Degree of Master of Science
Second cycle, 30 credits

Estimating CO$_2$ emissions with satellite and traffic data: a Swedish practical case study

MARGAUX BOUCHER
Abstract—The large carbon footprint of industries is one of the main issues raised when talking about climate change. Active carbon monitoring methods need to be implemented to give transparency to the industry market and to spread awareness and information. This thesis investigates multiple CO₂ emissions monitoring via satellite monitoring for four different industries in the EU. The pulp and paper industry was monitored in Sweden through smoke detection coming from the chimneys’ factories. The CO₂ aggregated emissions of 14 Swedish factories were calculated with a mean error of 12%. The metal ore roasting and sintering industry were successfully monitored as well through smoke detection. In Sweden with an error 4.6%, and in the EU with an error 9.5%. The production of lime and the calcination of dolomite were unsuccessfully monitored due to no suitable method found. Finally, coke industry emissions were monitored through burned gas monitoring. The CO₂ emissions were correlated to the real emissions with a mean correlation coefficient of 0.64. This study took part in a public information campaign lead by a Swedish start-up, and some results were displayed in Stockholm, Sweden.

Index Terms—CO₂ emissions monitoring, European Union industry monitoring, Swedish industry monitoring, Satellite imagery, Carbon footprint, space monitoring.


NOMENCLATURE

EGD The European Green Deal
ETS Emissions Trading System
EU European Union
IEA International Energy Agency
KTH Kungliga Tekniska Högskolan
LKAB Luossavaara-Kiirunavaara Aktiebolag
Mt Megaton
Mt CO₂ eq Equivalent Megaton of CO₂
NASA National Aeronautics and Space Administration
R&D Research and development
RGB Red Green Blue
VIIRS Visible Infrared Imaging Radiometer Suite
VNF VIIRS Nightfire

I. INTRODUCTION

SINCE the Paris Agreement in 2015, efforts to reach climate neutrality have intensified. The European Green Deal (EGD) aims to reach climate neutrality by 2050 [7]. Countries and citizens in Europe and across the world are watching their emissions more carefully. The industry sector is particularly under the radar of the world because of its important share in the carbon balance. According to the International Energy Agency (IEA), in 2017, the total emissions of the industry sector equaled 519 Mt CO₂, which represented 23% of the total carbon emissions [8]. However, no tool is capable of monitoring the carbon emissions of industries on a monthly basis to this day. There is a lack of information available to the countries’ institutions, companies, and the public. Yearly emissions are no longer sufficient information. To prevent the situation from getting worse, new methods need to be implemented to bring transparency to the industry market, strengthen decision makers’ ability to reach their goals, and raise global awareness toward citizens.

Space technology is nowadays a well-known and integrated tool to mitigate climate change. It is used in many different ways. A satellite can be used as a monitoring tool to monitor and prevent wildfires, it can give access to wind speed and direction to enhance windmills capacity and productivity. Satellite-based systems are fully active devices. They can detect very precise assets and give precious and accurate information on industrial installations like industry heater combustion [9]. All of these integrated sources of information are waiting to be used and exploited to develop new methods to bring clarity to the industry transition.

A. Disclaimer

The work, data, and methodology hereby described in this thesis relate to an R&D (Research & Development) internship at Kayrros from February to August 2022. The methodology does not necessarily accurately reflect all or part of the methods used in Kayrros commercial products. In addition,
B. Purpose of the project

The purpose of this project is to implement a prototype method to monthly monitor carbon emissions from the industry sector. To draw a precise carbon footprint model of different industries based in the European Union (EU). The goal is to bring transparency regarding carbon emissions and enable a better following and tracking inside the EU for companies and institutions. It is also to inform and communicate those results to the public for them to have a better understanding of the situation and help them become an integrated part of the direction we are heading.

This project has been developed as part of an internship in the French start-up Kayrros. Kayrros is a leading global asset observation platform. Its purpose is to combine satellite imagery, geolocation data, and multiple sources with machine learning. Its mission is “to give energy and industrial actors the data tools they need to optimize operations, tackle the climate challenge, navigate the energy transition” [19].

A part of this study has been realized as part of a partnership with Doconomy. Doconomy is a Swedish impact-tech start-up that seeks to develop new tools to educate and raise awareness regarding climate change [18]. The common goal was to give precise carbon emissions data of different Swedish sectors for the company to create a communication campaign around carbon emissions in Sweden. Therefore, when possible, a focus has been made on Swedish industries. However, the results aim to be extended to the EU.

Nowadays, satellite imagery offers a wide range of tools available to the public. The satellite constellation offers consistent imagery with high resolution and high refresh time. These tools demonstrate accuracy and robustness in the imagery they deliver and are now an essential part of climate change monitoring. The objective of this project is to use satellite imagery as a primary source of monitoring and to complement it with algorithms based on satellite onboard devices. These space-based monitoring methods can be used as a complement to other monitoring methods already implemented, or as a fully independent tool.

This use case has been carried out as part of a private company. It will serve as a preliminary study for future projects inside the company. Depending on the results obtained, some methods will be selected to further develop fully autonomous algorithms that will be able to monthly monitor specific industries.

C. Thesis structure

This thesis falls into four parts. Firstly, an overview of the available spatial resources and tools used during this project was described. Then, the different methods chosen to evaluate yearly industry emissions were explained in detail. Afterwards, four industries were selected according to their emissions during those past six years in Sweden or Europe as well as the feasibility of the method, and the previous works done in the company. A suitable method was chosen and implemented for each of those industries. Next, the results of this study were given. They include an estimation of the error compared to official emissions reported by the EU for the period studied. Depending on the industries a focus was given either on Sweden or on the EU. Finally, the results and the use of the model were discussed and examined in the discussion section. The sustainability aspect of the project was also discussed.

II. METHODS

A. Literature review

1) Resources available: The goal of this project was to use spatial tools available to the public to monitor industries’ emissions from space. The satellites chosen had to give enough information for monthly monitoring, and access to their information had to fall under the company’s agreement. After discussions, four tools were selected for this study. The Sentinel-2 satellite, the Visible Infrared Imaging Radiometer Suite (VIIRS) tool, and the Landsat-8 and -9 satellites. These satellites were commonly used within the company project and showed different characteristics that fit the purpose of this study. Table I gives out these characteristics [9] - [11].

<table>
<thead>
<tr>
<th>Owner</th>
<th>Type</th>
<th>Spatial resolution (m)</th>
<th>Refresh Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentinel-2</td>
<td>ESA</td>
<td>Optical</td>
<td>Up to 4 days</td>
</tr>
<tr>
<td>Landsat-8/9</td>
<td>ESA</td>
<td>Optical</td>
<td>Up to 16 days</td>
</tr>
<tr>
<td>VIIRS tool</td>
<td>NASA</td>
<td>Optical</td>
<td>1/2 images per day</td>
</tr>
<tr>
<td>Google Earth</td>
<td>Google</td>
<td>Optical</td>
<td>&lt; 30 image per year max</td>
</tr>
</tbody>
</table>

2) The European Union Emissions Trading System: The EU Emissions Trading System (ETS) is a system implemented in 2005 by the European Union to monitor and regulate the European carbon market [12]. It is one of the measures intended to help reach carbon neutrality by 2050 in the scope of the EGD.

Every year, industrial installations across Europe are allocated a certain amount of free carbon allocations. These allocations represent the quantity of carbon they are allowed to release for free, and their right to pollute. One allocation is equivalent to 1 t CO₂ released.

The number of allocations received is based on an industry benchmark system of EU installations. Each installation is compared to the 10% most efficient installations in their industry in terms of average greenhouse gas emissions and risk of carbon leakage. Each industry is associated with
one benchmark, regardless of the method, technology, and processes used. These 10% serve as a reference for the allocations. They intend to encourage other industries to follow their paths and reduce their emissions. At the moment, the system is composed of 54 benchmarks [12].

At the end of each year, companies report their total emissions in terms of equivalent CO₂ released. At this point, three potential scenarios come into play. If the company's emissions match its allocations, the company has polluted what it was entitled to pollute. However, if the company's emissions do not match their allocations, they are entitled to either sell or buy their extra or missing allocations. The aim is to encourage companies to reduce their emissions to avoid paying for extra credits each year, or to even increase their profit by selling their extra credits.

All those emissions are publicly available thanks to the EU and give an insight into carbon emissions per sector and country. A more detailed report gives the carbon emissions from the EU ETS for every installation of every industry. This result is used as a reference for installations emissions across the EU.

3) Algorithms:

Company platform
All the satellite images were retrieved via the company platform. They were downloaded and processed by processes already implemented by the company. All the algorithms described in the following parts were directly used through this platform and could be combined if needed. Once the chosen algorithms were launched, the results were retrieved via codes implemented with Python.

Thermal anomaly
A thermal anomaly algorithm using Sentinel-2 imagery to detect important changes in temperatures has been implemented during this project. This algorithm has been designed to detect temperature changes around 200/300°C at best. The multispectral Imager carried by the Sentinel-2 satellites delivers thirteen bands total [21]. The algorithm is based on the combination of various bands with a resolution ranging from 10 to 60 m. Alone, they give access to vegetation index or coastal aerosol for instance [21], information that cannot be directly used in most cases. However, these bands can provide much more information if used as a base for more complicated calculations. Through this algorithm, once the geographic coordinates of an area are implemented as well as more required inputs, the desired bands are collected and processed via calculation. This algorithm based on the Sentinel-2 satellite delivers results from 2015 to nowadays.

Cloud detection
To gain precision in the imagery analysis process, a cloud detection algorithm has been used. Many algorithms used alone do not take into account the cloud coverage of the studied area. This can lead to incorrect or imprecise results if many clouds are visible during the day the satellite passes. To fully process these results, the cloud coverage has to be obtained and taken into account.

The cloud detection algorithm used gives a qualitative assurance of the probability of the presence of clouds in a selected area. It is based on the VIIRS Cloud tools from the National Aeronautics and Space Administration (NASA) [20]. For each image retrieved, a cloud index from 0 to 3 is given according to the confidence level given in Table II. Other cloud masks are used to improve the granularity and spatial resolution of the cloud coverage estimation.

<table>
<thead>
<tr>
<th>Score</th>
<th>Confidence degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Confidently clear</td>
</tr>
<tr>
<td>1</td>
<td>Probably clear</td>
</tr>
<tr>
<td>2</td>
<td>Probably cloudy</td>
</tr>
<tr>
<td>3</td>
<td>Confidently cloudy</td>
</tr>
</tbody>
</table>

This algorithm is systematically and consistently used to gain information on studied areas and to refine the results primarily obtained.

Fire detection
A small fire detection algorithm using VIIRS infrared imagery has been implemented. This algorithm was based on VIIRS Nightfire (VNF). VNF is produced with data from VIIRS, a device on board two satellites from NASA, NPP, and NOAA-20 [1]. This algorithm has been designed to detect for a set area any variations that could indicate a small fire. A small fire is defined as an industrial fire, started in the process with no “risk” to expand. Small fires are produced by a defined and located source for the industrial process. They are commonly the result of gases created during the industrial process that we want to get rid of through burning.

The small fire detection algorithm has very specific processes to achieve precision in the detection. The VIIRS bands are selected to detect only small fires and not wildfires. The intensity of the fire is not registered, the algorithm only gives an occurrence answer. Once the geographic coordinates of the area are implemented, for each image collected, the algorithm gives the results according to Table III.

<table>
<thead>
<tr>
<th>Score</th>
<th>Situation detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>Fire detected</td>
</tr>
<tr>
<td>Off</td>
<td>No fire detected</td>
</tr>
</tbody>
</table>

This algorithm based on the VIIRS satellite delivers results from 2018 to nowadays.

B. Chosen resources
In this project framework, the first step was to choose the best tools to observe and study the chosen industries. Each
industry was different, and a set of diverse sources was needed.

This study had a specific focus on Sweden, the resources chosen had to cover both the European Union and the entirety of Sweden. It had to be noted that Northern Sweden and Swedish Lapland being closer to the North pole were areas sometimes less covered by satellites. Therefore the chosen satellites had to comprise these areas in their covered area in order not to limit the study.

In terms of imagery, Google Earth was used for the preliminary work. This tool grants access to high-resolution images and covers the vast majority of the world. However, it allowed little historical data and the number of images highly depended on the areas. For this last reason, it was only used mainly for the preliminary study and not as a recurrent tool.

Regarding more consistent sources, it was necessary to have access to a reasonable amount of images per month to study and draw valuable conclusions for monthly monitoring. Taking into account the inoperable images and the cloudy days, the satellites were chosen based on a time resolution between 1 and 5 days. This time frequency can be optimized using more than one source together.

With a resolution of 10 m and a delivery image frequency of a maximum of 4 days, Sentinel-2 was one of the best satellites on the market for industry monitoring investigation. Studying its Red Green Blue (RGB) mode, a combination of its Red, Green, and Blue bands enabled enough precision to see the majority of the factories and gave enough images to pursue monthly and yearly investigation.

In the same scope, the European Landsat-8 and 9 satellites were good complements to Sentinel-2. With a spatial resolution of 30 m along a 185 km swath. The spatial resolution was a bit lower than the Sentinel-2 but used combined, these satellites enabled a higher image frequency overall.

Figures 1, 2, 3 show three view of Kungliga Tekniska Högskolan (KTH) University in Stockholm, Sweden for the three sources selected. It was clear that the three sources had different resolutions and therefore did not contain the same amount of information. The resolution of Landsat-8/9 that can be seen in figure 3 is quite low, but if the area studied was well-known, it was still possible to get some visual information.

In addition to a simple visual investigation, some algorithms could be implemented thanks to the different wavelengths of the bands available with the satellites. The algorithms were described in previous sections and were the following. Using Sentinel-2, the thermal anomaly algorithm was used during this project. The cloud detection algorithm was also implemented thanks to the VIIRS tools on board Suomi NPP and NOAA-20 NASA satellites. The fire detection algorithm using VIIRS was also used.

The purpose of this project was to bring clarity regarding carbon emissions in the Swedish and European Union industry sectors. We were looking at implementing methods to monthly monitor industries’ emissions. Each industry is unique by the products, processes, and mechanisms used and therefore needed a customized monitoring method that would ensure precision and consistency. Firstly, each industry was studied to have a clear idea of the processes used and the production unit architecture. Then, according to the industry and the available sources, three monitoring options were considered. The period studied varied between 2016 and
2021. It could be adjusted for some industries depending on the availability of the sources and the method chosen.

Overall, all the monitoring methods implemented had one goal, to give quantitative indicators of the activity that could be related to the carbon emissions for each installation. To find one process or one signal that could indicate if a plant was “On” or “Off”. Once this part was achieved, the production and the emissions were calculated and compared with a reference data source. For some industries, multiple data sources were used to provide more significant results.

C. Visual methods

The first set of methods implemented were the visual methods. A lot of industrial installations are visible and can be studied from space. Google Earth and Sentinel-2 were used as the first tools to study them. First, the installations’ architectures were studied to have a clear understanding of the areas and the surroundings. Each part of the installations had to be known and had to fit the theoretical installation plan studied in the preliminary research. Once a clear idea of what was seen from space was set, the work on the choice of the adequate method began. The choice of the methods had to help visually determine the opening frequency of the installations over the months using only Sentinel-2 and/or Landsat-8/9. To know whether or not the installations were working at the time the imagery was taken.

1) Labelling: Once the method has been chosen for one industry, all the European or Swedish installations allocated with EU ETS were located and delineated. RGB imagery from Sentinel-2 or Landsat-8/9 was retrieved for the set period studied. For each image, a label was assigned according to the set of labels previously defined by Table IV.

   

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>Visual indications that the installation is working</td>
</tr>
<tr>
<td>Off</td>
<td>Visual indications that the installation is not working</td>
</tr>
<tr>
<td>Cloudy</td>
<td>Clouds hiding partially or completely the installation</td>
</tr>
<tr>
<td>Unclear</td>
<td>Image not focused enough to determine the installation activity</td>
</tr>
</tbody>
</table>

This labeling gave access to crucial information for each plant and especially gave access to the opening/closing frequency over the months. It also gave information on the cloud coverage of the area. A very cloudy area is more likely to give imprecise or even false results. The cloud coverage of an area had to be taken carefully into account in the result analysis to fully understand the situation. This “manual” method is part of the R&D process. If the method is approved, further investigation will be carried out to find an autonomous process.

2) Calculation: Once the labeling was properly done, weights were associated with each label according to Table V.

<table>
<thead>
<tr>
<th>Label</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>1</td>
</tr>
<tr>
<td>Off</td>
<td>0</td>
</tr>
<tr>
<td>Cloudy</td>
<td>np.nan</td>
</tr>
<tr>
<td>Unclear</td>
<td>np.nan</td>
</tr>
</tbody>
</table>

The days labeled as “Cloudy” and “Unclear” were considered days with no information. They were associated with the “Nan” value of Python library “Numpy”.

When using satellite imagery, the number of images highly depends on the satellite characteristics as well as the location of the area studied. Some days, no image was provided to find the opening stage of the installation. The missing days have to be dealt with to have a better view of the monthly and yearly opening frequency. To do so, the missing days were replaced by proxy data available through different sectors. This choice was made following a method implemented in a study to estimate the patterns of daily CO₂ emissions reductions in the first year of COVID-19 [2]. This gave a new yearly opening frequency more precise than the previous one where a lot of days were left without information.

Once the labeling and counting were done, the goal was to find a common emission factor that could be re-used throughout the years. By determining the opening frequency every month thanks to the visual method, the total emissions by installation were found by Equation 1.

\[
\text{Emissions} = \text{Opening frequency} \times \text{Emission factor}
\]

Custom emission factors reflecting overall activity and coverage were used throughout this study. Then, for each plant and every year, an estimation of the emissions based on the labeling results from Table IV and the emission factor was calculated.

Finally, the relative error between the emissions’ estimations and the EU ETS emissions was calculated according to Equation 2.

\[
\text{Relative Error} = \frac{\text{Estimated Emissions}_{year \times} - \text{Emissions ETS}_{year \times}}{\text{Emissions ETS}_{year \times}} - 1
\]

If the relative error was found low and constant throughout the years, it meant that the method gave coherent emissions and could be re-used to monitor the different plants.

To gain precision in the calculation, an aggregated error, resulting from the sum of the emissions of all the installations per industry was calculated following Equation 3. In some cases, some installations might be easier to study than others. Depending on their sizes, their locations, or the quality of
the images retrieved. Aggregating the results over all of the installations can smooth the error while giving a more precise result for the whole industry. Because this study focused on emissions at a country or European scale, there was no special need for facility scale results and aggregating results were sufficient.

\[
\text{Relative Error sum} = \frac{\sum \text{Estimated Emissions}_{year} \times \sum \text{Emissions ETS}}{\sum \text{Emissions ETS}}
\]

(3)

D. Algorithm

In some situations, visual methods were not the most efficient or precise way to determine a plant’s activity. In those cases, some algorithms could be processed on the plant surroundings to study the different parts of the installations.

1) Thermal anomaly: Some industries include in their production some processes where materials need to be heated or melted at very high temperatures. When held outdoors, those processes lead to strong heat diffusion, and therefore to an important thermal difference between the material used and the outside.

The thermal anomaly algorithm using Sentinel-2 imagery to detect important changes in temperatures was used during this project. As mentioned before, this algorithm can detect changes in temperature around 200/300°C. Lots of industrial processes deal with much higher temperatures. To implement this monitoring, the first step was to identify the area of interest. To do so, the Sentinel-2 RGB mode and Google Earth were used in the same way as in the visual method explained in the previous part. Once the element that was going to be heated was identified, this element was isolated and the thermal anomaly algorithm was processed on the area.

For each image, a label was assigned according to the set of labels previously defined in Table VI.

<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>Heat detected</td>
</tr>
<tr>
<td>Off</td>
<td>No heat detected</td>
</tr>
</tbody>
</table>

The labeling also gave access to the opening/closing frequency. The same calculation method as the one used for the visual method was then applied.

2) Burning gas: Some industries include burning-gas processes. This technique can occur at different steps of the process, usually to get rid of some gases created along the way. A special unit of the plants is often dedicated to these processes. It can be found next to the main unit or a little aside. It can be assumed that the burning process is proportional to the production of the installation, and therefore could be an indication of the installation’s emissions.

To monitor the plant activity through the burning-gas process, the first step was to locate the unit in charge of burning the gas. Google Earth was used to visually find this special plant. The high resolution of Google Earth enabled us to have a clear picture of the installations and to see the small fire created with the flared gas. If this was too difficult, the thermal anomaly algorithm was processed on the whole installation area, and the hot spots were studied on Google Earth to see if any of them corresponded to the burning-gas unit.

Once the burning-gas unit was located, the fire detection algorithm was used on the area. The cloud algorithm was also used on the unit area to increase the accuracy of the results. Then, the cloud detection algorithm was applied in a binary mode to the results. For every day with a cloud detection between 0 and 1, the fire results from the fire detection algorithm were kept unchanged. However, for days where the cloud algorithm gave results higher than 1, the results from the hot spot detection were set to “no results”. This was done considering that clouds detected meant that the hot spot algorithm could not work properly.

The results obtained were then compared to reference data. The reference data sources were chosen to fit more precisely the industry studied.

E. Industries

For the scope of this study, industries were selected based on two criteria. The first industries selected were the main industries contributing to the Swedish carbon footprint that had not been studied by the company in previous projects. They were studied either only in Sweden or on a European scale. Then, more industries contributing globally to the European carbon footprint were added. Some major industries had already been studied in the past and therefore were eliminated from this study. It was the case with cement, steel, and the refining of oil. All the industries studied can be found in Table VII. The industries were named according to the EU ETS category they are receiving. The rank corresponds to the rank of the industry in the most polluting industries in Sweden/EU. The percentage corresponds to the proportion of CO₂ emissions in Sweden/EU. The results given by Table VII are calculated based on the past six years, from 2016 to 2021.

1) Pulp and paper: Sweden is an important manufacturer of pulp and paper, therefore it was decided to study it entirely in Sweden and not study other countries inside the EU. The results obtained can be extended to the EU.
### TABLE VII
<table>
<thead>
<tr>
<th>Industry</th>
<th>Rank</th>
<th>% Sweden</th>
<th>Rank</th>
<th>% Europe</th>
<th>Mean (Mt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulp and Paper</td>
<td>5</td>
<td>3.28</td>
<td>10</td>
<td>1.6</td>
<td>25.5</td>
</tr>
<tr>
<td>Metal ore roasting, sintering</td>
<td>7</td>
<td>3.14</td>
<td>23</td>
<td>0.13</td>
<td>2.24</td>
</tr>
<tr>
<td>Lime production, calcination dolomite/magnesite</td>
<td>8</td>
<td>3.12</td>
<td>9</td>
<td>1.8</td>
<td>29.3</td>
</tr>
<tr>
<td>Coke production</td>
<td>-</td>
<td>-</td>
<td>16</td>
<td>0.6</td>
<td>9.54</td>
</tr>
</tbody>
</table>

### Situation in Sweden

Sweden is well known for its high-quality pulp and paper production. This industry is divided into two categories, the production of pulp, and the production of paper and cardboard. Combined, these industries were allocated an average of 695,931 EU ETS per year during the last 6 years and represented 2.7% of EU allocations for pulp and paper production for a total of 37 paper plants and 14 pulp plants. This industry is quite important in Sweden and contributes actively to Swedish carbon emissions.

### Process

The pulp production factories are divided into multiple units and areas. The wood is usually delivered in pre-cut chips and stored in an outside area called the wood yard. The tree bark is then removed and the wood is cleaned and chipped. The cleaning process can use a lot of water depending on the method chosen, and an important amount of steam can be produced. Then goes the pulping process which includes several hours of cooking around 170°C. Again, this process can release a lot of steam. Once the pulp has been cooked, it is usually bleached with chemicals. The most prominent emissions generally associated with pulp bleaching are chlorine and chlorine dioxide, carbon monoxide, chloroform, and methanol. The smoke released by the chimney includes CO₂ from wood combustion to create energy, SO₂, and NOₓ from chemical combustion. Regarding the paper-making process, once the water is extracted from the pulp, the pulp is refined during a process using a lot once steam once again. Then, it is cleaned and dried. It is important to note that the plants in the Production of paper and cardboard ETS category welcome a broad diversity of products manufacturing, from kraft paper to tissue papers and facial tissues. Therefore, a multitude of different processes and architectures are found.

After an intensive study of the installations, two visual methods and one algorithm method were investigated. They are stated in the following content.

### Wood stocks monitoring

Pulp factories usually stock their wood in lines next to the factory in the wood yard. An example can be seen circled in blue on Figure 4. Figure 4 is a Google Earth picture of the Östrands pulp factory located in the East of Sweden.

The evolution of the wood stocks gives a clear indication of the quantity of wood used for production and therefore can be used to determine the production and emissions of the installations.

The first visual method proposed was therefore to monitor the evolution of the wood stocks for each plant. The tools chosen for this method were Sentinel-2 and Landsat-8/9. These two satellites combined enable access to pictures of the factory every two or three days. By comparing every image with the last one, an estimation of the wood used during the lapse of time could be found.

### Smoke emissions monitoring

For this industry, the majority of the installations have multiple chimneys for the different processes used inside. It was decided to use a visual method using Sentinel-2 and Landsat-8/9 to monitor the smoke released by the chimneys. It was assumed that chimneys releasing smoke meant an active installation and that a chimney not releasing smoke meant that the installation was closed or inactive.

In pulp and paper factories, two types of chimneys could be found. Small chimneys released mainly water steam used to dry the paper, and bigger chimneys released smoke and chemicals used for the pulp-making process. For this use case, it was decided to separate the chimneys and label them independently. Each chimney is referred to in the future as a sub-asset. To determine the state of a whole installation, all the sub-assets could be considered. For labeling and counting, two methods were implemented.

- Label and count the number of open days for all the sub-assets and consider the whole installation open as soon as one sub-asset is open.
- Label and count the number of open days for all the sub-assets and consider the mean of all the sub-assets as the reference for the whole installation.

It was decided to only use Sentinel-2 for this study because of the important amount of Swedish factories, and the diffi-
difficulty to identify smoke. Indeed, Landsat-8/9 resolution was too low to enable a precise analysis of the chimneys. As an example, Figure 5 and Figure 6 represent Östrands factory from Sentinel-2 and Landsat-8/9. The smoke can be observed circled in yellow in both pictures.

![Fig. 5. Östrands pulp factory view from Sentinel-2](image1)

![Fig. 6. Östrands pulp factory view from Landsat-8](image2)

Other options were considered but were dismissed during the preliminary research. The cooking process mentioned earlier can be held outside. It was envisaged to use the thermal anomaly algorithm from Sentinel-2. However, this process takes place around 170°C and therefore was under the detection threshold.

Also, some ESA and NASA satellites are designed to detect particular chemicals like carbon monoxide. However, their spatial resolutions were often too low to work at an installation scale. This idea was therefore dismissed.

2) Metal ore roasting and sintering:

Situation in Europe

Metal ore roasting and sintering is the sixteenth most EU ETS-allocated industry in the European Union. Between 2016 and 2021 it represented 13.4 ETS - Mt CO₂ eq. In the EU, the metal ore roasting and sintering ETS category holds nine installations in seven different countries for an average of 2.2 ETS- Mt CO₂ eq.

Process

Metal ore roasting is a metallurgical process based on a chemical reaction between gas and metal. It takes place at very high temperatures and aims at purifying metal. Sintering is a heat treatment process used with metal to fuse it with another element. [13] [14]

In Europe, these processes are used in many industries including zinc, steel, molybdenum, or iron production. These industries are very different in terms of processes and the metal roasting and sintering take place under different conditions. Therefore a common and systematic method needed to be found and implemented. For the scope of this study, it was decided to split the study in two between the Swedish installations and the rest of the EU.

Situation in Sweden

Sweden is well known for its iron mines in the Northern part of the country. It has been a pioneer in iron manufacturing for many years. Three mines are owned by the worldwide company Luossavaara-Kiirunavaara Aktiebolag (LKAB) and are contributing to CO₂ emissions in the region, they are therefore allocated an important amount of EU ETS every year.

Metal ore roasting and sintering is the sixth most EU ETS-allocated industry in Sweden. Based on the last 6 years, this industry was allocated an average of 666 999 ETS- t CO₂ eq per year, constituting 29.8% of the total EU allocations of 2.2 ETS- Mt CO₂ eq.

Process in Sweden

The iron mining process falls into three parts, exploration, mining, and processing. In terms of CO₂ emissions, the last part is the interesting one because it is the one considered for the EU ETS. Therefore, the focus was made on the last step of the processing, the pelletization.

After being mined, the iron ore is sorted to remove the residual waste rock and smashed into little pieces. Then, through mechanical and chemical treatment, the ore is cleaned from impurities. Even though these steps are energy-consuming, they are not producing CO₂ directly. [15]

The CO₂ released is mainly due to the pelletization process, and especially the sintering. Sintering here is the process of heating. It increases the strength of the iron by melting partially the iron ore together. The focus of this study was therefore made on the six pelletization plants owned by LKAB. Three of them are located in the biggest mine in Kiruna, two are located in Malmberget, and the last one is located in Svappavaara.
Smoke emissions monitoring
The entirety of the process is held behind closed doors, therefore it is difficult to have access to valuable satellite imagery of the process itself. However, the installations release a lot of white, dense smoke through the plant’s chimneys. The choice was made to use a visual method based on Sentinel-2 and Landsat-8/9 imagery. The same chimney monitoring method as the one used for the pulp and paper industry was implemented. In this case, each installation was equipped with only one or two chimneys, which make the labeling easier.

The chimney’s smoke of Kiruna plant can be located circled in blue in Figure 7.

Fig. 7. Kiruna plant view from Google Earth

The production of Lime and the calcination of dolomite/magnesite is the seventh most allocated industry in Sweden. Based on the last six years, this industry was allocated an average of 662 221 ETS- t CO₂ eq per year, constituting 2.5% of total EU allocations of 2.9 ETS- Mt CO₂ eq for its 9 production units.

Process
Limestone or calcium carbonate is a sedimentary rock. It is used in a variety of different fields, from agriculture to industry or architecture. Once extracted, the rock is processed to be reduced into smaller pieces. These steps do not emit much. The majority of CO₂ emissions occur at the final step of the process, during the calcination.

Calcination is a thermal process where the rock is heated around 900°C in burners. At this temperature, a chemical reaction occurs and the stones are transformed into calcium dioxide. The product is then shipped to various industries.

This industry has a unique carbon profile, unlike most carbon-intensive industries, most lime emissions are not generated by heat or power generation but by chemical reaction. An example of the architecture of this kind of factory is found Figure 10.

Burner monitoring
The first method implemented was to detect thermal anomalies around the burners. The majority of the burners are located outside and heat the stones at around 900°C. This algorithm is usually used for other industries using heaters in their processes as the cement industry. After locating the burners of the calcination plants and quarries, the Sentinel-2 images were processed with the algorithm.

Waste monitoring
After a careful study of the factories with Google Earth and the RGB mode of Sentinel-2, it was noticed that half of the plants showed piles of black residue. The thermal
anomaly was briefly used, and it was found that those piles were emitting heat. This pile can be seen in the lower left corner of Figure 10. It was still unclear what these piles of residue exactly consisted of. For each of these factories, the piles of waste were located. The area was then processed through the thermal anomaly detection algorithm.

4) Coke production:

Situation in Europe

Coke comes from coal. It is a solid residue produced by heating coal at high temperatures around 900/1100°C in oven batteries for an extended period of time. The residue is a mixture of mainly carbon, hydrogen, nitrogen, sulfur, and oxygen [5]. Coke is then used in the steel-making process for instance. Coke production was responsible for an average of 12.4 Mt CO₂ eq during the past six years, for 14 installations in the EU.

Process

The coke production process falls into four steps. Once the coal has been delivered to the plant, it has to be crushed in a crusher for consistency to fit the coke oven batteries. It is then heated in the oven. This heating is called “thermal distillation” or “pyrolysis.” To produce coke that will be used in steel manufacturing, coal is usually thermally distilled for around 15 to 18 hours. However, the process can take up to 36 hours. The temperature of the ovens ranges between 900 and 1100°C. Figure 11 represents a Google view of the Polish Przyjaźń coke plant. The blue rectangle stakes out the ovens of the plant.

Fig. 11. Koksownia Przyjaźń view from Google Earth [24]

Two kinds of distillation methods and therefore ovens exist. The first method consists of a recovery process in which the coal is heated in a completely reduced atmosphere and the volatile products are recovered in an associated by-product plant. The by-product recovery plant is located in the same plant as the coke processing unit. It is where the coke oven gas is driven off as effluent gas. It is a flammable gas that is extracted, treated, and purified by turbo-exhauster [16]. The yellow circles in Figure 11 as well as Figure 12 show the by-product recovery plant of Przyjaźń. Gas is being burned in Figure 12.

Fig. 12. Koksownia Przyjaźń by-product recovery plant view from Google Earth [25]

The second method uses “simple” ovens. In this case, the carbonization of coal is carried out in a non-recovery coke oven battery. In the non-recovery process, the air is introduced above the top of the coke bed in the coke oven. The volatile products generated during the carbonization are combusted in the oven itself for providing the required heat for the coal carbonization process. This method is only used in one of the 14 installations in the EU. the focus of the study will therefore be on plants using the first method [16].

For both methods, once the thermal distillation is completed, the hot coke is pushed outside of the oven into a “quench car.” This quench car transports the hot coke to the quench tower, where it is showered with water to prevent the coke now exposed to open air from igniting [16]. Figure 13 represents the hot coke being retrieved from the oven battery and being pushed in a quench car.

After this study of the manufacturing plant, here are the methods implemented to monthly monitor the coke plants.

Thermal anomaly monitoring of the ovens

The first attempt was to monitor the heat released by the oven battery with the thermal anomaly algorithm. The temperatures reached by the ovens during the cooking process are around 1000°C and the ovens are placed outside in every coke factory. These conditions fell within the algorithm limits and were encouraging to provide significant results of the plant activity.

Visual and Thermal anomaly monitoring of the coke

The second method implemented was to monitor the heat released from the hot coke once retrieved from the oven battery. Because the coke is very hot at this point, the thermal anomaly algorithm could spot the quench cars. This method had already been implemented in previous projects inside the company for different industries where a hot product was produced at one point in the process.
Fire detection on the by-product recovery plant

The last method chosen was to monitor the gas burned from the by-product recovery plant. The majority of European coke factories release a lot of visible flames in those units as can be seen in Figure 11. The fire detection algorithm had a good chance of detecting the gas being burned by detecting the localized fire.

III. RESULTS

A. Pulp and paper

1) Wood stocks monitoring: The wood stocks were firstly observed with Google Earth. Thanks to the high-quality images, they were located around the plants. The wood stocks were then observed with the RGB mode from Sentinel-2. It was possible to locate the stocks thanks to the brown color of the wood. However, the quality of the images was not high enough to observe clearly if the wood stocks were changing or not. The same problem appeared with Landsat-8/9. The picture retrieved from those satellites can be seen in Figure 6 and 5. The wood stock area is circled in blue. The resolution was the key problem for this method and therefore no monitoring was possible.

2) Smoke emissions monitoring: The Swedish pulp and paper industry is composed of 37 paper plants and 14 pulp plants referenced in the EU ETS. Not all of those plants produce the same products and use the same processes. Therefore, there were some differences in their architectures. Due to some very small chimneys, and the differences in density of the smoke released, the labeling part was difficult. The resolution of Sentinel-2 and Landsat-8/9 did not enable us to identify clearly the state of some sub-assets, especially the one releasing steam water. The smoke was not dense and white enough to be clearly differentiated from the installation. It was decided to focus on the installations where dense clouds of smoke were released. For this study, 14 plants were labeled and studied. The majority was pulp plants. The yellow circles in Figure 5 and Figure 6 circle the smoke released from the factory.

The labeling and counting procedure was done according to Table 1. Table VIII gives the number of days labeled as cloudy or unclear over all the plants. At this point, the days with no picture were not counted and the missing days were not filled.

<table>
<thead>
<tr>
<th>TABLE VIII</th>
<th>CLOUDS AND UNCLEAR DAYS IN DAYS METHOD 1 AND 2 - PULP AND PAPER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2016</td>
</tr>
<tr>
<td>Cloudy and unclear days</td>
<td>41</td>
</tr>
</tbody>
</table>

Once the 14 plants were labeled for the two methods previously selected, the missing days were filled according to the method stated earlier.

Method 1 - One asset “On” = Installation “On”

The first opening frequency was found considering that if one sub-asset is “On”, then the whole installation is “On”.

The aggregated error on all the factories, calculated according to Equation 5 is found in Table IX. The mean and the median were calculated by taking the absolute values of the errors.

<table>
<thead>
<tr>
<th>TABLE IX</th>
<th>AGGREGATED ERROR METHOD 1 - PULP AND PAPER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>2016</td>
</tr>
<tr>
<td></td>
<td>0.031</td>
</tr>
</tbody>
</table>

Method 2 - Mean on the sub-assets

The second opening frequency was found by calculating the mean value of the sub-assets labeling.

The aggregated error on all the factories, calculated according to Equation 5 is found in Table X. The mean and the median were calculated by taking the absolute values of the errors.

<table>
<thead>
<tr>
<th>TABLE X</th>
<th>AGGREGATED ERRORS METHOD 2 - PULP AND PAPER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>2016</td>
</tr>
<tr>
<td></td>
<td>0.024</td>
</tr>
</tbody>
</table>

B. Metal ore roasting and sintering

1) Smoke emissions monitoring: The chimney monitoring was implemented for six installations. The labeling and
counting procedure was done according to Table IV. The days labeled as cloudy or unclear were counted. At this point, the days with no pictures are not counted or filled.

The opening frequency was then obtained with the filling method previously defined and shown.

The aggregated error, error obtained from the sum of the emissions and calculated according to Equation (3) is given for Sweden and then for the EU in Table XI and in Table XII. The mean and the median were calculated by taking the absolute values of the errors.

### TABLE XI

<table>
<thead>
<tr>
<th>Year</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>-0.046</td>
<td>-0.068</td>
<td>-0.042</td>
<td>0.0</td>
<td>0.086</td>
<td>-0.035</td>
<td>0.046</td>
<td>0.044</td>
</tr>
</tbody>
</table>

### TABLE XII

<table>
<thead>
<tr>
<th>Year</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>0.079</td>
<td>0.075</td>
<td>-0.15</td>
<td>0.0</td>
<td>0.20</td>
<td>-0.11</td>
<td>0.095</td>
<td>0.076</td>
</tr>
</tbody>
</table>

To analyze the results, a coefficient linking the opening frequency and the number of cloudy days detected was calculated. The correlation between this ratio \( A \) and the emission errors is given in Table XIII.

### TABLE XIII

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant 1</td>
<td>0.2</td>
</tr>
<tr>
<td>Plant 2</td>
<td>0.6</td>
</tr>
<tr>
<td>Plant 3</td>
<td>0.6</td>
</tr>
<tr>
<td>Plant 4</td>
<td>0.9</td>
</tr>
<tr>
<td>Plant 5</td>
<td>0.6</td>
</tr>
<tr>
<td>Plant 6</td>
<td>0.8</td>
</tr>
</tbody>
</table>

D. Coke production

1) Oven monitoring: In each installation, the ovens were located with the help of Google Earth. Once the ovens were located, the plants were processed by the thermal anomaly algorithm. For each plant, a clear hot area was matching the oven locations and delimitation. However, the algorithm gave 100% of days “On”. It was found that the ovens were turned on 24/7 with no break or no “resting mode”, a mode where the oven would have been cooler and not detectable by the algorithm. Therefore, detecting when the ovens were working did not indicate the opening of the coke production plant.

2) Hot coke monitoring: The installations were studied with various satellite sources with different resolutions. Sentinel-2 did not enable us to locate the hot coke pushed out of the ovens in the quench cars, the resolution was too low to be able to confidently identify the quench cars. Google Earth enabled us to spot some quench cars but with difficulties. Because they were moving every day they were hard to consistently spot and identify with satellite images that were delivered only once a day at a set time. The thermal anomaly method was also tried to spot the quench cars but with no results.

3) By-product recovery plant: As seen before in Figure 11, the by-product recovery plants release a lot of flames due to the combustion of coke oven gas. Those plants are usually located a bit outside the plant perimeter. They were located using Google Earth.

The small fire detection algorithm using VIIRS infrared imagery was processed on each plant. The cloud detection algorithm was also processed and applied to the results as mentioned in the method part. Figure 14 and Figure 15 are two example of the difference before and after applying the cloud mask. In blue are the raw results from the small fire detection algorithm. In red are the results once the cloud mask has been applied.

C. Production of Lime and the calcination of dolomite/magnesite

1) Burner and waste monitoring: After selecting and labeling the limestone factories across Sweden and Europe, the burners were isolated and studied. The Sentinel-2 images of the areas were processed by the thermal anomaly detection algorithm and the cloud detection algorithm. No thermal anomaly was detected around the area of the burners, probably due to a lack of temperature differences. The piles of waste identified earlier were also processed with those algorithms but no consistent results were found.

Half of the Swedish plants were located next to steel plants and it was therefore impossible to clearly differentiate one from the other. The architectures of the plants allocated with this type of ETS varied a lot and no common technique was found to study their production and activity.
Fig. 15. Burning gas detected VS Burning gas detected with the cloud mask - example 2

The results were then compared with the EU ETS from the coke production category. The mean value for the cloud detection is given for each coke factory in Table XIV.

### Table XIV
Mean cloud detection per asset

<table>
<thead>
<tr>
<th>Plant</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant 1</td>
<td>1.23</td>
</tr>
<tr>
<td>Plant 2</td>
<td>0.52</td>
</tr>
<tr>
<td>Plant 3</td>
<td>0.72</td>
</tr>
<tr>
<td>Plant 4</td>
<td>0.57</td>
</tr>
<tr>
<td>Plant 5</td>
<td>0.62</td>
</tr>
<tr>
<td>Plant 6</td>
<td>0.69</td>
</tr>
<tr>
<td>Plant 7</td>
<td>1.15</td>
</tr>
<tr>
<td>Plant 8</td>
<td>0.88</td>
</tr>
<tr>
<td>Plant 9</td>
<td>0.44</td>
</tr>
<tr>
<td>Plant 10</td>
<td>0.74</td>
</tr>
<tr>
<td>Plant 11</td>
<td>0.49</td>
</tr>
<tr>
<td>Plant 12</td>
<td>0.76</td>
</tr>
<tr>
<td>Plant 13</td>
<td>0.74</td>
</tr>
<tr>
<td>Plant 14</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Figure 16 and Figure 17 correspond to yearly burning detection compared to annual EU ETS results from 2018 to 2021. Figure 16 is from a plant were the correlation coefficient was high. Figure 17 is an example were the correlation did not work. In red is the yearly burning detection obtained from the small fire detection algorithm. In blue is the annual EU ETS for the plant studied. The correlation coefficients between the flaring detection and the EU ETS can be found in Table XV.

### Table XV
Correlation coefficient between EU ETS and flaring detection

<table>
<thead>
<tr>
<th>Plant</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant 1</td>
<td>-</td>
</tr>
<tr>
<td>Plant 2</td>
<td>0.90</td>
</tr>
<tr>
<td>Plant 3</td>
<td>-0.13</td>
</tr>
<tr>
<td>Plant 4</td>
<td>0.3</td>
</tr>
<tr>
<td>Plant 5</td>
<td>0.5</td>
</tr>
<tr>
<td>Plant 6</td>
<td>-0.8</td>
</tr>
<tr>
<td>Plant 7</td>
<td>-0.3</td>
</tr>
<tr>
<td>Plant 8</td>
<td>-0.23</td>
</tr>
<tr>
<td>Plant 9</td>
<td>0.6</td>
</tr>
<tr>
<td>Plant 10</td>
<td>0.76</td>
</tr>
<tr>
<td>Plant 11</td>
<td>0.98</td>
</tr>
<tr>
<td>Plant 12</td>
<td>0.88</td>
</tr>
<tr>
<td>Plant 13</td>
<td>0.91</td>
</tr>
<tr>
<td>Plant 14</td>
<td>0.97</td>
</tr>
</tbody>
</table>

If new satellites with higher resolution and RGB mode are launched.

### IV. Discussion

#### A. Result analysis

1) **Pulp and paper:**

**Wood stocks monitoring**

The wood stock monitoring was not successful with the sources available for this study. Both Sentinel-2 and Landsat-8/9 had too little resolution for the method to be implemented. However, the results with Google Earth showed that a high-resolution satellite could enable wood stock monitoring. This is something to keep in mind for the future.

**Smoke emissions monitoring**

Looking at opening frequencies, method-2 which considered the mean opening of the different chimneys gave lower results. This makes sense because most of the plants had multiple sub-assets, usually 2 or 3. So it is possible that only one
or two were “On”, or that only some were detectable. The emission factors were higher for the second method too. It is not surprising to find very different industry factors because they were strongly correlated to the capacity and production of each plant and therefore were specific to each plant.

In the case of the Swedish pulp and paper industry, the goal was to obtain an overall CO$_2$ monitoring of the factories. Therefore, the focus was made on the aggregated errors, and not on the individual emission errors. Also, the labeling part was quite difficult for some plants, and aggregating results enabled us to smoothen the errors link to those difficult sub-assets.

Table IX and Table X give the aggregated errors of the calculated emissions from 2016 to 2020 for method 1 and 2. The first thing to notice was that the results from both methods were close, a bit smaller for method 2. This means that the second method enabled a bit more precision. However, the first method, more simple was also valid.

Except for 2017, where both methods significantly overestimate the emissions, the other emissions were estimated with an error of less than 12%. The year 2016 was particularly good with an emission error of around 3%.

As can be seen in Table VII, 2016 was the least cloudy year with 41 days recorded as cloudy. This was one explanation for the very good results obtained for this year. If the number of cloudy days is significantly higher, this can lead to worse results. However, the number of cloudy days was not the only element that was important in this method. The distribution of the clouds was also important. An illustration of this kind of problem will be given in the result discussion of the metal ore and roasting industry.

These results left room for improvement. The aggregated results were correct, but they did change from one year to another. The labeling part was difficult on some sub-assets, especially the chimneys releasing steam. A closer look at the results showed that their labeling was imprecise. Therefore a better method would be only to study the images at an asset level. To consider the installation “On” the moment smoke is detected, wherever it comes from in the factory, like method 1, but without zooming on the sub-assets. Also, a better resolution would considerably help gain more precision but was at this day out of reach.

2) Metal ore roasting and sintering: Smoke emissions monitoring was the only selected method for this industry.

The emission factors strongly depended on the plant. In this industry, there was no common manufacturing method. They all processed different metals. The emissions varied a lot depending on the production of the plant, the processes used, and the capacities.

In 2021, Plant 4 was allocated −1 EU ETS. This can happen because the EU ETS emissions were still not available or because the plant was in the process of shutting down. The error resulting from this plant was therefore not taken into account.

Looking at individual results, Plant 1, 2, and 3 showed more consistent results. Over the years, their estimations had a lower mean error, and especially a lower median around 7-8%. This was better compared with the median of the three other plants that stayed above 10%. This could be due to the location of these three plants. Indeed, Plant 1, 2 and 3 were located in Northern Sweden whereas plants 4, 5, and 6 were located in Eastern Europe. The weather in Northern Sweden is very cold, snow is present a large part of the year, and days were also sunny. In these conditions, the smoke released by the chimneys was very visible and easy to spot. Therefore the labeling was precise and consistent throughout the study.

Another factor that could influence the results was the opening frequency calculated with the filling method. As can be seen in Table XIII, for some plants, there was a correlation between the ratio $\alpha$ and the error calculated. Plant 4, 5 and 6 had a strong correlation between the ratio $\alpha$ and their emission errors. They were the plants with the most significant errors that reached respectively 30% and 70% at their maximum. For a negative error, the CO$_2$ emissions were under-estimated. The opening frequency found by the model was too low to match the real one. On the contrary, a positive error, as the one found in 2016 for Plant 6, showed that the CO$_2$ emissions had been over-estimated. This meant that the more the opening frequency was over- or under-estimated, the less precise the estimated emissions were. This can be due to the filling method chosen and the estimation of the known days. One limit appears when there are a lot of cloudy days.

Plant 6 was taken as an example. It had the most differences in its range of error. Its emission errors went from 6% to 68%. Figure 19 to 23 in the Appendix represent the distribution of clouds and unclear detections from 2016 to 2021. On these graphs, 1 represented days with a cloudy or unclear image, whereas 0 represented a day with an “On” or “Off” detection. Figure 19 has a lot of constant stages at 1, meaning that multiple days in a row, no practical information was found. On the opposite, Figure 23 had a lot more variations, meaning that even if there were a lot of cloudy days, they were spread over the year, and therefore the knowledge of the state of the installation was more consistent. From 2016 to 2021, the error calculated from Table XII for plant 6 decreased every year whereas the corresponding graphs show less and less plateau.

This example showed two of the main limitations of satellite imagery monitoring. First the weather, the more cloudy a region, the less exploitable the images. Second, the time-frequency of the satellites used. According to the time resolution, solutions and methods must be implemented to estimate data during lapses of time in which data is not available. Every method chosen has its limitations.
To gain precision, another filling method could be implemented. The method implemented was chosen after careful discussion inside the company and regarding previous work done. It is important to note that when using satellites, data availability strongly depends on exterior factors that cannot be changed.

Looking at the Aggregated error from Table XII, it was clear that the emissions aggregated to the EU were more consistent than the individual results. The mean emission error fell indeed at 9%. But, the emission error in 2020 was still high at 26%. Aggregating the results enabled smooth errors that could appear along the way. The scope of this project was to study industries at a European level, and not asset level. Therefore the aggregated error was the main object.

To have a closer look at Swedish industries, Table XI gave the emission error aggregated over the three Swedish plants. The emission errors fell under the 8% for the six years. Compared to European scaling, Sweden gave better results throughout the years and generally throughout individual assets. This showed that the method enabled consistent estimations of the company’s emissions, both at the asset and country level. The main reason for these results was the easy labeling. As stated before, the Swedish factories had the easiest labeling due to the smoke density and the weather conditions. The stage of each factory was easily determined by the visual method. This means that these industries could be a good model to try implementing an autonomous smoke detection algorithm.

3) Production of Lime and the calcination of dolomite/magnesite: At this date, no efficient method was found to monitor the industries producing lime or using calcination on their dolomite or magnesite processes. Notably, the European industries in this category were very different in terms of the architectures and processes used. Therefore it was hard to find a common method to study them all at the same time.

4) Coke production: The coke production was monitored using the small fire detection algorithm and the cloud detection algorithm.

The results of the correlation between the flaring detections corrected with the cloud algorithm and the EU ETS were given Table XV. They were quite fluctuating depending on the plants studied. Plants 7 and 8 gave very low results with a correlation coefficient of around −0.3. Four of the plants gave pretty weak results too, with correlation coefficients around 0.4 and 0.6. Surprisingly, plant 6 showed a strong anti-correlation at −0.8. However, five plants gave very good results with correlation coefficients above 0.88 for four of them and one above 0.76. It is relevant to note that adding the cloud detection algorithm to the small fire detection algorithm smoothed the flaring detection. It increased the correlation of the results by eliminating false detections due to the clouds. Also, as can be seen in Figure 14 and Figure 15, using the cloud mask smoothed the fire detection results by removing the detection that happened during cloudy days. Indeed, one limit of the VNF is that it can give false negative detections during cloudy days.

The raw results given by the small fire detection algorithm were combined with the results of the cloud detection algorithm. As can be seen in Table XV, plant 1, 7 had a mean cloud detection index superior to one. Plant 8 also had a quite high index around 0.88. This meant that over the years, many days were detected as cloudy. This could explain the low results of the correlation with the EU ETS, −0.3 for plants 7 and −0.23 for plant 8. They can be seen in Table XV. The other plants had different cloud indexes lower than 0.8. At this level we were confident regarding the cloud coverage, therefore the errors were not correlated anymore. Burning gas was not detected during the study of plant 1. Except for some very weak detections in 2019 that were removed by the cloud algorithm results. The study of this plant did not enable a comparison with the EU ETS.

As in the previous industry studies, one limitation of the model was the number of clouds and our ability to confidently detect them. For this study, the cloud detection algorithm was a robust tool that could be used with a good level of confidence. Here, the main limitation was the resources used. Indeed, the small fire detection algorithm based on the VIIRS satellite only provided results from the year 2018. This reduced the study to only four years, which decreases significantly the interpretation of the correlation. Also, the quantity of the official data plays an important role in the study. The EU ETS only gave yearly emissions figures. Access to monthly production or emission data would have given more significant and precise results. They would have been compared with the monthly results from the small fire detection algorithm over the four years and enabling a more robust correlation.

These results were promising, especially if the study focuses on individual assets that showed a good correlation with real emissions. However, it has to be continued with the 2022 data to be more significant.

B. Use of the model

As mentioned before, the goal of this project was to bring transparency to the industry market and to bring awareness regarding industries’ CO₂ emissions. To do so, a part of this study has been realized as part of a partnership with a Swedish start-up seeking to develop new tools to raise awareness regarding climate change. A three-day campaign was created by this start-up. This campaign gave the variations of CO₂ for various sectors, including the industry sector. Some of the emissions calculated in the project were combined with other emissions previously obtained inside the company. The results were then published in Europe and displayed around Stockholm to be seen by all people living or enjoying the
capital. One of the billboard located at Stockholm Central Station can be seen Figure 18.

Another goal of this project was to use the results as a feasibility test for such monitoring method. If the results were good and consistent, the idea was to try to develop algorithms that could autonomously implement the same methods. After discussion, it was decided to use the Swedish metal industry to develop a smoke detection algorithm for narrow chimneys. This could help monitor more factories and more industries in the future.

C. Sustainability

This project represented a straight continuation of the European guideline toward sustainability and the fight against climate change. The main goal of this study was to define a clear method to watch and monitor industries’ CO₂ emissions. The final purpose was dual. Firstly to inform the institutions and the general public regarding industries’ emissions and to create knowledge. Secondly, to encourage and enable industries to change their methods and processes, diminish their emissions, and reach carbon neutrality as soon as possible. To this day, this study is still in the R&D state but aims to become an integrated part of the company’s products.

V. CONCLUSIONS

The purpose of this project was to develop CO₂ monitoring methods for industries using satellite resources. It focused on the most polluting industry in Sweden and Europe. Some of the Swedish industries studied aimed to be part of an awareness campaign to spread information regarding CO₂ emissions.

Four industries were studied, the Swedish pulp and paper factory, the Swedish and European metal ore roasting and sintering industry, the production and calcination of lime and the calcination of dolomite/magnesite in Europe, and, finally, the coke production in Europe.

Two methods were selected and implemented. For the two first industries, the emissions were estimated using a visual method to monitor chimneys’ smoke. The opening frequency of the facilities was deduced from visible smoke signals. Then the emissions were estimated through emissions factors.

For the pulp and paper industry, two methods were implemented to determine the opening frequency of the installations. The installations were either considered opened as soon as one of their chimneys was emitting smoke or the opening frequency was calculated considering the mean opening frequency of each chimney. The emissions were estimated with a mean error of 12% compared to the EU ETS for both labeling methods. However, the second method showed slightly better results at the installation level. One of the weak points of this method was the consistency of the results at the plant level. The visual method implemented was difficult to set up. Zooming on the chimneys of each plant was revealed not to be the best choice due to the resolution of the Sentinel-2 satellite, therefore leading to imprecise results. The results could maybe be improved by changing the image analysis. If the focus is made on images of the plant directly and not images of the chimneys, it could be easier to determine the stage of the plants.

The metal industry was as well monitored through this method, but with better and more consistent results. The emissions were calculated with a mean error of 9.5% for the EU and 4.6% for Sweden. The main difference came from the easy monitoring of the Swedish industry. They released dense white smoke in good weather conditions. With this mentioned, smoke monitoring of the Swedish iron industry was considered the best case to go further in the study, and to try to implement an autonomous algorithm that could detect smoke coming from these kinds of chimneys.

The emissions linked to the production of coke were monitored via the burning gas process that occurs during the manufacturing process. The emissions were compared with the EU ETS emissions with a mean correlation coefficient of 0.64. Some of the plants had better results than others and therefore could be monitored independently. The next step of the research would be to add some more granularity to the data used to have more points in the correlation computation.

To this day, no method using satellites has been found to monitor the production and calcination of lime and the calcination of dolomite/magnesite.

Finally, some of the results were published around Europe and across Stockholm as part of the partnership with Doconomy, to be available to the general public. This followed one of the purposes of this thesis, to make people and institutions aware of where their emissions are coming from and therefore give them the tools to fully understand the situation.
AFFILIATION

This report was submitted for review and approval on September 22, 2022.

Margaux Boucher is a Master of Science degree student in Aerospace Engineering (space track) at Kungliga Tekniska Högskolan, Royal Institute of Technology in Stockholm, Sweden. She is also a Master’s student at Institut d’Optique Graduate School in Palaiseau, France. She is a Junior Energy Analyst at Kayrros in Paris, France.

Gunnar Tibert is associate professor in space technology at the Engineering Mechanics Department at Kungliga Tekniska Högskolan in Stockholm, Sweden. He served as the examiner for this thesis project.

Clément Licour is product manager at Kayrros in Paris, France. He served as the supervisor for this thesis project.

ACKNOWLEDGMENT

I would like to thank Kayrros for this incredible opportunity. Being part of this company during this six-month thesis was a chance for me. I have found people eager to share knowledge and experiences. Being able to join such an important and impactful project was an amazing and growing experience. I would like to thank the company and the staff for their support and trust all along this thesis.

I would also like to thank Dr. Gunnar Tibert for his faith and his trust as my supervisor during my thesis, as well as for the various courses that I had a chance to attend at KTH. They give me the tools I needed to fully comprehend and start my career in the space field.

Finally, I would thank my family and friends that showed great support during the writing phase of the process. I am very grateful for their limitless confidence and support in my projects.
REFERENCES


[23] Google Earth 7.3.4.8738 (64-bit). KTH Royal Institute of Technology. 50°20'49, 35°N, 18°04'26, 57°E eye alt 511 m. [Accessed: 13- Aug- 2022]


**APPENDIX**

Fig. 19. Detection of cloudy and unclear days 2016 - Plant 6

Fig. 20. Detection of cloudy and unclear days 2017 - Plant 6

Fig. 21. Detection of cloudy and unclear days 2018 - Plant 6

Fig. 22. Detection of cloudy and unclear days 2020 - Plant 6

Fig. 23. Detection of cloudy and unclear days 2021 - Plant 6