<td>$v_c = (40-80)\text{m.min}^{-1}$</td>
<td>$T= \frac{37358}{v_c^{2.18} \cdot f^{0.38} \cdot a_p^{0.32}}$</td>
<td>0.9 - 6</td>
</tr>
<tr>
<td>PCBN „A“</td>
<td>- 10 30 135 0.5 6 -</td>
<td>$f = (0.04-0.16)$ mm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(25x16) mm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN „B“</td>
<td>-6 6 75 90 1.6 6 0.2</td>
<td>$v_c = (40-80)\text{m.min}^{-1}$</td>
<td>$T= \frac{620487}{v_c^{2.7} \cdot f^{0.51} \cdot a_p^{0.32}}$</td>
<td>0.5 – 1.6</td>
</tr>
<tr>
<td>SNMN 12 0316 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(CSBNR 2525 M123)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cem. carbide (HW K10)</td>
<td>-6 6 70 90 1.6 6 -</td>
<td>$v_c = (10-20)\text{m.min}^{-1}$</td>
<td>$T= \frac{35739}{v_c^{3.06}}$</td>
<td>0.8 – 1.8</td>
</tr>
<tr>
<td>SNMN 120416</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAREX – 223850.1</td>
<td>$\gamma_I = -6$, $\gamma_P = -6$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(25x25) mm</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN „C“</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>157/II.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNMMN 090300</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLRNR 2020 K 09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN – D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNMMN 130300F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRGNR 2525 L13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_I = -6$, $\gamma_P = -6$</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN „A“</td>
<td>PCBN segment brazed onto tool</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN „B“</td>
<td>Solid PCBN insert</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN „C“</td>
<td>Solid PCBN insert</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCBN „D“</td>
<td>Full-face type</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 4. Microstructure of PCBN „B“.

Fig. 5. Wear of PCBN „B“ cutting edge.
4. SURFACE INTEGRITY BY HARD TURNING

4.1 Surface integrity

Properties and performance of all manufactured parts are related to the quality-or integrity-of the part’s surfaces. The surface integrity is affected by manufacturing processes employed.

The surface finish influences not only the dimensional accuracy of machined parts, but also their properties. Whereas surface finish describes the geometric features of surfaces, surface integrity pertains to properties such as fatigue life and corrosion resistance, which are influenced strongly by the type of surface produced.

Factors influencing surface integrity are temperatures generated during processing, residual stresses, metallurgical (phase) transformations, and surface plastic deformation, tearing, and cracking.

Surface integrity describes not only topological - geometric - aspects of surfaces, but also their mechanical and metallurgical properties and characteristics [7].

The quality of machined surface is characteristic by:

• surface texture
• directionality, roughness, waviness
• hardness
• residual stresses
• the changes of structure
• chemical changes in surface layer, etc.

The amorphous layer on work-piece caused by grinding is also found with hard turning in a similar form. It is called a white layer.

The white layer zone is a structural transformation on work-piece surface. The high pressure between cutting edge and work-piece causes a rise in temperature at the rim. Due to the rapid cooling down after cutting, renewed hardening of work-piece surface occurs. Additionally, there is growing internal stress in tempered zone [4].

This very thin layer must be removed in many cases through a suitable manufacturing process before the work-piece can be put into service (e.g. by honing or superfinishing [8, 9]. One of the most commonly used techniques for testing surface integrity is metallography.

Destructive and none destructive techniques are used to observe and test surfaces.

4.2 Experimental results

With new cutting materials the surface integrity was researched by our experiments, too. The roughness of the surface after hard turning is shown in Table 1. Such roughness we can consider to be good. Other possibilities bring modifications of geometry in the form of Wiper geometry. Wiper inserts are capable of turning at high feed rates - without losing the capability for generating good surface finishes [11].

The ceramic and CBN wiper inserts are the object of further research also at our department [12, 13].
Fig. 6. The machined surface morphology by turning with PCBN „B“ ($v_c = 80\text{m.min}^{-1}$, $f = 0.08$ mm, $a_p = 0.25$ mm).

Fig. 7. Microstructure and places of microhardness measurement by turning with PCBN „B“ ($v_c = 23.9\text{m.min}^{-1}$, $f = 0.109$ mm, $a_p = 0.5$ mm).

The machined surface morphology by turning with PCBN „B“ is on Fig. 6. We have not found the white layer on machined surface after turning with PCBN „B“ (Fig. 7).

Right under the surface the residual stresses were pressure like. However during the experiments by hard turning with mixed ceramic we found the white layer zone on workpiece surface (Fig. 8).

Fig. 8. Residual Stresses after hardening / tempering (A25) and turning with PCBN „A“ (A13) and PCBN „B“ (A2).
Fig. 9. Microstructure of surface layer after dry hard turning of steel STN 19221.4 with mixed ceramic (CC6050). [10]

5. SUMMARY

Based on the experiments there was investigated the tool life of various kinds of PCBN by hard turning. Taylor equations can be used for the cutting conditions setting. The measured values of surface roughness, residual stresses and microhardness show possibility of hard turning usage in practice.

In case of the white layer zone appearance this needs to be eliminated by superfinishing.

ACKNOWLEDGEMENTS

This paper was elaborated within the project VEGA 1/0615/12 Influence of 5-axis grinding parameters on shank cutter’s geometric accuracy.

REFERENCES

Wear and renewal of cutting tools properties

Ivan Baranek, Ivan Buransky, Stefan Vaclav

Slovak University of Technology in Bratislava,
Materials Science and Technology in Trnava,
Institute of Production Technologies
ivan.baranek@stuba.sk

ABSTRACT

The article presents resharpening of tools by CNC Cutting Tool Grinding Machine WZS 60 REINECKER in The Centre of Excellence of 5 – axis Machining (CE5AM) at the Faculty of Materials Science and Technology in Trnava. Utilizing of the programming system NUMROTO allows quick and cost effective production and regrinding of precision cutting tools. The testing and inspection of tools has become increasingly important from quality and process reliability point of view. There is used the equipment Zoller Genius 3s in the CE5AM. The wear of resharpened mill is more than 4 times higher than the wear of the new tool. The reason is that there is no coating on the face of the mill after resharpening. The coating increases the wear resistance of the cutting edge. It is necessary to do the complete tool reconditioning including recoating.

KEYWORDS: wear, production, renewal, cutting tools

1. INTRODUCTION

The efficiency and quality play the important role in the today engineering production. These production characteristics are connected with words like wear and tool life when producing machine parts. I practice it is easy to find out that the cutting tool is worn by surface texture, dimensional accuracy, ability to control chips satisfactorily etc.

The cutting tools are subjected to high localized stresses, high temperatures, sliding of the chip along the rake face and sliding of the tool flank along freshly cut surface. The combination of loading factors on the cutting edge induces the tool wear.

The basic wear mechanisms are abrasion, adhesion, diffusion, oxidation, static or dynamic fatigue and plastic deformation [1].

The rate of wear depends on a tool material, workpiece material, geometry of active part of tool, cutting environment, machine-tool characteristics, cutting conditions.

In practice, the useful tool life is the period during which the cutting tool is capable to produce acceptable workpieces of the required dimensional accuracy and the desired surface characteristics.

The worn cutting tool has to be resharpened.
The article presents resharpening tools by CNC Cutting Tool Grinding Machine WZS 60 REINECKER in The Centre of Excellence of 5-axis Machining (CE5AM) at the Faculty of Materials Science and Technology in Trnava [2]. This grinder is built on a polymer concrete base, which is characterized by high rigidity and excellent dampening. Utilizing the programming system NUMROTO allows quick and cost effective production and regrinding of high precision cutting tools.

The testing and inspection of tools has become increasingly important from quality and process reliability point of view. There is used the equipment Zoller Genius 3 in CE5AM. This equipment also supports:
- production of new tools
- resharpening

Therefore it allows cooperation with CNC grinding machines.

2. DESIGN AND RESHARPENING OF TOOL

2.1. Grinding machine

Reinecker WZS 60 is tool and cutter grinder. Utilizing of the new clamping system and the programming system NumrotoPlus® in conjunction with linear glass scales allows quick and cost effective production and regrinding of high precision cutting tools. WZS 60 is used for production of tools up to 25mm diameter [3].

2.2. Measurement instrument

Zoller Genius 3s is machine for tools inspection [4]. Typical process of preparation and measurement of different kind of tool is shown in Fig. 1. The device was used for the measurement geometry. Table 1 shows the values obtained for the mill. Measured parameters of tool are input in to the software NUMROTO.

![Fig. 1. The process of preparing and measuring.](image)

Table 1. Parameters and their values for end mill obtained by Zoller Genius 3s.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix angle</td>
<td>50°</td>
</tr>
<tr>
<td>Rake angle</td>
<td>11°/9°</td>
</tr>
<tr>
<td>Radial relief angle on the cylinder</td>
<td>12°</td>
</tr>
<tr>
<td>Land width on the cylinder</td>
<td>1.8 mm</td>
</tr>
<tr>
<td>Tip relief angle - primary</td>
<td>6°</td>
</tr>
<tr>
<td>Tip relief angle – secondary</td>
<td>16°</td>
</tr>
<tr>
<td>Land width on the tip</td>
<td>1 mm</td>
</tr>
</tbody>
</table>
2.3. Tool

For measuring and resharpening was used end mill from company Walter with label H3021317-12. Tool is used for face milling, and milling grooves in the material up to 48 HRC. Shape of mill can be seen in Fig. 2, and the catalogue dimensions in Table 2. The material of the tool is carbide with TiAlN coating.

![Fig. 2. Dimensions of the end mill.](image)

![Fig. 3. The 3D model of the end mill from NUMROTO Plus.](image)

Table 2. Parameters of end mill.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutting diameter - $D_c$ h10</td>
<td>12 mm</td>
</tr>
<tr>
<td>Cutting length - $L_c$</td>
<td>26 mm</td>
</tr>
<tr>
<td>Tool length - $l_1$</td>
<td>83 mm</td>
</tr>
<tr>
<td>Length - $l_4$</td>
<td>38 mm</td>
</tr>
<tr>
<td>Shank diameter - $d_1$ h6</td>
<td>12 mm</td>
</tr>
<tr>
<td>Number of teeth - $z$</td>
<td>4</td>
</tr>
</tbody>
</table>

2.4. Software for design of end mill

NumrotoPlus is a comprehensive software package for producing and resharpening a whole range of tools. NumrotoPlus uses leading edge technology to exploit all the capabilities of modern PCs and operating systems. It can be easily networked and also used with in-house software [5]. Result of end mill design in software NumrotoPlus is a list of grinding operations (Fig.4). NumrotoPlus includes support of 2D and 3D grinding simulation. There are used three types of grinding wheels for resharpening: periphery wheel (1A1), cup wheel (11V9) and disk wheel (12V9). Periphery wheels are mostly used for flute machining, gash out and radial clearances. Cup wheels are used for relief machining and sometimes for circular grinding. Disk wheel is used for gash out end mill tip. Result of 3D simulation is shown in Fig.3.

![Fig. 4. The list of grinding operations.](image)
2.5. **Resharpening**

It is necessary to set data for touch probe before resharpening. These parameters were measured by touch probe: length of tool clamping, flute depth, lead of tool. Layer with thickness 1.8 mm was removed from the face of the end mill, because all the end mill teeth were worn. The mill was sharpened on the face of tool. The view on the face of tool is in the Fig.5.

![Fig. 5. The view on the face of tools.](image)

Cutting speed and feed when sharpening are in Table 3.

<table>
<thead>
<tr>
<th>Operation</th>
<th>Cutting speed (m.s⁻¹):</th>
<th>Feed (mm.min⁻¹):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tip gash out V2</td>
<td>19</td>
<td>50</td>
</tr>
<tr>
<td>Tip gash out</td>
<td>19</td>
<td>50</td>
</tr>
<tr>
<td>Tip relief 1</td>
<td>25</td>
<td>70</td>
</tr>
<tr>
<td>Tip relief 2</td>
<td>25</td>
<td>70</td>
</tr>
<tr>
<td>Chamfer 1</td>
<td>25</td>
<td>70</td>
</tr>
</tbody>
</table>

### 3. TESTING OF THE CUTTING PROPERTIES OF TOOLS

The tool after sharpening was tempted to testing of cutting properties. The wear of the resharpened tool was investigated and it was compared with the wear of new tool. The selected workpiece material was medium-carbon steel (ISO C45). Workpiece dimensions were 100 x 100 x 100 mm. The 5-axis machine tool DMG HSC 105 linear with Heidenhain iTNC 530 control system was used for the wear test. Cutting parameters for milling are in Table 4.

<table>
<thead>
<tr>
<th>Cutting parameters:</th>
<th>Values:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of revolutions – n</td>
<td>3978 min⁻¹</td>
</tr>
<tr>
<td>Cutting speed – vc</td>
<td>150 m.min⁻¹</td>
</tr>
<tr>
<td>Feed per tooth – fz</td>
<td>0.08 mm</td>
</tr>
<tr>
<td>Depth of cut – ap</td>
<td>1 mm</td>
</tr>
<tr>
<td>Width of cut – ae</td>
<td>9 mm</td>
</tr>
</tbody>
</table>

The maximum width of flank wear on the face teeth (VBₘₐₓ) was measured after 15 minutes of milling on both end mills. The measurement was done for every individual tooth. The average rate of tool wear after 15 minutes was used for the comparison (Fig. 6). The wear of resharpened mill is more than 4 times higher than the wear of the new tool (Fig. 7). The
reason is that the coating is not on the face of the mill after resharpening. The coating increases the wear resistance of the cutting edge [6].

![New tool – tooth 1](image1)
![Resharpened tool – tooth 1](image2)
![New tool – tooth 2](image3)
![Resharpened tool – tooth 2](image4)
![New tool – tooth 3](image5)
![Resharpened tool – tooth 3](image6)
![New tool – tooth 4](image7)
![Resharpened tool – tooth 4](image8)

Fig. 6. The wear of individual tooth of new tool and resharpened tool.
4. CONCLUSION

The paper is focused on the resharpening of cutting tools by the Cutting Tool Grinding Machine WZS 60 REINECKER. This machine is a part of The Centre of Excellence of 5-axis Machining at The Faculty of Material Science and Technology in Trnava. There were verified possibilities of the milling tool after resharpening. Testing showed functionality of the resharpened tool. It also showed necessity to do the complete tool reconditioning including recoating.

ACKNOWLEDGEMENTS

This paper was elaborated within the project VEGA 1/0615/12 Influence of 5-axis grinding parameters on shank cutter’s geometric accuracy.

REFERENCES

The reverse engineering to optimize the assembly of a conical spur gear by CAD

F. Belarifi¹, E. Bayraktar²

¹I.M.S.I/ L.S.M.C- Oran University B.P 1524 El M’Naouer Oran, Algeria,
²SUPMECA (Paris) -3Rue Fernand Hainaut 93400 Saint- Ouen France
belarifi_farid@yahoo.fr

ABSTRACT

Defining and generating a working drawing of a conical spur gear to replace an original piece requires geometrical and dimensional constraints, especially when the cutting module is determined by reverse engineering. The approach consists on drawing real tooth model, with dimensioning characteristics by CAD; extracting a database of tooth drawing that permits the determination of the real volume by mass destruction. It also allows creating a virtual model, by theoretical geometric characteristics, to calculate the volume. The numerical program gives, solution of optimization problem which consist on searching an extremum by minimising a constrained, unstrained function f(x) such that min f(x)= -Max f(x).

We have developed the model drawing according to the International Standard Organisation (ISO), in which we introduce a new non-existent coast into the current dimensioning and check the assembly using the Tredgold method, which transforms a conical spur gear into a model cylindrical spur gear, to which the equations for parallel cylindrical gearing can be applied. This method determines a good estimate if the couple of gear verifies the specified geometrical conditions of assemblies. We present the developed algorithms, as well as the results of applications of various programs.

KEYWORDS: Reverse engineering/conical spur gear / Optimization/ Tredgold method/CAD

1. INTRODUCTION

The reverse engineering is a process, used, in applied research to increase the performances of a mechanism and in maintenance to establish the working drawings of definitions for their manufacture. This process relies on the measurement techniques and their
reliabilities as to the numerical techniques of computational optimization and the CAD tools. The work that has been given to us consists to establishing the working drawing for the pairs of conical spur gear, after worn [1]. We introduce on a simulation of the assembly check [2] using the Tredgold method, which consists on substituting the conical gear couple with a model cylindrical parallel gear [1-3].

Our simulation demonstrates the geometric characteristics of a gear couple whose number of dimension is not normalised or non-existent according to ISO and NFE23-016 standards [1;4]. We use these standards to propose and establish a model that addresses this problem. This paper presents an approach that consists on using the destruction mass phenomena [3], to optimize the module of cutting, and the algorithmic procedure for the assembly check by considering the drawing model as well as its performance.

2. THE CONTEXT OF STUDY

The generation of definition drawings using reverse engineering applied to conical gearings demonstrates the difficulties of establishing the representation details according to established standards, as indicated in fig. 1a, and checking the assembly with a normal allowance with or without interference. Given that conical gear geometry is determined by spherical trigonometry theory [3,4], we distinguish two profile types that are indicated in fig. 1b.

a) ISO conical gear
b) Tooth geometry

Fig.1. Geometrical conical spur gear

The problem with gears is the correspondence between the actual component dimensions and the calculated dimensions; thus, the standard ISO introduces corrective factors to the calculations [1,2].

The difficulty with the conical gear necessitates the introduction of an approach to simplifying the assembly such that the model and theoretical gears are more similar; the Tredgold method consists on transforming conical gear with module m, Z teeth, pressure angle α, addendum cone angle δ and summit angle δ₁, into cylindrical spur gear model (Fig. 2).
3. OPTIMIZATION APPROACH

There exists the mass destruction, due to the matter transfer between the wheels, when two gears are working together. This different destruction is one of the causes of a progressive worn or broken to be under sliding [2-3]. Doing the change of variable following $m=X$

The objective function is written as:

$$V(X) = \text{Max} \sum R_i(X) b_i(X) e_i(X)$$  \hspace{1cm} (1)

Subjected to the constraints defined as follows:

- $V(X) > V_r$
- $d_a(X) < D_e$
- $\sigma_n(X) < \sigma_f$
- $\sigma_n(X) < \sigma_p$

4. THE DRAWINGS DEFINITIONS MODEL

Our model is established according to the standard ISO (Fig. 1a) with the addition of the coast $h_\beta$ (Fig. 3) at the level of the additional cone. The introduction is necessary for our model; the remainder of the coasts are defined in figure 1a and normalised according to the NF E 23-016 [1;4].
• Determination of the coast $h_{\beta}$

This coast was chosen by considering the height of the tooth $h$, to which we added three allowances of the teeth set ($j$):

- \[ h = 2.25 \text{ m} \]  \hspace{1cm} (2)
- \[ j = 0.25 \text{ m} \]  \hspace{1cm} (3)
- \[ h_{\beta} = h + (3.j) \]

We deduce that

- \[ h_{\beta} = 3 \text{ m} \]  \hspace{1cm} (4)

5. RESULTS AND DISCUSSION

Our application concerns the pairs of conical spur gear of a glass machine molding.

5.1. Optimization results

The table I, indicates the characteristics of the machine to calculate the optimization module by equation (1).

This calculation is established for different geometry of teeth.

We obtain the module of cutting $m=4$ for 25CD4 material

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z$</td>
<td>24</td>
</tr>
<tr>
<td>$P$ (Watt)</td>
<td>2285 min - 10000 Max</td>
</tr>
<tr>
<td>$N$ (tr/mn)</td>
<td>45min - 218 Max</td>
</tr>
<tr>
<td>$V_g$ (m/s)</td>
<td>0.24min - 1.094 Max</td>
</tr>
</tbody>
</table>

5.2. Check assembly results

Optimization module cutting remains inadequate if we do not check mounting operating clearance established by equations (2, 3 and 4). This part was done by a CAD (Autocad 2000)

Initial Assembly

This assembly is obtained for the moderate characteristics of the worn couple, according to the dimensions in table 1.

We obtained the distance from the axis, $a = 285.305 \text{ mm}$

The Assembly of the optimised module

This assembly is realised from the coasts of the definition drawing from the optimised module, $m = 4$ [3, 4]. It allows the allowance ($J$) between the pinion and the wheel to be verified, which has to meet the standard ISO; thus, $J = 0.25 \text{ m}$, shows that the functioning allowance is $J_f = 4.576 \text{ mm}$, with the distance from the axis $a = 268.326 \text{ mm}$.

We note that, for the optimised module, we should have an allowance $J = 1 \text{ mm}$. Given that the allowance exists, it is necessary to decrease it according to the function

- \[ Dif = J_f - J = 3.576 \text{ mm} \]

This difference automatically influences the distance from the axis of functioning. Thus, the corrected assembly is necessary.
Corrected Assembly

The guarantee of the normal allowance ($J = 1 \text{ mm}$) suggests that we manipulate the optimised assembly to abolish the difference between actual function and the normal allowance ($\text{Dif} = 3.576 \text{ mm}$). This will lead us to deduce the corrected distance from the axis.

\[ a = 264.853 \text{ mm} \]
\[ J_f = 1.018 \text{ mm} \]

6. CONCLUSION

The elaboration of the assembly check algorithms allowed us to use parametric programming with AutoCAD\textsuperscript{®} 2000. The proposed algorithms model the assembly couple, which allows us to verify the geometric characteristics: function allowed and the distance between the axis and the complementary angles must be equal to 180°. The generation of the drawing of definition according to the standard ISO and AFNOR demonstrates the benefits of knowing the module cut. The improved performance and maintenance of the machine elements demonstrates that other criteria are not automatically normalised. Thus, we introduce a new geometric model of the drawing of definition, which considers the function allowed to determine the coast $h_\beta$.

REFERENCES


APPENDIX I - SYMBOLS

\[ a : \text{distance from the axis} \quad \text{mm} \]
\[ b : \text{The width of the tooth} \quad \text{mm} \]
\[ da : \text{Diameter of head} \quad \text{mm} \]
\[ de : \text{Diameter area of site} \quad \text{mm} \]
\[ e_m : \text{Thickness of the tooth} \quad \text{mm} \]
\[ m : \text{Module of cutting} \quad \text{mm} \]
\[ P : \text{Power the machine} \quad \text{W} \]
\[ Ri : \text{Ray tooth} \quad \text{mm} \]
\[ V_g : \text{Slip velocity} \quad \text{ms}^{-1} \]
\[ V(X) : \text{Modeled volume} \quad \text{mm}^3 \]
\[ V_r : \text{Worn volume} \quad \text{mm}^3 \]
\[ \sigma_n : \text{Normal constraint} \quad \text{GPa} \]
\[ \sigma_f : \text{Practice constraint in the flexion} \quad \text{GPa} \]
\[ \sigma_p : \text{Hertzian practice constraint} \quad \text{GPa} \]