KTH ROYAL INSTITUTE OF TECHNOLOGY

Degree project in Civil Engineering and Urban Management

Second cycle, 30 credits

Evaluating an automated method for digitizing detailed plans

Using a Swedish municipality as test case

LINNEA BERGMARK

Stockholm, Sweden 2022

Evaluating an automated method for digitizing detailed plans

Swedish title: Utvärdering av en automatiserad metod för digitalisering av detaljplaner

Degree Programme in Civil Engineering and Urban Management Master's Programme, Transport and Geoinformation Technology

Author: Linnea Bergmark

Date: June 17, 2022

Supervisor: Dr. Takeshi Shirabe
Examiner: Dr. Yifang Ban
School of Architecture and the Built Environment
Host company: Agima Management AB
External supervisor: Mikael Grönkvist

Cover image: AI generated image using the key phrase "Automatic Digitizing of Detailed Plans".
Ai generator source: https://creator.nightcafe.studio/

Abstract

With new directives from Swedish authorities imposing municipalities to digitize sections of their plan archives, the question of digital detailed plans is becoming more relevant than ever in Sweden. Digitizing already existing detailed plans is time consuming, so effective automated digitizing methods will be valuable to save time in this process. However, in order to know if a method is effective it first has to be evaluated. This study aims at evaluating a recently introduced method for automated digitizing of detailed plans, and it is the first one evaluating this method in a quantitative manner. The questions to be answered within the study is whether the implemented method is effective and if it has any weaknesses. Additionally, whether a number of defining characteristics of the detailed plan maps influence the quality of the result. As the quality of digitized detailed plans have not been subjected to systematic evaluation before, a novel contribution of this study is also suggesting a framework for how this can be measured and evaluated.

The method consists of 3 steps and the first 2 steps, namely automated georeferencing and automated vectorization, have been performed on a set of 75 detailed plans. Using manually digitized versions of the same detailed plans as ground truth, the results of these two steps have been compared and evaluated using a set of quantitative measures.

Findings from this study have shown that about 70% of the detailed plans tested can be georeferenced, and 44% of relevant areas in the plan maps can be vectorized using the method. However, the results have displayed a significant disparity of quality, with error values for georeferencing ranging between under 5 meter for the best results and over 100 meters for the worst.

The weaknesses that have been identified for the method are mainly that the georeferencing procedure requires extensive manual supervision, that the vectorization produces polygons of ambiguous belonging, and that the method is limited to multicolor detailed plans. Furthermore, a small plan area has been identified as the most influential factor for a low quality result. Main conclusions of this study have been that the method can be considered effective for digitizing detailed plans to some extent. Additionally, the method for evaluating the quality of digitizing could be expanded by reviewing more factors such as shape and gaps between polygons in future work.

Sammanfattning

Frågan om digitala detaljplaner är mer relevant än någonsin i Sverige efter nya direktiv från svenska myndigheter som ålägger kommunerna att digitalisera delar av sina planarkiv. Att digitalisera redan existerande detaljplaner är en tidskrävande process, vilket innebär att effektiva automatiserade metoder kan bli värdefulla för att kunna spara tid. Men för att att veta om en metod kan sägas vara effektiv behöver den först utvärderas. Denna studie syftar till att utvärdera en nyligen presenterad metod för automatiserad digitalisering av detaljplaner, och är den första kvantitativa undersökningen av metoden. De frågeställningar som undersöks är huruvida metoden är effektiv och om den har några svagheter. Dessutom analyseras resultatet utifrån ett antal egenskaper hos detaljplanerna, för att se om dessa egenskaper påverkar kvalitén. Eftersom kvalitén på digitala detaljplaner inte har studerats systematiskt i något tidigare sammanhang har ett av studiens bidrag också varit att ta fram ett ramverk för hur detta kan mätas och bedömas.

Metoden består av totalt tre steg och de två första stegen som innefattar automatiserad georeferering samt automatiserad vektorisering har applicerats på totalt 75 detaljplaner. En kvantitativ utvärdering av metodens två första steg har sedan genomförts med hjälp av jämförelsedata i form av manuellt digitaliserade detaljplaner.

Studien visar att ungefär 70% av detaljplanerna i testen kunde georefereras, och 44% av bestämmelseytorna kunde vektoriseras med hjälp av metoden. Resultaten har dock stor spridning i fråga om kvalitét, med felvärden på 5 meter för de mest lyckade resultaten och över 100 meter för de sämsta.

Svagheterna som identifierades för metoden handlar främst om att georefereringsprocessen krävde omfattande manuell granskning och att den var begränsad till detaljplaner med plankarta i fyrfärg. Dessutom hade de vektorer som genererades ingen självklar tillhörighet till något bestämmelseområde. Gällande vilka egenskaper som påverkade resultatet fann studien att det framför allt var en liten detaljplanearea som hade ett negativt inflytande. De huvudsakliga slutsatserna från studien var att metoden till viss del kunde anses digitalisera detaljplanerna på ett effektivt sätt, och att utvärderingsmetoden kan byggas vidare på i framtida studier genom att ta hänsyn till ytterligare faktorer såsom form på- och avstånd mellan polygonerna.

Acknowledgements

I would like to express my deepest appreciation to Dr. Takeshi Shirabe for supervising this project and generously giving feedback on the content which has been incredibly valuable. I am also grateful for my external supervisor Mikael Grönkvist at Agima Management AB who presented this intriguing topic for me, and have provided valuable feedback, shared knowledge and contacts that this endeavor would not have been possible without.

Thank you Staffan Kjellman, consultant at Agima working as project leader at Lantmäteriet, for providing the necessary software and reports and sharing your knowledge. I want to thank Arpine Hakobyan, Daniel Morosini and Mats Åberg at Upplands Väsby municipality for contributing with detailed plans for the study. I'd also like to acknowledge Nora Andersson Eklund who has helped me with proofreading the paper.

Contents

1. Introduction **7**
1.1 Detailed plans 7
1.2 Digital detailed plans 10
1.3 The current status of digitization of detailed plans 11
1.4 Areas of use for digital detailed plans 12
1.5 The method to be evaluated 12
1.5.1 Key terms and definitions 13
1.5.2 The three steps of the method 14
1.5.2.1 Step 1 15
1.5.2.2 Step 2 17
1.5.2.3 Step 3 19
1.5.3 Results and past evaluation of the method 19
1.6 Scope limitations 20
1.7 Research questions 20

2. Literature review - digitization of maps and evaluation methods **21**
2.1 Map digitizing in related work 21
2.1.1 Challenges of digitizing maps 21
2.1.2 Georeferencing 22
2.1.3 Evaluation of georeferencing methods 23
2.1.4 Color image segmentation and vectorizing of maps 25
2.1.5 Evaluation of color image segmentation and vectorization of maps 27
2.2 Automated digitization of detailed plans 28
2.2.1 Challenges of digitizing detailed plans 28
2.2.2 Accuracy specifications for digital detailed plans 30

3. Methodology **32**
3.1 Study area and Data 32
3.1.1 Study area 32
3.1.2 Raw data 33
3.1.3 Preprocessing 33
3.1.4 Accuracy 35
3.2 Procedure 35
3.2.1 Step 1 35
3.2.1.1 Output for Step 1 37
3.2.1.2 Evaluation of Step 1 38
3.2.2 Step 2 38
3.2.2.1 Output of Step 2 40
3.2.2.2 Evaluation of Step 2 41
3.2.2.2.1 Matching procedure 42
3.2.2.2.2 Evaluation measures for Step 2 46

4. Result **47**
4.1 Evaluation of Step 1: Georeferencing 47
4.2 Evaluation of Step 2: Vectorization 49
5. Discussion **51**
6. Conclusions and future work **55**
6.1 Conclusions 55
6. 2 Future work 56
7. References **58**

1. Introduction

Detailed plans are important juridical documents regulating building as well as land and water use in Sweden (Boverket, 2021b). From 2022, a new directive stated that all new detailed plans in Sweden had to be in a digital format (Sweco, 2020). The digital plans were to be available in a database administered by Lantmäteriet, which is the Swedish authority regulating geodata, property formation and property registration (Lantmäteriet, n.d).

In a report about digitization of detailed plans, Sweco (2020) stated that it was important to digitize the already existing detailed plans, since it would make the information more accessible for the public and different types of urban planning stakeholders. In addition, it would save time and reduce costs in the urban planning process. However, since there are over 100,000 detailed plans in Sweden that are not digitized, it would take over 400 years in work hours to digitize them all manually. Therefore, methods automating parts of the digitizing process can be important for saving time (Lantmäteriet, 2019).

A method called Automatisk Digitalisering av Detaljplaner (Automatic Digitizing of Detailed Plans, abbreviated ADDP), has been presented by Lantmäteriet for this purpose, but it has not yet been evaluated in a quantitative manner. The method consists of a three-step process for georeferencing and vectorizing detailed plans, intended to save time for municipalities in the digitizing process (Sweco, 2020). But if ADDP is to be used by municipalities in their digitizing work, it is important to first know what level of accuracy can be achieved, and what limitations there are to the method.

1.1 Detailed plans

Boverket (2022a) describes detailed plans as legal documents regulating what can and cannot be built. In addition, they specify the intended use of land and water within a specific area. Detailed plans are created and administered by municipalities, and the municipalities are required by law to establish a detailed plan for new exploitations or change of land use, if the public or neighborly interest motivates it. The document can then be used as a basis for municipality decisions on whether changes in characteristics, or preservation of existing buildings should be granted permission. Additionally, it is used for deciding if an area is suitable for a certain exploitation.

According to Boverket (2022b), regulations of detailed plans can be divided into zonal and land use regulations. Land use regulations describe how land and water areas can be used and are divided into three main categories: water, public spaces and development districts. Water areas refers to open water bodies or other types of water areas with minor facilities such as bathing piers. Public spaces are areas intended for public use such as streets, parks, nature or squares. Development districts can be residential areas, offices, commercial areas, technical facilities or other types of land uses that do not fall under either of the first two categories.

Furthermore, detailed plans describe zonal regulations, which are limitations in building design for certain areas. These are more specific than the land use regulations, and serve as complementary information about what kind of development that is allowed in an area. Limitations can take the form of specified maximum height, area or roof angle of buildings. Zonal regulations can also specify visual characteristics such as allowed colors, materials or shape of the buildings. Additionally, the placement of buildings can be restricted through zonal regulations. Boverket (2021c) defined “dotted ground” as areas where no built structures are allowed. These are marked with black dots in the plan map. With the help of dotted ground it is possible for planners to decide the exact placement of new structures within a certain area.

Boverket (2021b) have stated that the detailed plans always consist of two compulsory documents; a map called ‘plan map’ (Figure 1) and a text document called ‘plan description’. The plan description contains information about the intention of the detailed plan and how the map should be interpreted. In addition to these two, other documents are commonly included, usually additional maps displaying property subdivision or topography. These should be available if they are deemed necessary for the context.

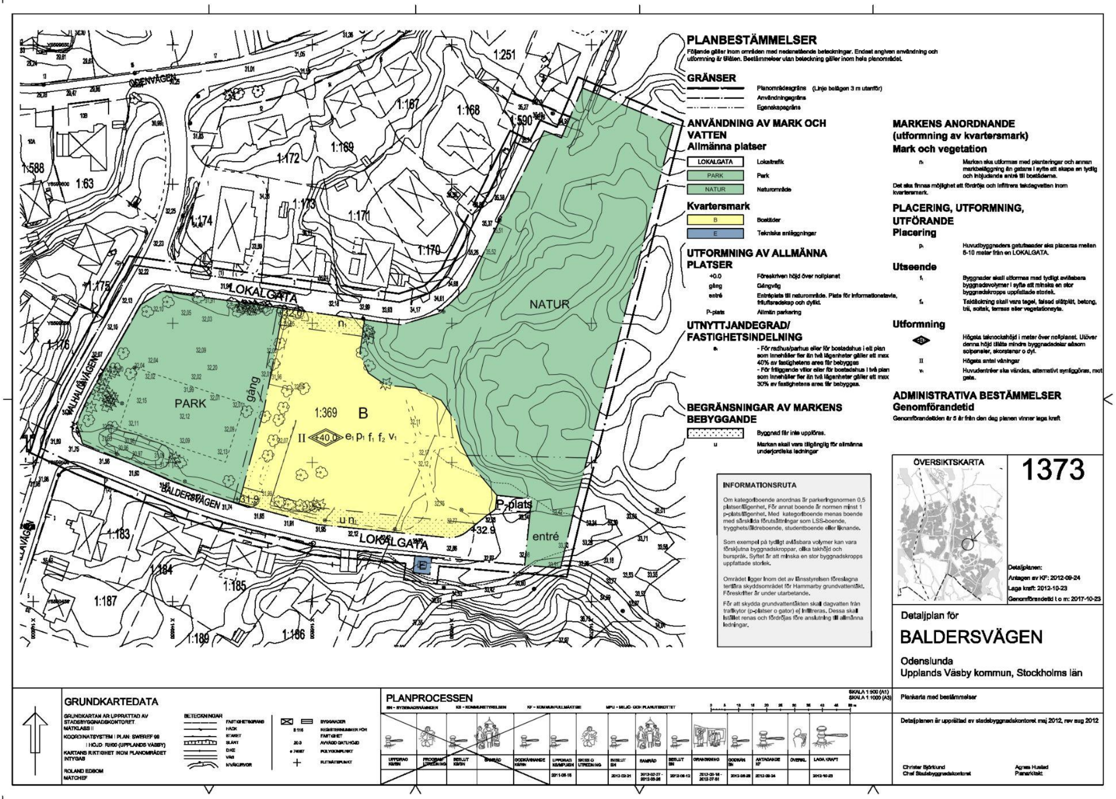

Figure 1. Example of detailed plan map.

The plan map marks the detailed plan’s location and displays regulation areas through different visual characteristics that are gathered in the plan map’s legend (Boverket, 2022a). Boverket (2021d)’s regulation catalog (Planbestämmelsekatalogen) contains the directives for how the plan map features should be displayed. Sub-areas of the map with their own specific regulations should be clearly displayed using different colors (Figure 2), dots to mark dotted ground (Figure 3), or by a text label describing the regulations (Figure 2). Lines with different styles are used to mark the plan border and regulation area borders. In some maps, an additional line is also marked three meters outside the plan border (Figure 3). Aside from this information, the plan map is also commonly provided with cross markings along with corresponding coordinates, which can be used as control points when georeferencing the map (Figure 4). The plan map is the only part of the document that has legal bearing, the rest of the information is only meant to aid the interpretation of the plan map (Boverket, 2022a).

Figure 2. Legends presenting regulation areas marked with color and text labels.

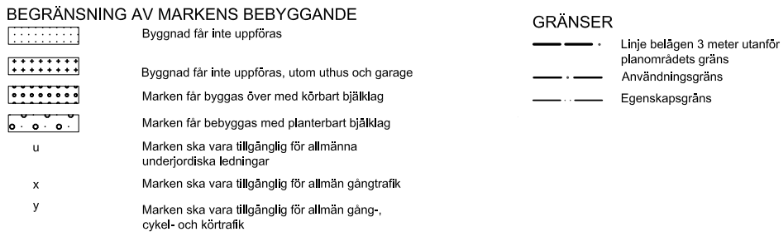

Figure 3. Left: Legend presenting regulations marked with dotted ground and symbols. Right: Legend presenting border lines for regulation areas, plan area and 3 meter outside plan area.

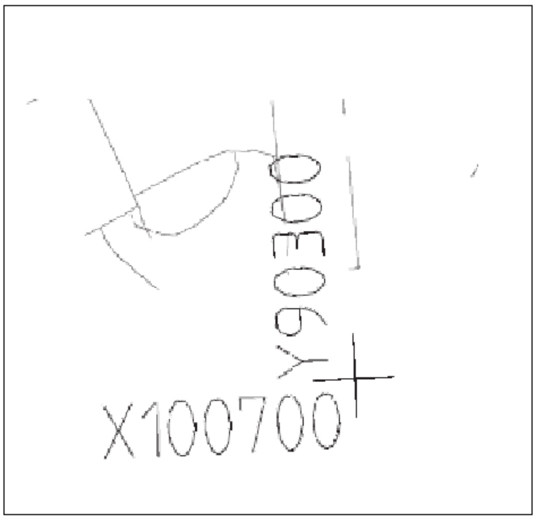

Figure 4. Control point cross marking and corresponding coordinates found in a plan map.

1.2 Digital detailed plans

A detailed plan is a set of documents that can be either in physical paper form, scanned to raster format or created digitally in raster format. A digital detailed plan does not replace the original detailed plan map that has been decided by the municipality. The physical map or raster plan map is still the juridical document that holds legal bearing, which means that the digital detailed plan is not juridically binding (Thulin, 2022).

Boverket (2021a) laid out directives in 2021 for what kind of digital information should be available in new detailed plans. They stated that the following information should be digitally connected to each plan area: name of the municipality, diary number for the plan, name of detailed plan, reference to the decision protocol, date for initiation and date for legal force. Additionally, information about implementation time and regulations should be digitally connected to the plan. This includes land use and zonal regulations which are present within the detailed plan area. The digital information should be structured in a uniform manner to make it easier to reuse and transfer. Additionally, the information should be linked to a specific spatial location. It is not enough to have a scanned raster version of the detailed plan documents for it to be considered digitized (Boverket, 2020).

Boverket (2017) conducted an investigation on technical specifications for how digital detailed plans can be designed, which suggested that the desirable format for digital detailed plans should be georeferenced vector polygon format in two dimensions. Both the detailed plan area and the regulation areas should be represented in this manner, with attributes such as data quality and dates for validity and legal force assigned to each polygon. This suggestion has since then been made into a national specification by Lantmäteriet (2022) who established that digital detailed plans should be in vector format, with the above mentioned characteristics.

1.3 The current status of digitization of detailed plans

According to a study made by Thydén (2021) where municipalities were asked to describe their current status of and attitude towards digitizing detailed plans, the average detailed plan takes eight hours to digitize manually. The study showed that 43% of municipalities in Sweden had not started digitizing their detailed plans. 9% had finished the work and 21% did not plan to digitize their plans. The municipalities that had started the process usually had begun with the newest detailed plans and then worked their way back. Among the municipalities that stated that they would not digitize their plans, the two main reasons given for this was lack of resources (mainly financing and personnel) and that it was not prioritized. Only 9% stated that the reason was that there was no need for digitized plans. The vast majority (86%) of municipalities that planned on digitizing their detailed plans stated that scanned or rasterized PDF files would be their basis in this process.

1.4 Areas of use for digital detailed plans

As the digital detailed plans hold no legal bearing, they are not meant to replace the original detailed plans. However, a number of complementary fields of application have been presented by Boverket (2020) for digital detailed plans. For example, digitizing the detailed plan information makes it possible to arrange and present it to the public in a more accessible way compared to having it divided into separate documents for each plan. A digital format makes it possible to sort what regulation areas should be visualized, and GIS (geographic information system) analyses can be made based on the data.

One example of this that has been presented by Boverket (2020) is how stakeholders could have an overview of which areas' detailed plans allow for new land exploitation, saving time when initiating new construction projects. By having the regulation areas accessible as digital data, it could also be possible to perform a multi-criteria analysis weighting in factors like distance to public transport when making such decisions, without having to do the time consuming work of compiling this information each time. Another example from the same author is how detailed plan administrators working at municipalities could have their work simplified when having access to a digital overview of detailed plans from other areas of Sweden. In this way, decision making could be aided by being able to search through a nationwide database to find out how questions regarding certain detailed plan issues have been solved in the past. This would save time and make more use of the accumulated knowledge of the detailed plan process. Thus even though the digital detailed plan does not replace the original documents, it could have complementary functions such as the ones listed above, aiding the overall urban development process.

1.5 The method to be evaluated

The ADDP method report by Sweco (2020) has described the background, technical contents and result of the method that will be evaluated in this thesis. The purpose of the method that was ordered by Lantmäteriet, was to fortify the digitalization of existing detailed plans in Sweden using a Proof Of Concept (POC). This means that ADDP was an early-stage investigation of possible technical tools and methods to automate the digitizing process. The tools that produced the most promising results were then set together and presented in a three-step method. The intention was that the municipalities would be able to use the method to aid their digitizing process, but it has not yet been adapted by any municipality. The report stated that input data for the method were the detailed plan map, along with a georeferenced polygon in the shape of the detailed plan border (Table 1 and Figure 5). The detailed plan map could either be a scanned physical document or digitally created and was retrieved from Arken (Lantmäteriet's plan database) or municipalities' own WebGIS. The plan border vector set containing the polygons was accessed from Lantmäteriet's database for geoinformation. Moreover, the final product of the digitization method was two dimensional vector polygons marking the border of plan regulations.

Input data	Format	Source
Detailed plan map	PDF (raster)	Municipality WebGIS
Detailed plan border polygon	Shape (vector)	Lantmäteriets database

Table 1. Input data for the ADDP method.

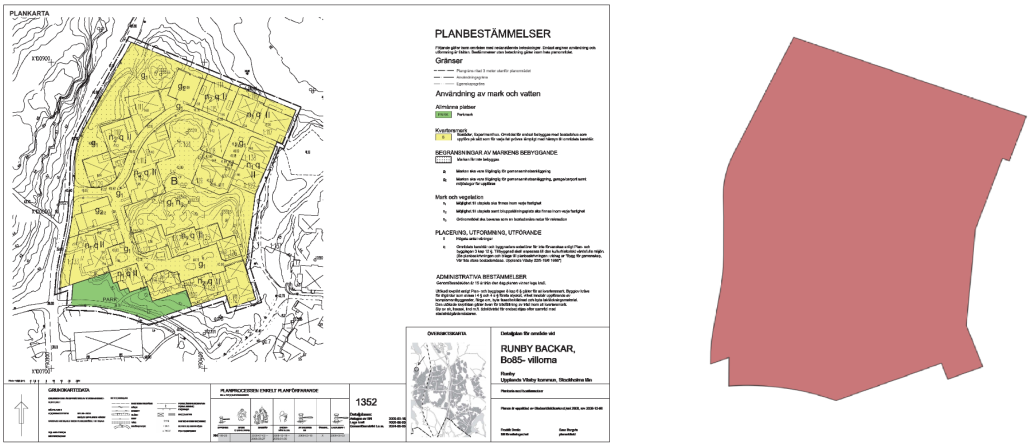

Figure 5. Example of input data. Left: Detailed plan map. Right: Corresponding plan border polygon, containing spatial information.

1.5.1 Key terms and definitions

Some key terms are introduced for the ADDP method and the context of the thesis.

- Map size relation: the relation between the document size of the detailed plan map and the area on the document that contains the map illustration (see Figure 6).
- Rasterized border shape: The contour line of the plan border polygon that has been rasterized (see Figure 7).
- Regulation area: an area within a detailed plan where a specific set of land use or zonal regulations are valid. Regulation areas on the plan map are visualized through color, dotted ground, text or symbols.

- Plan border polygon: The vector polygon that has the same shape and geographic location as the detailed plan area.
- Plan border vector set: the full vector layer containing all plan border polygons.

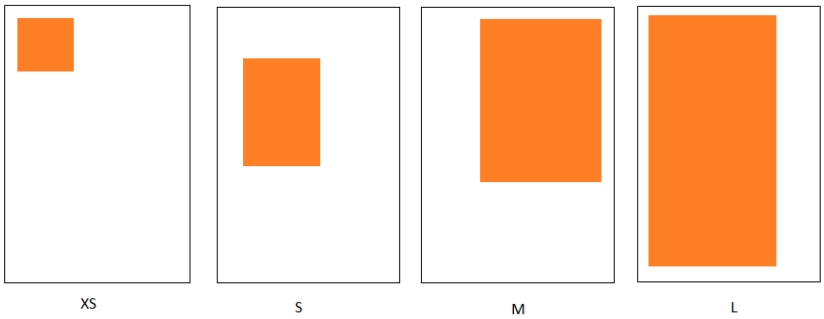

Figure 6. Map size relation. The area in the document that contains the map illustration (orange) and the total map document size (black border).

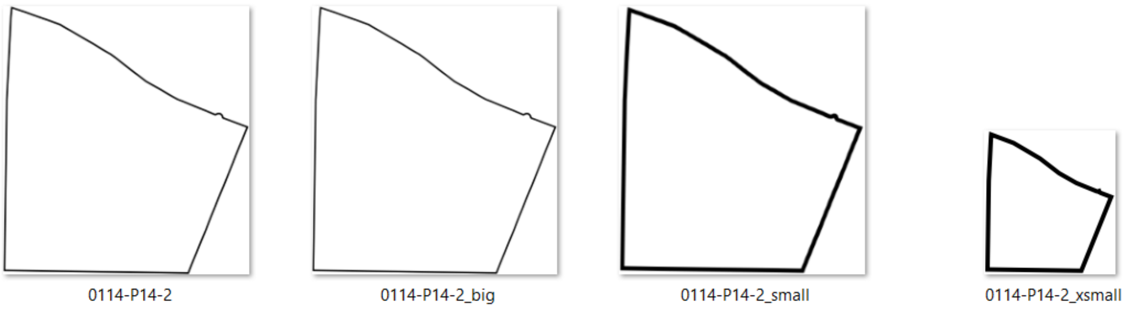

Figure 7. Rasterized border shapes of the same detailed plan in the four different sizes.

1.5.2 The three steps of the method

In its entirety, ADDP consisted of three steps. The first step was to georeference the plan map, the second step was to vectorize it, and the third step was to identify and attach regulation area attributes to each vector. A decision was made to limit the study area to the first two steps of the method (Figure 8), to keep the scope feasible within the time frame of the project. This means that the third step of the method was not evaluated within the scope of this thesis.

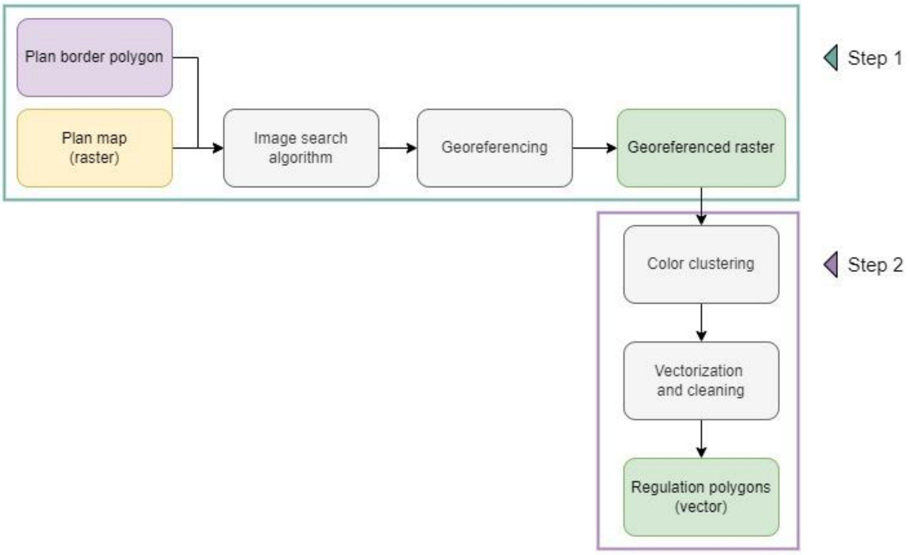

Figure 8. Overview of the first two steps of the ADDP method.

1.5.2.1 Step 1

Step 1 refers to the first step of the ADDP method, which was to georeference the plan map. First the concept of this step will be briefly described, and then each stage of the procedure will be explained in detail. An overview of Step 1 can be seen in Figure 9 below. The output of this step was a georeferenced raster (number 7), produced by transferring the geographical information from the polygons in the plan border vector set to the plan map. The polygon was placed at the correct position in the plan map using an image search algorithm. For search input the rasterized border shape of the polygon was used.

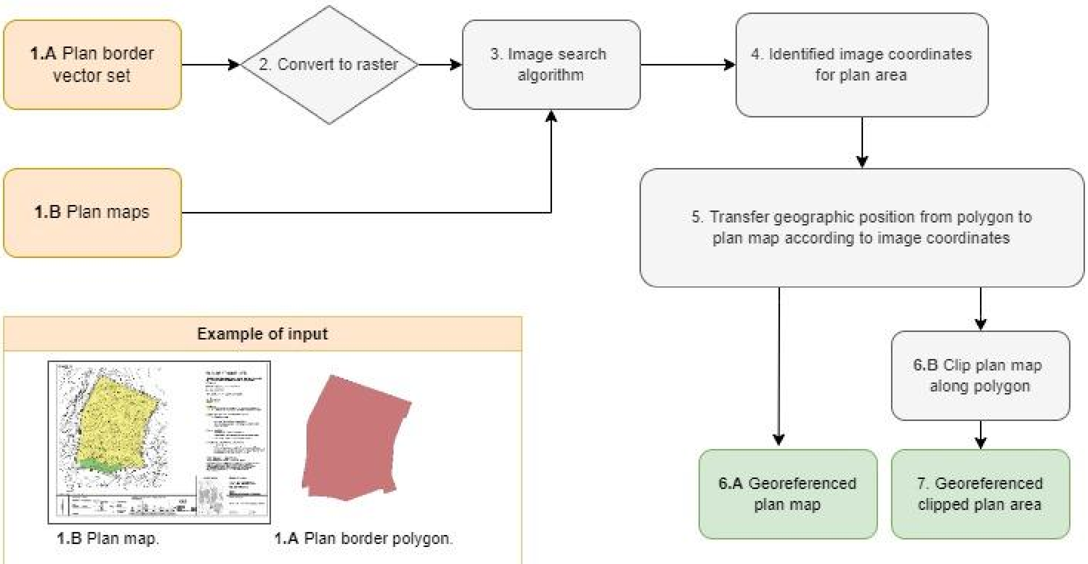

Figure 9. Illustration of the georeferencing process named Step 1.

Step 1 started with taking 1.A (plan border polygon) and 1.B (plan map) as input, and then converting each polygon in the plan border vector set to a raster image as described in Task 2. Only the contour of the polygon was transferred to the raster, creating a black border shape on transparent background, as seen in Figure 6. The rasterized border shape was then resampled into four different sizes, and the map size relation determined which size would be used as input for the image search algorithm. The map size relation, which refers to the relation between the total map size and the detailed plan size on the map, was given as metadata input in the form of an attribute table for each plan map. This metadata also contained two other attributes, which was plan ID to be able to match each plan map to the right plan border polygon, and information about the coordinate reference systems for any necessary reprojections.

Using the plan ID and map size relation, each plan map was then paired with the corresponding rasterized border shape of correct size, and the rasters were fed pairwise into the image search algorithm (Task 3). The border shape was used as search input, and the algorithm aimed to identify the area within the plan map that was most similar to the given shape. If this was done correctly, the algorithm identified the plan border in the map.

Multi scale template matching is the name of the search algorithm that was used in Task 3 of the procedure, as described by Sweco (2020) in the ADDP report. It is based on an implementation of Template matching, a predefined function within the programming function library OpenCV which matches two rasters to each other based on similarity. This specific implementation allowed for matching raster features even if they had slightly different scales, and it returned image coordinates for a box containing the matched area, as seen in Figure 10. As the algorithm had limited capacity for handling too large differences in scale, having the rasterized border shape in four different sizes was necessary, and the map size relation was used to determine the relevant border shape size for each map.

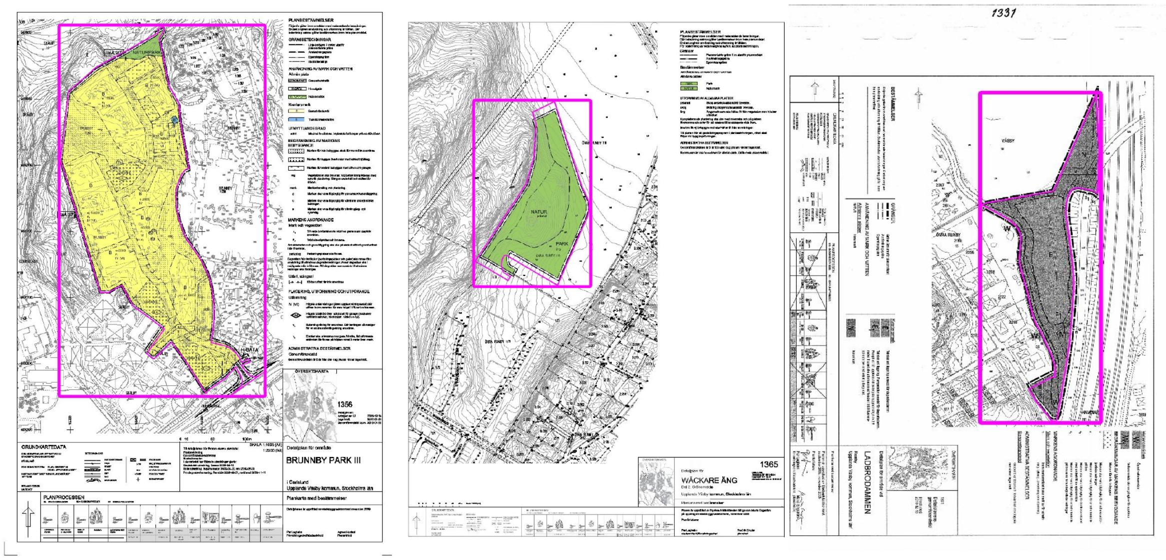

Figure 10. Three examples of plan borders as identified by the image search algorithm, each marked with purple lines.

Using the image coordinate output from the multi-scale template matching, the plan map was placed in the correct relation to the georeferenced plan border polygon. In this way, the plan map raster was georeferenced. Then, as preparation for Step 2, the georeferenced map was clipped using the plan border polygon, resulting in a raster containing only the plan area.

1.5.2.2 Step 2

In the following section, Step 2 of ADDP will be explained in detail. In this step, the georeferenced raster underwent a color image segmentation procedure. The aim was to identify all regulation areas that were visualized in different colors, and segment all pixels based on their color. The output from this step was georeferenced vector polygons of regulation areas for each color field in the plans.

With Figure 11 below as the basis, the procedure will be described stepwise. As seen in the figure, input was the georeferenced clipped plan area from Step 1. This input was a raster showing only the part of the plan map that contained the detailed plan area.

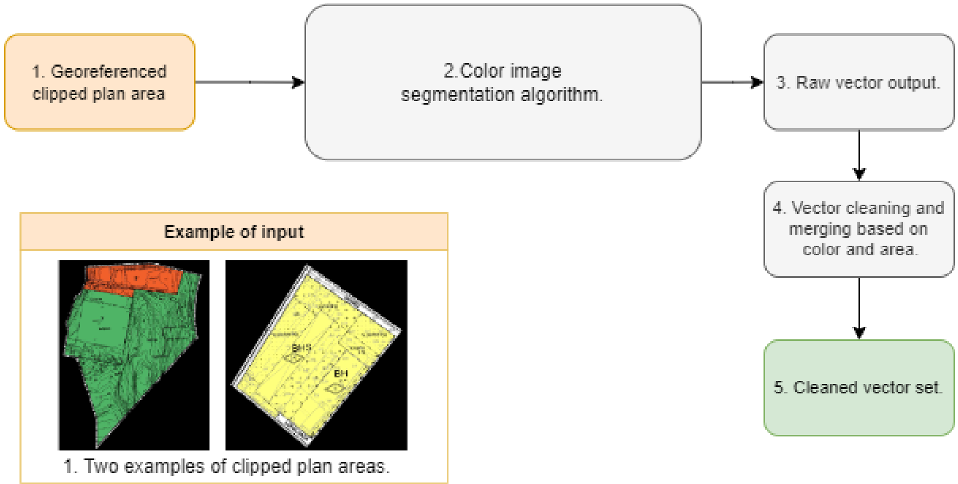

Figure 11. Illustration of Step 2.

Task 2 was a color image segmentation algorithm named K-means. The clipped plan area rasters were given as input to the algorithm, and it started by cleaning the rasters from lines and symbols to prepare them for the segmentation. The algorithm needed a predefined integer as input for how many clusters it should identify. The integer was the same as the number of colors in each plan, and it was specified in an attribute table that was given as input along with the rasters.

The algorithm then processed the color RGB-values for all pixels in each raster, and defined thresholds for how similar colors pixels should have to be clustered together based on the pre-defined number of clusters that had been given as input. Each group of pixels which the K-means algorithm had identified as having a coherent color was then clustered, with the aim of minimizing the in-variance of each cluster. After this, the clusters were vectorized, and the average RGB-value for each vector was annotated as an attribute. The clustering algorithm provided a raw vector output from this task. This vector output consisted of a large number of small polygons assigned with RGB-value attributes. The large amount was due to the fact that the color areas in the plan maps were usually not represented consistent enough for the program to identify them as single solid colors.

Therefore in Task 4, the vectors were merged together based on closeness, color and area to reduce their number and create larger coherent polygons. A number of vector operations compared the RGB-value attributes of neighboring polygons, and merged them if the values were similar enough. The new polygons inherited the mean RGB–value of the merged polygons. After these operations, polygons that did not pass a lower area limit were sorted out, and the rest were kept, which is illustrated in Figure 12 below. After this Step 2 was finished, and the output was a vector polygon set with RGB-value attributes.

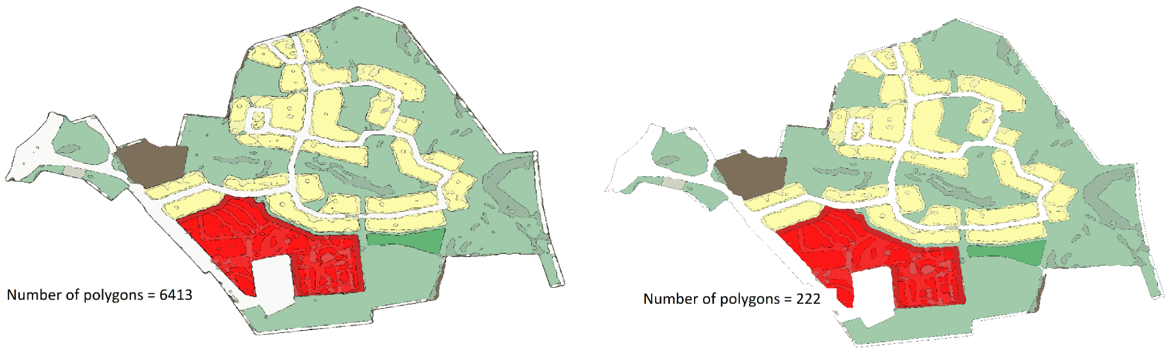

Figure 12. To the left, an example of raw vector output from the color clustering program. To the right, the same vector set after cleaning and merging had been performed.

1.5.2.3 Step 3

Finally, the third step consisted of text and symbol identification. The components of this step were a Optical Character Recognition (OCR) to identify text, a neural network trained with typical detailed plan symbols to identify markings, and a combined method of vectorization and vector grouping to identify dotted ground. The purpose was to identify map symbols that provided further information about what regulation areas were present in each vectorized part of the plan map.

1.5.3 Results and past evaluation of the method

The ADDP method has been presented along with initial results in the POC by Sweco (2020). Here the method was tested on a set of detailed plans and the result was then evaluated by visual inspection. This is the only evaluation of this method that has been presented yet, which means that the performance of the method has not been quantitatively evaluated. For the municipalities, it is relevant to know how well the method performs the task of digitizing, as well as what weaknesses there are of the method. In the conclusion of the POC, a number of factors identified as influencing the result of the method (Sweco, 2020):

- Number of colors: Whether the plan map is multi color or monochrome.
- Map layout: Whether the plan map area is divided into several map pages or not.
- Area: The true size of the detailed plan area, in terms of square meters.

- The map size relation.
- Resolution and scanning quality of the plan map.

1.6 Scope limitations

In order to evaluate the performance of the method, manually digitized plans were used as ground truth for comparison. As the manual digitizing was a time consuming process, all three steps could not be evaluated within the time scope of this study. Therefore, the performance of only the first two steps were evaluated. As the third part of the method was not evaluated within the scope, it was only possible to evaluate the identification of the regulation areas that were based on color. This excluded many regulation areas on detailed plans that were marked with symbols, lines or as dotted areas. This limited the insight from this study in how well the overall method performed.

1.7 Research questions

The purpose of this study is to quantitatively evaluate the performance of the first two steps of the ADDP method. Furthermore, the aim was to identify what limitations there are to the method, and how they influence the result. Finally, the factors that were identified in the study by Sweco (2020) as influencing the result were compared to the result of this thesis, to find out whether the conclusions agree on this point. Three research questions were prepared in order to reach these objectives, and the aim of the thesis was therefore to answer:

- Is the ADDP method effective for digitizing detailed plans?
- Are there any weaknesses of the method?
- Do any of the factors that were identified in the previous study (presented in section 1.5.3) influence the quality of the result?

2. Literature review - digitization of maps and evaluation methods

This literature review aims to give an overview of the current state of digitization of detailed plans in Sweden and directives concerning this question, as well as the general issue of digitizing maps and how the results from map digitization have been evaluated in related work.

2.1 Map digitizing in related work

Uhl & Duan (2021) explained the term map processing as an interdisciplinary field which covers research areas such as geoinformatics, computer vision, cartography and geographic information science to extract information from scanned maps. The purpose is to convert physically stored geographic information into digital data that can be stored and processed by computer software. Extraction of map content includes both recognition and extraction of features such as lines, labels, single map symbols and composite symbols. Chiang et al. (2014) identified a number of digital map processing techniques, dividing them into three major categories, namely separation of raster layers, georeferencing and extraction of map content. Separation of raster layers refers to the process of dividing the map into segments based on color, commonly known as color segmentation.

Different aspects of map information extraction has been explored, and some other recent examples are Aurelie & Jean (2021), who extracted linear features to georeference old maps without annotations and Ciolli et al. (2019) who used an object based image analysis to semi-automate the text identification and segmentation of cadastral maps. The potential of map text content extraction has also been further investigated by Chiang et al. (2017).

2.1.1 Challenges of digitizing maps

There are some main challenges associated with digitizing scanned maps which have been thoroughly described in literature. One issue that was summarized by Liu, Xu & Zhang (2018) is the distortion or bleaching of physical maps due to age, printing settings, archivation or inappropriate handling. This distortion leads to graphical quality issues that are also transferred to digital format when the maps are scanned. Additionally, as pointed out by the authors, the scanning image resolution affects graphical quality, and there is a trade off between image resolution and file size if the storage space is limited. Large files also significantly impair running times for digital processing, while on the other hand too low resolution results in pixel distortion effects of the map.

Chiang et al. (2014) noted that the complexity of maps compared to other types of documents makes applying techniques for graphics identification or document analysis challenging. Usually maps are created using consistent rules for visualization, but these rules vary greatly between maps and drawn map features are commonly not clear enough for existing document recognition methods. For this reason, most methods for digitizing maps have to be tailored to fit the particular map to be analyzed and its specific properties, and the authors highlight that there is an absence of general approaches that can be applied to a wider variety of maps.

Manzano-Agugliaro et al. (2013) pointed out a temporal aspect of challenges associated with this subject, namely that old maps can have outdated map features if the map area has been rebuilt or the land use has changed since the map was created. Because of this, it might be hard to match map features for georeferencing when the data sources have a significant age gap. For the same reason, it can also be challenging to identify what coordinate reference system has been used to create the map. Sometimes it is not specified at all or has a low accuracy because of the methods used in that age (Lelo & Baiocchi, 2014). Different methods have been developed to estimate coordinate reference systems for old maps and reproject them when this is unknown (Bayer, 2016) (Bayer & Kočandrlová, 2018) (Janata & Cajthaml, 2020).

2.1.2 Georeferencing

Georeferencing describes the process of projecting a raster image to a coordinate reference system, and an overview of methods for this purpose have been presented by Hackeloeer et al. (2014). They explained that this makes it possible to arrange the rasters in relation to other objects that also have a spatial reference, for instance to create a raster map mosaic. Furthermore, older maps that have been scanned need to be georeferenced in order to be useful for GIS analyses, and this can be done in different ways depending on the type of map and what information it contains. If the map is annotated with some kind of ground control points (GCP), these can be identified and used for projection. Otherwise, the normal approach is to match the scanned map with a georeferenced map covering the same area, to transfer the spatial information from the latter.

Uhl & Duan (2021) have identified a number of procedures which together constitute the process of digitizing maps. Automated georeferencing is here divided into two types, which are called metadata-based and matching-based approaches. Metadata-based approaches, as described by the authors, can be applied when there is coordinate information associated with the map intended for georeferencing. Usually this means that the map has been annotated in advance with GCPs that have known coordinates. Using GIS software, the GCPs can be placed manually on the map, which is then transformed into the desired coordinate reference system. This approach has been used to georeferenced historical maps by Brovelli & Minghini (2012) and Molnár (2010), among others. Some methods have also been proposed for automating the annotation of the points. For instance, Burt et al (2020) have developed a method for automatically identifying the GCPs in topographic maps using spatial computing and pattern-matching.

Matching based-approaches aim to match map features to another already georeferenced map, this can be geometric features such as building shapes or road network (Chen et al., 2004), or semantic features such as geographic names in the map texts (Wolter et al., 2017), (Luft, 2020), (Arriaga-Varela & Takahashi, 2019). Yan et al. (2017) used a polygon-based registration approach where building shapes in a cadastral map are used as control primitives and matched to a newer georeferenced map of the same area. Zhang et al. (2020) used another feature match method where the polygonal objects are first converted to a graph before the matching and georeferencing.

The ADDP method could be considered a matching based approach as it is based on matching the shape of the plan border polygons to the scanned map. In this way, it was similar to Yan et al. (2017)'s method of matching building polygons to old cadastral maps.

2.1.3 Evaluation of georeferencing methods

Evaluating a performed georegistration of a map is usually done by visual inspection (Hackeloeer et al, 2014). But to be able to compare and evaluate the result quantically, it needs to be based on numerical computations. Zitova & Flusser (2003) described a number of approaches to image registration, and presented a corresponding set of methods for accuracy assessment. They did acknowledge that accuracy assessment of georeferencing is a non-trivial problem, as a number of error sources from each registration step can affect the result. Additionally, it can be difficult differentiating between georeferencing errors and actual differences between the map content. The authors described three main error sources for methods using GCP matching. Firstly localisation errors, concerning the identification of reference features such as control points in the wrong location. Secondly, matching errors which occur in the matching step when false matches of control points arise. Generally, they stated, the aim is to avoid these two types of errors as the georeferencing cannot be considered correctly done if they are present. The third error type is called alignment error and is the most relevant when evaluating the quality of a successful georegistration. Alignment errors are always present to some degree as the transformation parameters are not calculated precisely enough, and the transformation model chosen might not correspond 100% to the maps.

According to Zitova & Flusser (2003), there are several ways to measure alignment errors, where the simplest one is control point mean square error (CP-MSE). As this method only evaluates how well the translation fits points to each other, there is a risk for overfitting if the control points were also used in the registration process. This is a well-known problem in numerical analysis that occurs when choosing too many degrees of freedom for a model. A large CP-MSE can also be an indication of a control point localisation error, which does not necessarily equate to poor overall alignment. A similar method which was also described by the authors is the test point error method (TPE). Here, some control points are excluded from the fitting procedure and their alignment is then used as a sort of control group for accuracy evaluation. This error estimation method eliminates the risk for model overfitting, but still holds the risk of being affected by localisation errors. It is also based on the presumption that there are enough control points that the fitting procedure can be done with good results even with a share of the points excluded.

Some methods for evaluation that are relevant for this study will be described further. For instance, Luft & Schiewe (2021) have used an evaluation method where they annotated the corners of the original map, and then interpolated their true spatial position based on reference data. By the help of Multi-scale template matching they could then determine the map corner position of the map that had been georeferenced by the help of their method. The geodetic distance between reference corner position and its actual location was then calculated, and root mean square error (RMSE, see Equation 2) was computed. RMSE is calculated as the square root of MSE (Mean Square Error), the latter which is computed using **n** as number of data points, **Y** as observed value vector and $\hat{Y}$ as predicted value vector, as seen in Equation 1.

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2$$ *Equation (1)*

$$\text{RMSE}=\sqrt{MSE}$$ *Equation (2)*

Yan et al. (2017) have proposed a, for the GIS context, novel accuracy assessment model, which they retrieved from the medical image registration scientific area where it is commonly used. It is called Dice-coefficient (DSC or Dice) and quantifies the overlap between two areas (see Equation 3). The Dice-coefficient has a range from 0 to 1 where 1 indicates perfect overlap and 0 indicates no overlap, in other words, a high score indicates better result. This coefficient is calculated for two partially overlapping areas **A** and **B**. The union of both areas is multiplied by a factor of 2 and divided by the sum of both areas.

$$\text{Dice-coefficient} = \frac{2\left|A\cap B\right|}{\left|A\right|+\left|B\right|}$$ *Equation (3)*

2.1.4 Color image segmentation and vectorizing of maps

In order to convert a scanned map to GIS data, one major focus is usually to convert and divide the raster image into vector surfaces. Chiang et al (2014) have observed that man made maps, compared to the remote sensing field where satellite imagery forms a complex raster image, usually have been substantially simplified. This has usually been done in order to depict the relevant information, such as land cover or cadastral structures. The simplified illustrations in the form of lines and color surfaces are visually similar to GIS vector data, and several different approaches to vectorizing scanned maps have been developed, usually by segmenting the color surfaces.

Color image segmentation, which is the main operation in part two of the ADDP method, starts with segmenting the map into homogenous pixel clusters based on color (Chiang et al., 2014). Some recent examples of map vectorization are Auffret et al. (2017) who have performed color image clustering on historical maps using HistMapR. Furthermore, Ståhl & Weimann (2022) have used neural networks to identify land use areas in historical maps, and Herrault et al. (2013) have used unsupervised classification to extract forest areas from similar data.

K-means, which is used in ADDP, is a simple algorithm that is one of the most commonly used unsupervised methods for color image segmentation, as described by Dhanachandra, Manglem, Chanu (2015). It requires a predefined whole number of clusters as input, and is able to handle large values for K. The algorithm works by picking the given number of cluster centers randomly over the set, then the clusters are computed iteratively.

When the vectorization process has been implemented, the result is a set of vectors, and commonly it needs to be matched to another corresponding vector set for accuracy assessment or georeferencing (Xavier, Ariza-López & Ureña-Cámara, 2016). This is commonly called geospatial data matching, linking or alignment, and has been used for investigating land use- and landscape change (Costes, 2014), to merge attributes from one dataset to another (Fan et al., 2016) and to evaluate the quality of a vector set by comparing it to ground truth data (Ai et al, 2014), (Fonte & Martinho, 2017), to give some examples.

Xavier, Ariza-López & Ureña-Cámara (2016) have compiled findings within this field of research, presenting a taxonomy of similarity measures consisting of geometric, semantic, topological, attribute and context similarities which can be used for matching data sets. They highlighted the close connection between choice of similarity measure and the matching method, as the first influences the output of the latter. In other words, depending on what types of similarities are considered, widely differing matching results can be achieved. Geometric similarities were described as the most commonly used, and consist of: area overlap, measures for distance, geometric properties and shape. Attribute measures refers to non-spatial attributes such as text, numbers or lists. Furthermore, another important aspect of the matching process is how the sets are matched in terms of numbers, for instance one-to-one, one-to-many etcetera, which is referred to as case of correspondence by the authors.

Saalfeld (1998) has presented a simple one-to-one method for geospatial data matching, where the aim is to georeference a vector set by using rubber sheeting to align it to another vector set. This method is based on the following approach: first test criterias are defined. The test criterias will then be applied to the geospatial data pairs to decide if they match, and they can be attributes such as spatial position, number of intersections, address or area. They can consist of number or character string attributes that are linked to each object in the data sets. Then a combination of test criterias is decided to be able to narrow the selection. In the iterative process of matching the data sets, different combinations of test criteria for matching can then be used in each iteration. The strongest criterias should be applied first, whereas ambiguous criterias can be applied last as they only give an indicator of similarity. Based on this the two datasets can be iteratively aligned to each other, according to the author.

2.1.5 Evaluation of color image segmentation and vectorization of maps

The evaluation of similarity between two vector sets has been described by Xavier, Ariza-López & Ureña-Cámara (2016), who states that Recall is a common measure for this purpose. Recall as presented by Rijsbergen (1979) evaluated both non-matches, compared to the full vector set. Some have also suggested further developments of this equation, such as assigning different weights to a parametrized version called F-measure, to be able to adjust what influences the result the most (J'erome Euzenat et al., 2009) (Do & Rahm 2002).

$$\text{Recall} = \frac{correct}{unmatched + correct} \qquad \textit{Equation (4)}$$

For calculating spatial overlap between vectors, Dice-coefficient and Jacquard constant were suggested by Zhang (1996). The two methods were further described by Taha & Hanbury (2015) who divided commonly used methods for evaluation of color image segmentation in medical science into six different categories. Here the use of confusion matrices was presented, describing the relation between false negatives, false positives and vice versa. Dice-coefficient and Jacquard constant are two of the most commonly used measures in this context.

When compared, Jacquard is providing a larger number in all cases except for the extremas (0 and 1) where they are equal. As the both measures are correlated to each other, the authors conclude that it is not meaningful to use them both. For evaluating color image segmentation, in general there is no gold standard method, as mentioned by numerous authors (Lucchese & Mitra, 2001) (Zhang, 1996) (Wang et al. 2020). Instead, the evaluation method of choice should be based on the needs of the specific experiment or study.

2.2 Automated digitization of detailed plans

The issue of automated digitization of detailed plans has been discussed before in a Swedish context. A pilot study made by Lantmäteriet (2019) stated that out of the existing detailed plans, a large number are possible to digitize automatically by the help of AI. However, as stated in the study, older hand drawn detailed plans cannot be processed in this method, and a presumption for the digitizing process is that both metadata and georeferenced plan borders are available. In the study, a digitized detailed plan is defined as having two dimensional vector polygons representing the total plan area as well as the regulation areas. Furthermore, the polygons should be georeferenced and associated with attributes such as data quality, regulations and period of validity.

SWECO (2020) published a report in conjunction with the digitizing method they presented. The first two steps of the method were evaluated through visual inspection using 61 detailed plans, and they stated that 67% of the plans could be digitized successfully in this case. In the report it was not specified what factors were considered when implementing the evaluation, and neither what defined a successful digitization. They highlighted the challenge of finding a method that performs well on different types of detailed plans, in terms of size, image resolution etcetera, and suggested that further tests should be done to explore techniques that can be applied to improve the results.

Some of the mentioned suggestions were to implement morphological transformations for image analysis to improve vectorization performance. Additionally, automatic identification of the plan legend using color clusters, to be able to perform OCR text identification for regulation areas present in the legend. Finally there was one suggestion to create a web application for more effective collection of the detailed plan metadata.

2.2.1 Challenges of digitizing detailed plans

The challenges of digitizing detailed plans have been thoroughly described by Örebro Kommun & Smart Built Environment (2019) in their handbook for digitizing detailed plans.

They stated that aspects such as the judicial interpretations of map design and knowledge about the background of map creation need to be considered in the digitizing process. This includes delicate assessments of past interpretations of the plan map, corresponding cadastral register maps and possible discrepancies between these two, together with changes and additions that have been made to the detailed plan since its establishment.

For instance, plan borders can have been interpreted differently (or incorrectly) in the past, resulting in placement of buildings on dotted ground, or structures that are technically outside of their dedicated regulated area. One question that arises in such a situation is if the digital plan borders then should be adjusted to interpretations like this that have been made in the past, to avoid future misunderstandings when buildings are found to be in the "wrong" area, even though their placement once have been sanctioned by the municipality.

The authors concluded that it is not always possible to base the digitization solely on the plan documents, as property formation and past plan interpretation can also need to be considered. This can be an opportunity to compile and screen out the relevant information about the detailed plans, to simplify for further use and interpretations. Nonetheless, these estimations are if not impossible then at least much more challenging to automate, compared to for example georeferencing. Consequently, this limits what aspects of digitization are possible to automate.

Moreover they stressed out that digitized detailed plans can give a false impression of accuracy. A raster map without geographical information has plan borders whose exact location are open for interpretation, whereas a digital vector dito has an exact geolocation. Depending on the quality of the data and digitizing method, the result will also have varying accuracy.

One significant example that the authors highlighted is how some plans have been created in outdated local coordinate reference systems. Conversations from them to up-to-date systems such as SWEREF99 can result in warping of lines. When plans are digitized, errors in the process or data might also result in overlaps or gaps in between plans that are intended to be placed edge to edge. Likewise, the detailed plan might be incorrectly georeferenced, or the shape of the regulation area might be wrongly translated to vector format. According to the authors, if this possible uncertainty is not clearly communicated to the users of the information, the digital format can be deluding as it creates a false sense of exactness. This illustrates the double edged consequences of digitizing detailed plans - it opens up for new possibilities of exactness when presenting map information, but also leaves no room for interpretation when the basis of information in itself is inexact or uncertain.

2.2.2 Accuracy specifications for digital detailed plans

Detailed plans started before 2022-01-01 that have been digitized have some directives concerning accuracy. The national specifications for detailed plans (Lantmäteriet, 2022) stated that these plans should not have gaps or overlaps between plan polygons that are intended to lie edge to edge. Acceptance limit of such gaps or overlaps are 0,25 meters or less measured in perpendicular distance between the polygon edge lines. Conversely, detailed plans that do not share borders should have a distance between their polygons of at least 3 meters measured in the same way. If this condition is not fulfilled, they have to be controlled manually to investigate if the locations are correct.

Regarding regulation area polygons, the same specifications (Lantmäteriet, 2022) stated that they should also have no overlap or gap of more than 0,25 meters for polygons that are intended to lie edge to edge. Overlaps of less than 5 square meter area are regarded as incorrect. If the polygons on the other hand are not intended to share borders, they should have a distance of at least 2 meters from each other, measured in perpendicular distance between the polygon edge lines.

The specifications also stated that location uncertainty of detailed plans that have been digitized based on scanned analogous maps should be calculated based on the directions of the handbook HMK-Digitalisering 1998, chapter 3.2. The handbook (Lantmäteriet, 1998) determined that location uncertainty for digitized objects from analogue maps should be calculated based on accuracy of the original survey measurements that the map is based on, as well as the accuracy of mapping the measurements, map scale and accuracy of digitization. All these factors were weighed together in the equation for total mean error (Equation 5) below, which was presented in the handbook:

$$\textit{Total mean error} = s_{tot}^2 = s_u^2 + \left(\frac{m_k}{1000}\right)^2 \cdot \left(s_k^2 + s_d^2\right) \qquad \textit{Equation (5)}$$

In Equation 5, s_{tot}= total mean error.

s_u= Mean error of original survey measurements on the ground (in meters).

s_k= Mean error of mapping the measurements (in millimeters).

s_d= Mean error of digitizing process (in millimeters).

m_k= Scale factor of map.

The handbook (Lantmäteriet, 1998) suggested the following standard values for estimations based on photogrammetric surveying: s_u = 15% of flight height, and s_d = s_k = 0.15 mm on the map. For geodetic measurements, these values vary depending on how old measuring equipment has been used.

3. Methodology

This section describes the method chosen for the thesis. It consists of a pilot study that was conducted to test out the performance of the ADDP method on detailed plans from a chosen municipality. The study was performed by applying Step 1 of the ADDP method, which is automatic georeferencing, and Step 2, which is automated vectorizing, on a set of detailed plans. Then the output of each step of the method was evaluated independently, and the output was also compared to manually digitized reference data using a set of evaluation procedures. Arrangements were also made in effort to eliminate consequential errors from being transferred between the first and second step, influencing the result of the latter. This meant Step 2 was also performed on manually prepared reference data. In this way, both the overall method performance seen in Step 2 and the performance of each step could be independently evaluated.

3.1 Study area and Data

In this section, the data selection, preprocessing and choice of study area are presented. The data quality is also briefly described based on the information that was available regarding this.

3.1.1 Study area

Detailed plans from Upplands Väsby municipality were used as the base for this study. The municipality was chosen mainly because the relevant data was easily accessible, and as the municipality had already digitized their detailed plans manually, the ambition was to be able to use this as ground truth data for the evaluation, even if this turned out later not to be possible. Additionally, the ambition was to cover a variety of map types that were representative for a municipality in need of digitization, to create an overview of how the procedure performed depending on the different characteristics of the maps.

Upplands Väsby provided a wide range of detailed plans, and the selection was made to fit in as many different factors as possible that could influence the result. Out of the 75 maps, 18 were monochrome and 9 had a multi-page layout, as seen in Table 2. This time span included plan maps of several different sizes, both in terms of raster resolution and detailed plan area. All maps earlier than 2006 were monochrome, and as this made it impossible for them to be vectorized if they contain multiple regulation areas, they were of limited interest for the study. The entire plan database is publicly accessible in Väsbykartan, which is Upplands Väsby municipality’s WebGIS, and from there all current detailed plans established between 2019 and 2006 were retrieved. As Väsbykartan had no function for displaying an overview of all detailed plans, the selection of maps used in the study was based on the maps that were present in Lantmäteriets plan border vector set.

3.1.2 Raw data

The raw input data for the study consisted of the detailed plan maps, and a vector set of plan border polygons. All detailed plan maps were retrieved in raster format, and the map legends contained information about coordinate reference system and number of regulation symbol colors which was information that was needed for the study. The other main data source was the vector set accessed from the layer 'plan_yta' in the dataset 'Markreglerande bestämmelser'. This vector set contained georeferenced polygons in the shapes of the detailed plan borders, and was retrieved from Lantmäteriets database. As the plan border polygons were georeferenced in correct scale, their areas corresponded to the detailed plan areas.

Monochrome plan maps	**Non-monochrome plan maps**	**Plan maps with multi-page layout**	**Plan maps with single-page layout**	**Maximum plan area (m^2)**	**Minimum plan area (m^2)**	**Minimum resolution of plan map (pixels)**	**Maximum resolution of plan map (pixels)**
18	*57*	*9*	*66*	*221855*	*154*	*501000*	*4277000*

Table 2. Characteristics of input data.

3.1.3 Preprocessing

In order to implement the ADDP method, some information about the plan map characteristics was needed as input. This map metadata was present in the map legends, and it was collected manually from each plan map and stored in an attribute table (see Table 3) during the preprocessing step. As the input detailed plan maps contained no information about the original survey method, the evaluation method presented in Section 2.2.2 could not be used in this study.

Metadata	**Coordinate reference system**	**Rotation of map**	**Plan ID**	**Map layout**	**Number of colors in legend**	**Map size relation**
Metadata value	ST74 SWEREF99 1800	0 90 180 270	Unique ID	Single-page Multi-page	0 (mono-chrome) and ascending	XS S M L

Table 3. Metadata collected manually from each plan map.

To be able to overview what factors influence the result, the detailed plans were divided into groups based on the following characteristics: map size relation, color, layout, coordinate reference system, plan area (which corresponds to the detailed plan area), and resolution of the raster map. In the evaluation, mean, median and standard deviation results were then calculated for each group. These characteristics were chosen to match the ones that had been identified as influential in the former study (see section 1.5.3). Out of these factors, four already had natural groups: color and coordinate reference system had two alternatives each, layout was divided into single-page and multi-page layouts and map size relation was divided into the four sizes, as seen in Table 3.

For the resolution of the plan map and the plan area, a manual division into groups was necessary to be able to compare if these factors influence the result. Quartiles were computed for the plan area groups, which meant that all values were sorted and then divided into four equal sized groups based on the numbers presented in Table 4. The reason for choosing quartiles was that the values were evenly distributed. For plan map resolution, natural breaks were chosen as most values were distributed around either 1000, 2000 or 4500. The delimitators and names of each group are presented in Table 5 below.

	Median	**Lower quartile**	**Upper quartile**
Plan area (m^2)	9593	3728	23797

Table 4. Group division of area factors.

Factor	**Group name**	**Lower limit**	**Upper limit**	**Number of detailed plans in group**
Plan area (m^2)	Group 1	0	3728	16
	Group 2	3729	9593	17
	Group 3	9594	23797	17
	Group 4	23798	-	17
Plan map resolution (100 pixels)	Group 1	0	1500	10
	Group 2	1501	2500	11
	Group 3	2501	-	46

Table 5. Delimitators and group names.

3.1.4 Accuracy

The input data contained some main error sources, limiting the achievable accuracy from the very beginning. The plan border vector set from Lantmäteriet had no location uncertainty stated in the product description, which means that the accuracy of the vector border position could not be considered. Additionally, as seen in Figure 13, the exact border positions were marked with a relatively thick line and it was not stated whether the borders should be interpreted as the center, outer or inner edge of the lines.

Figure 13. Example of border and 3 m outer border of detailed plan area in plan map.

3.2 Procedure

The procedure of evaluating both steps will be thoroughly explained within this section. Each of the steps were evaluated independently, and as the Step 2 used two different input data sets resulting in an equal number of outputs, they were, apart from the comparison to the manually digitized result, also compared to each other.

3.2.1 Step 1

In this section, the approach for evaluating the georeferenced maps produced by Step 1 is described. It consisted of first georeferencing the detailed plan maps using the automated method, and then a manual method, and then comparing the resulting maps to each other. The automated method was Step 1 of ADDP, where the plan maps were used as input, together with the plan border vector set and the Table 3 information to produce a georeferenced plan map. Then, the manual georeferencing was also performed on all the plan maps, to create a reference set of plan maps that can be used for comparison of the Step 1 output.

The manual georeferencing was performed in a GIS-software using control points. Most plan maps contained control point markings and coordinates which could be used for this purpose, and in the cases they did not, extracted coordinates from the corners of the plan border polygons were used instead (see Figure 14).

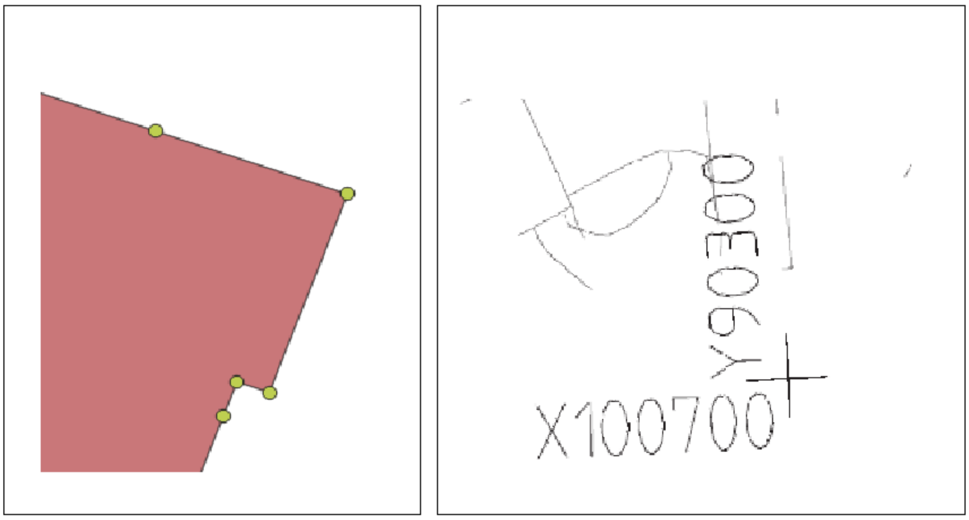

Figure 14. Left: Control points extracted from detailed plan border polygon corners. Right: control point cross marking and corresponding coordinates found in map.

Before evaluating the output of Step 1, all plans that for some reason could not be georeferenced correctly by the automated method were sorted out and presented separately in the result. A set of criteria for what maps should be removed in this sorting was defined:

- Layout: detailed plans that had the map on more than one page. This made georeferencing by ADDP method impossible as the plan area identification algorithm could not be run.
- Failed identification of plan area: defined as that the identified area fell completely outside of the true area (see Figure 15).
- Technical issues: defined as that the detailed plan could not have the method applied due to technical issues such as program crashing when georeferencing, no output produced by the algorithm and similar problems. This also included cases where the plan map could not be satisfyingly georeferenced by the manual method.

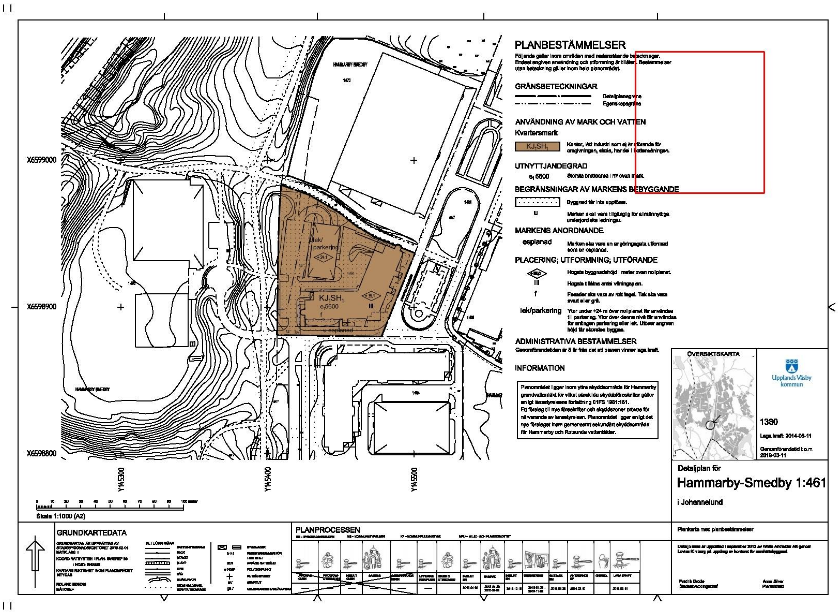

Figure 15. Example of failed identification of plan area. Red square marks the plan area identified by ADDP method.

After sorting out maps based on these criteria, preparations for the evaluation methods were carried out. Each map pair was matched based on plan ID, and the map sheet corners were annotated with reference points.

3.2.1.1 Output for Step 1

For Step 1, two rasters were produced as output. The georeferenced plan map was used for the accuracy evaluation of this step, and the cut-out plan area was the same raster that had been clipped along with the corresponding border polygon (see Figure 16 and Table 6). This clipped plan area raster was used in Step 2, so that the vectorization would only be performed on the actual plan area.

Output data step 1	**Description**	**Format**
Georeferenced plan map	The whole plan map raster that has been georeferenced.	Georeferenced raster
Clipped plan area	The georeferenced plan map that has been clipped by the detailed plan border polygon.	Georeferenced raster

Table 6. Output data for step one.

Figure 16. Left: Clipped plan area. Right: Georeferenced plan map. Images do not show true scale relation.

3.2.1.2 Evaluation of Step 1

To evaluate the first step of the method, the accuracy of the georeferencing procedure had to be investigated. To do this, the output georeferenced plan map was compared to the manually georeferenced plan map that will be used as a ground truth reference. This was done in two ways, using map corner annotation measured by RMSE (Equation 2) and map overlap measured by Dice-coefficient (Equation 3). Booth equations can be found in Section 2.1.2.1. As the georeferenced raster inherits the area of the plan border polygon, area was not a relevant factor to evaluate in this case.

RMSE provides an error value that can be numerically interpreted and evaluated in terms of distance. As the detailed plans varied in terms of area, this meant that small plans would be favored by the corner alignment evaluation. For this reason Dice-coefficient, which measures overlap of total map sheet areas, was also used as accuracy measurement.

3.2.2 Step 2

This section describes the approach for how the output vector set produced by the vectorization in Step 2 was evaluated. The automated vectorization aimed to create a polygon for each regulation area marked by color on the plan maps. This approach consisted of three parts, which are illustrated in Figure 17.

First, a vector set would be manually created to be used for comparison of the automated result. Then the automated vectorization was performed using the output data from Step

1 as input, which were the clipped and automatically georeferenced plan maps. The manually created and automatically generated vector sets were then compared to each other. Finally, the same procedure was repeated but with a different input raster set for the automated vectorization. Instead of using the output from Step 1, the manually referenced plan maps were clipped and used as input. In this way, the output was not influenced by the quality of output from Step 1. To summarize, two different inputs were used for the automated method, to create two different outputs. They could then be compared both to each other and to the manual method.

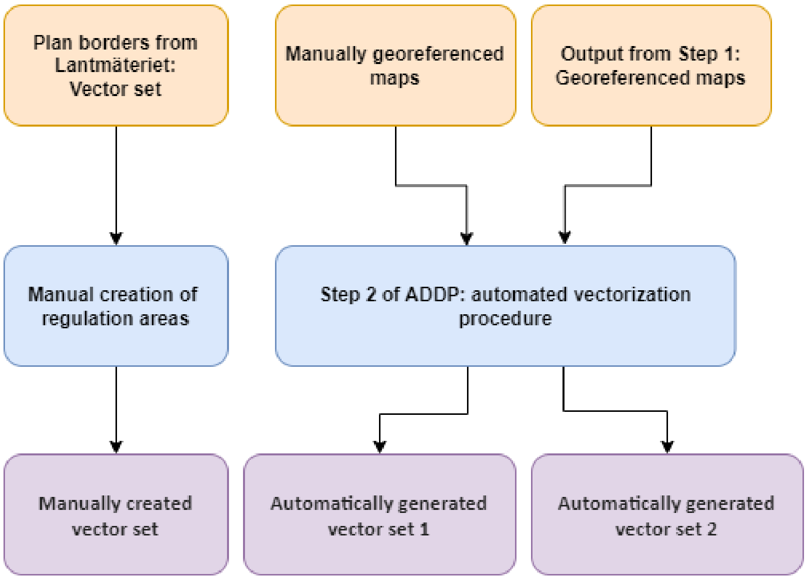

Figure 17. Creation of the three vector sets that were used for evaluating Step 2. The manually created vector set will be compared to the other two independently.

For the manual vectorization, the input plan border vector set was used together with the manually georeferenced plan maps to manually divide each polygon along the regulation area borders, as seen in Figure 18. Each regulation area polygon was annotated with a RGB-value representing the regulation color symbol, which was manually extracted from the map legend (Table 4).

Figure 18. Regulation areas were manually divided into individual polygons for use as reference data.

The attributes presented in Table 7 were annotated to both the manually created and the automatically generated vector sets, using different suffixes for the attribute names to differentiate between them. The RGB-values for the automatically generated polygons were generated during the Step 2 vectorization process.

Attribute	Content	Data type
Plan_ID	Unique name of each plan.	String
Polygon_ID	Unique ID for each polygon.	String
Area	Area of the regulation polygon.	Real
RGB	RGB-value for each regulation area. A number between 0 and 255 for each of three color bands (red, green, blue).	String

Table 7. Attributes of regulation vector sets for Step 2 evaluation.

3.2.2.1 Output of Step 2

The result from Step 2 of the ADDP method was a georeferenced set of polygons with the same attributes as in Table 7.

3.2.2.2 Evaluation of Step 2

In this section, the measures for evaluating Step 2 will be explained. The aim was to evaluate the results of step two of ADDP by comparing the outputs to a manually created vector set that was used as ground truth. Using spatial overlap analysis, and then performing a matching procedure based on a set of similarity criteria, a result was achieved. This result was then evaluated using two different methods which will be described in further detail below.

In Step 2 of the ADDP method, the output was a set of vector polygons. To evaluate how similar the automatically generated polygons are to the manually created ones, they first had to be matched to each other. The matching and similarity evaluation was performed on the individual polygon pairs. There was no matching based on the total similarity between the whole sets. As the automatically generated vector set contained more polygons than the manually created one, it was not possible to match all automatically generated polygons. Instead, the ones that did not clearly resemble any manually created polygon in terms of geometric and attribute characteristics (as defined below) were sorted out.

The characteristics that were used for similarity evaluation were the following:

- Spatial overlap: a geometric characteristic that was evaluated by two measures. The first measure was the number of polygons overlapping a specific area, starting from one and ascending. The second measure was Dice-coefficient, which gave a numeric score for how much two polygons are overlapping.

- Area relation: this was also a geometric characteristic to be evaluated by dividing the area of the reference polygon by the area of the generated polygon. In other words, it did not consider overlap, only total areas for each polygon.

- Color: this was an attribute characteristic expressed through RGB-values associated with each polygon. To evaluate whether the generated color deviated significantly from the reference color, a threshold value was set and each of three RGB color band values were evaluated based on this criteria.

- Attribute ID: this meant using the attributes *Plan_ID* respectively *Polygon_ID* to ensure the matching procedure was only performed on polygons belonging to the same plans, as well as that each polygon would have only a maximum of one unique match each.

Before starting the matching process, a spatial overlap analysis was performed of both vector sets together, where the condition was set so that only polygons with matching *Plan_ID* could overlap each other. The analysis produced a set of new polygons clipped along the overlap edges. Each overlap polygon inherited the attributes of all the overlapping polygons, and an attribute called *Overlap_area* was generated, representing the area of the overlapping vector. Additionally, Dice-coefficient was computed for each polygon pair according to Equation 3, and added as an attribute called *DICE*. To differentiate between attributes from the manually created polygons and the automatically generated dito, all attribute names were annotated with the suffixes *_manu*, for manual, or *_auto*, for automated.

3.2.2.2.1 Matching procedure

This matching procedure consisted of six tasks. The input of the procedure was a set of polygon pair matches that inherited attributes from both (or all) parent polygons. As each polygon could have a number of matches to other polygons, the following tasks were aimed to sort out the best matches and single out only one match per unique polygon in both sets. Task 1 to 4 are presented in Figure 19 below, and Task 4-6 in Figure 20.

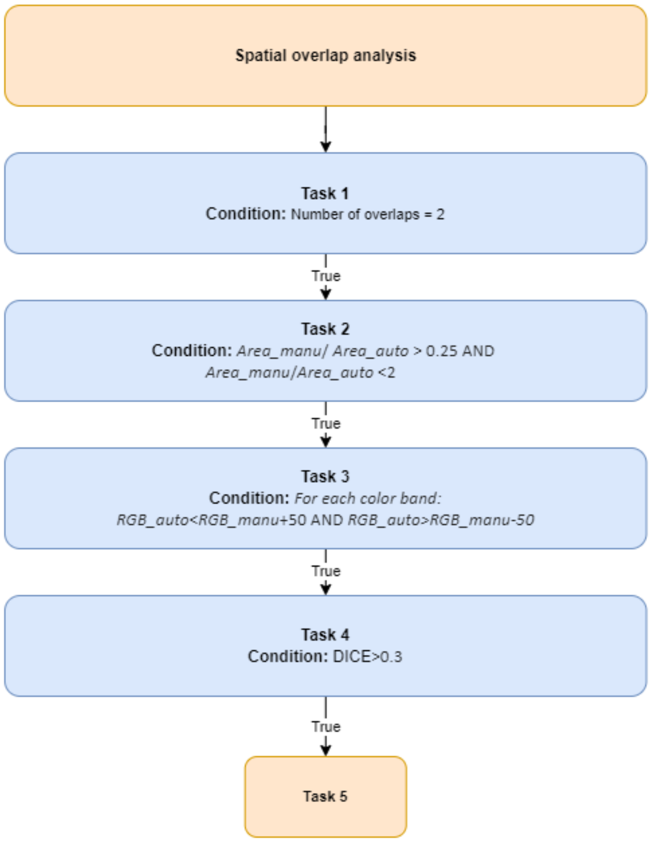

Figure 19. Task 1-4 of matching procedure. Only the polygon pairs that fulfill the condition of each step are kept.

The similarity evaluation was carried out starting with the strongest similarity indicators. Overlap and matching *plan_ID* had natural thresholds, but the other tasks would produce different matching results depending on the thresholds chosen. These had to be set based on the properties of the results of each case, and different limits were tested out to see which ones produced the most reasonable selection. The first four tasks did only act to sort out non-matches, and the final match was done in Task 5 and 6 by choosing which of the final polygons had the biggest overlap to the reference polygon.

Task 1 was to sort out all resulting polygons produced by the overlaps of two polygons. Polygons with no overlap were sorted out, together with overlaps that contained more than two polygons. As there were only two polygon sets that should contain no internal overlaps for each plan, triple overlaps were produced when the edges tangented each other due to inaccuracies in the automatic generation and were therefore excluded.

Task 2 consisted of computing the relation between areas for each polygon pair. Here, the threshold was set to 0.25 respectively 2, meaning that in the pairs, none of the polygons could be more than double the size of the other. The reason for choosing this threshold was that scaling errors in the georeferencing process sometimes resulted in inherited scaling errors in corresponding vectorization, and it was of interest to still be able to match the polygons even if differences in scale existed.

Task 3 aimed to identify if the color of the polygons matches. Each polygon was annotated with a RGB-value attribute, and they were compared within the pairs. Here, each of the three band values were evaluated independently, and the threshold was set to ±50 units from reference RGB value. This condition had to be fulfilled for all three bands in order to be considered passing. As the color code definitions of colors such as yellow, red and so on are subjective, it was challenging to define intervals for what would be considered "green", etcetera to evaluate how far a color could deviate before it was considered incorrectly rendered. Especially as the hues on most of the maps were muted, and the exact shades could also vary significantly between the maps, this threshold was chosen as a generic numeric interval of 50 centered around the reference RGB-value.

Task 4 acted as a threshold for excluding polygon pairs with a Dice-coefficient lower than 0.3. This number was chosen based on visual inspection of the result, where the polygons were deemed similar enough to be deemed relevant above this limit. The polygons that passed this threshold were passed on to Task 5, as seen in Figure 20.

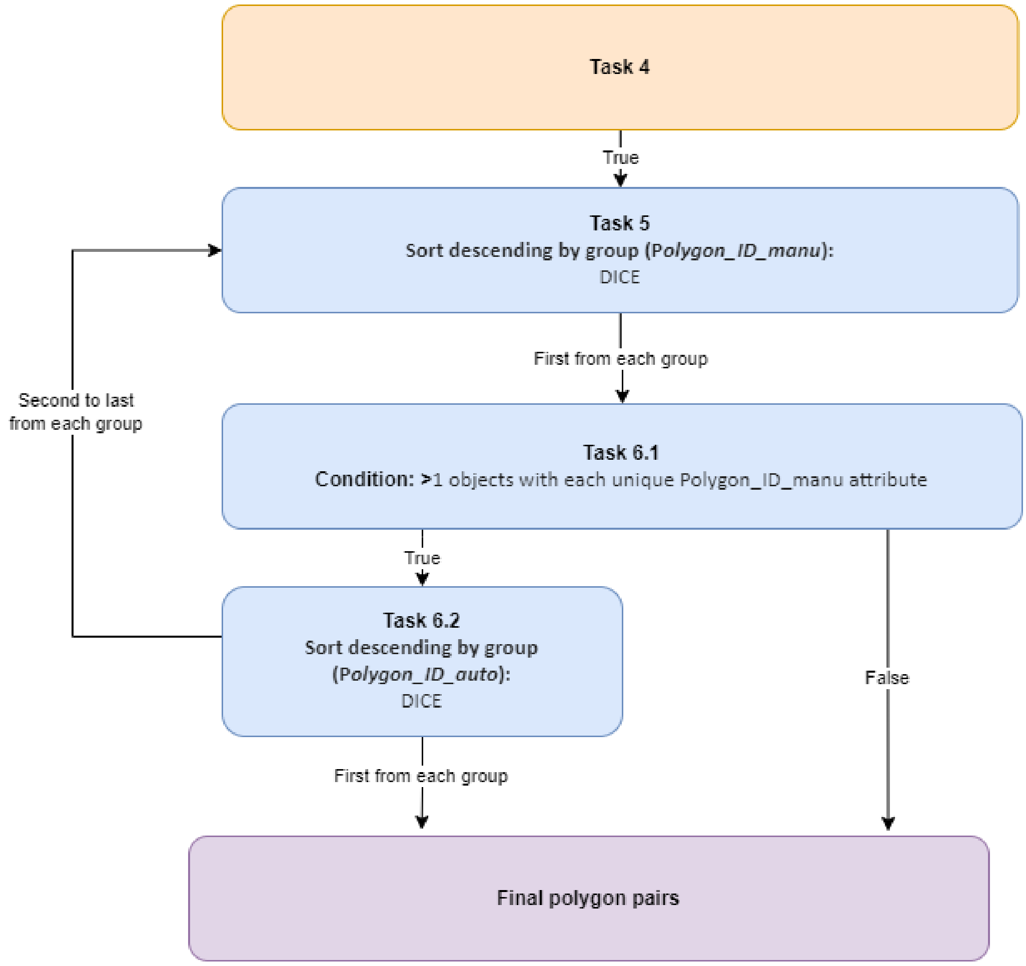

Figure 20. Task 5 and 6 in the matching procedure. Polygons that did not match the criteria for proceeding to any of the next tasks were sorted out.

Task 5 was also based on the Dice-coefficient, but here it was used to sort out only one match per reference polygon. If a reference polygon had more than one possible match, the match with the highest Dice-coefficient was kept. Similar to this is Task 6.1, which acted to make sure the generated vector set also only had one match per polygon. If a generated polygon was present in more than one matching pair, it was sent to Task 6.2 where it was once again ordered by Dice-coefficient, and the highest value pair was kept.

Then Task 5 and 6 were repeated for the corresponding reference polygons from the pairs excluded in Task 6.2. This time with the already used matches now removed, to find out if there were more matches that were rejected in the first Dice-evaluation. If there was a match to another generated polygon which passes all the criterias, they would then be added to the final result. This iterative process could be repeated as many times as necessary, until there were no unique polygon pairs left.

3.2.2.2.2 Evaluation measures for Step 2

The output from the matching procedure was a set of polygon pairs, where in each pair one polygon came from the automatically generated vector set, and one came from the manually created dito. Based on the matching procedure, these pairs had been identified as representing the same regulation areas in terms of color, size and location of the polygons. To evaluate how similar they were, their spatial overlap was presented using the Dice-coefficient generated during the matching procedure.

Additionally, three other measures were used to evaluate the performance of the automated vectorization implemented by Step 2. Firstly, Recall (Equation 5, see Section 2.1.3.1) was used to identify how many redundant polygons the method produced, by evaluating the number of matches in a set of objects. Each redundant polygon which did not correspond directly to a regulation area was considered unmatched. The manually created vector set was used as a basis for calculating the number of matched and unmatched values for the set.

Finally, the area difference between the polygon pairs, and the identified color of each automatically generated regulation area polygon were evaluated. This was done by computing the absolute value of the difference between the areas, respectively RGB attributes for each separate band (Red, Green, Blue) and polygon pair.

4. Result

In this section, the results of the experiments are presented. This consists of two sections, one for each of the steps that have been evaluated.

4.1 Evaluation of Step 1: Georeferencing

Using the evaluation methods chosen for Step 1, Dice-coefficient and RMSE were computed for each detailed plan. From this result total mean, median and standard deviation (SD) values were calculated and presented in Table 8 below, which also shows the share of detailed plans that could not be georeferenced. An overview of reasons for why the georeferencing failed can be seen in Table 9 below. Furthermore, the mean, median and SD of the Dice-coefficient and RMSE results for each group have been presented below in Table 10. The division into groups was based on the set of factors presented in Section 1.5.3, namely layout, raster resolution, plan area, map size relation, number of colors and coordinate reference system. Finally, to overview the dispersion between the groups for each factor category, SD was calculated for each category based on the results in Table 10. This can be seen in Table 11.

Detailed plans that could not be georeferenced	**Dice-coefficient**			**RMSE (m)**		
	Mean	**Median**	**SD**	**Mean**	**Median**	**SD**
28%	0,8740	0,9704	0,2081	34,09	10,06	62,11

Table 8. Overview of results from the evaluation of Step 1.

Number of plans	**Reason**
9	Map layout.
7	Failed identification of plan area.
3	Image search algorithm could not be run.
1	Other technical issues.
1	Faulty plan border polygon.
Total	**21**

Table 9. Summary of all plans that could not be georeferenced and identified reasons for this.

Factor grouped by	Group	Dice-coefficient			RMSE (m)		
		Mean	Median	SD	Mean	Median	SD
Plan area	Group 1	0,5783	0,6087	0,2997	114,4	53,38	103,2
	Group 2	0,9069	0,9700	0,1209	18,72	5,689	23,46
	Group 3	0,9561	0,9704	0,0327	14,98	7,670	7,670
	Group 4	0,9605	0,9800	0,0556	13,36	7,355	15,74
Resolution	Group 1	0,8208	0,8856	0,1488	29,69	25,23	18,79
	Group 2	0,8472	0,9753	0,2874	36,16	4,747	67,16
	Group 3	0,8920	0,9729	0,1968	34,76	7,986	66,90
Color	Monochrome	0,7541	0,9240	0,2966	78,48	30,232	100,5
	Multi-color	0,9091	0,9724	0,1541	21,09	7,3553	34,43
Map size relation	XS	0,6695	0,8762	0,3317	114,01	32,87	118,8
	S	0,9035	0,9685	0,1368	21,50	11,03	23,30
	M	0,8833	0,9740	0,1931	27,82	10,05	43,69
	L	0,9802	0,9811	0,0099	4,706	5,461	1,922
Coordinate reference system	ST74	0,8443	0,9352	0,2303	45,70	20,10	77,23
	Sweref 99 1800	0,9005	0,9735	0,1778	23,713	6,666	40,13

Table 10. Mean, median and SD of Dice-coefficient and RMSE for detailed plans in each group.

Standard deviations within each group, based on mean or median for the groups	SD of Dice-Coefficient values		SD of RMSE values (m)	
	Mean	Median	Mean	Median
Plan area	0,1585	0,1579	42,82	20,14
Resolution	0,0293	0,0417	2,777	8,990
Color	0,0775	0,0241	28,69	11,43
Map size	0,1152	0,0427	42,42	10,61
Ref	0,0281	0,0191	10,99	6,720

Table 11. Standard deviations for the group results of each factor.

4.2 Evaluation of Step 2: Vectorization

Step 2 was evaluated by vectorizing two different sets of input data, namely automatically georeferenced maps and manually georeferenced maps, and the results were then compared to each other. Monochrome maps could not be vectorized as the procedure is based on identifying color fields. The total share of maps that could be vectorized is presented in Table 12 below, and the number of input maps was determined by how many plan maps could be georeferenced in Step 1. To make the results comparable, also the number of manually georeferenced input maps were adjusted to match this number.

Number of input maps from Step 1	Monochrome maps	Share of maps that could be vectorized
54	13	75,93%

Table 12. Share of maps that could be vectorized in Step 2.

After the vectorization each output vector set was matched to a corresponding manually created vector set, and the similarities of the matched vector polygon pairs were evaluated using Dice-coefficient, a color offset measure, percentage and absolute value of the difference between polygon pair areas. The mean, median and SD of these results are presented in Table 13 below for the automatically georeferenced input, and Table 14 for the manually georeferenced input. Moreover, the recall of reference polygons that were matched to the generated polygons are presented in Table 15 for automatically generated input and Table 16 for manual input.

Dice-coefficient			Color offset			Area difference (m^2)			Area difference (percent)		
Mean	Media	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
0,7929	0,8505	0,1667	14,17	11	9,015	1712	440,1	5594	43.16	21.22	59.04

Table 13. Result from automatically generated input data.

Dice-coefficient			Color offset			Area difference (m^2)			Area difference (percent)		
Mean	Median	SD	Mean	Media n	SD	Mean	Median	SD	Mean	Median	SD
0,7961	0,8521	0,1612	14,49	11,17	8,911	1608	494,8	4514	40.07	21.63	52.88

Table 14. Result from manually generated input data.

Total number of reference polygons	Successfully paired polygons	Recall
228	99	0,4342

Table 15. Recall for automatically generated input data.

Total number of reference polygons	Successfully paired polygons	Recall
228	100	0,4385

Table 16. Recall for manually created input data.

A comparison was made between the output created with manually georeferenced input data and the one created with automatically generated input data, by arranging the results side by side. In Table 17 below, recall of the matching procedure is presented.

Input type	Total number of generated polygons	Recall
Manually georeferenced	776	0,1288
Automated georeferenced	810	0,1222

Table 17. Value of recall for both output vector sets of Step 2.

5. Discussion

In this section, the results from evaluation of both steps of the method are discussed based on the three research questions.

The first research question to be answered was whether the ADDP method is effective for digitizing detailed plans. Based on the evaluation results from Step 1 and 2, some observations can be made related to this question at issue. The results show that out of all plan maps, 72% were possible to georeference by the ADDP method, and 55% could be vectorized using the same method. This closely resembles the report from Sweco (2020), which had a 67% rate of successful georeferencing when evaluating this method. As that report contained no numeric result for how well the vectorization performed, it was not possible to compare the results from Step 2 to previous findings. Consequently, the results from this study extend our knowledge of the performance of this method for automated digitization of detailed plans. The results show that 44% percent of the regulation area polygons could be successfully vectorized by this method, with a mean Dice-coefficient of 0,79.

Concerning the accuracy for Step 1, the median RMSE for all detailed plans was 10 meter, but the mean was almost 34 meter, and the standard deviation was 62 meters. This implies that there were significant variations in accuracy between the plans. The same tendency can be observed for the Dice-coefficient, where the median Dice-coefficient for all detailed plans was 0,97, whereas the mean was 0,87 and a standard deviation of 0,21. As the results are on meter RMSE level with a median area difference of 440 m^2, the accuracy is not high enough to perform GIS-analyses. The output data might also be misleading if it is to be presented to citizens or officials, as digital maps can give a false sense of accuracy. The resulting detailed plans can still be used for visualization, but the significant error levels currently limit the possible uses for plan maps digitized by this method.

Problems that occurred during the implementation of the study were mainly regarding the first step, and more specifically the Multi-scale template match algorithm. Step 1 required extensive manual supervision as the performance of the algorithm was irregular and sometimes did fail identifying the plan area. The proposed map size relation that was used as input data for the plan area identification also did not provide the best identification in all cases, and in those cases different input sizes had to be tested out in order to find which one gave the best identification. As this could be necessary to repeat multiple times for some plans that then failed the plan area identification nevertheless, it became more time consuming that it would have been to georeference the plan map manually.

By visual inspection, it also became clear that the algorithm did not match the plan border polygon to the plan map more exact than what could also be done manually. For this reason, the ADDP method cannot be considered time efficient, even though it manages to georeference in an generally effective way. Based on these findings, one suggestion is to improve the georeferencing in terms of time efficiency and accuracy, by identifying plan borders manually in a digital interface. A similar suggestion has already been presented by Sweco (2020), and our findings provide further indication that this would be a relevant improvement of the method.

Regarding the evaluation of Step 2, the comparison between the result from manually georeferenced maps and the result using the output from Step 1 shows that the differences are miniscule. The manually created input, which is used as ground truth, displays only a minimal advantage, and the results are overall very similar. As the Results in Step 1 shows that the georeferencing has a shifting performance, this high degree of similarity is unexpected and indicates that the performance of Step 2 is at large independent of the performance of the previous step.

However, the number of detailed plans that could be vectorized is significantly smaller than the total number of maps used in the study, and this is mainly because detailed plans that failed to be georeferenced in Step 1 have to be sorted out, as the vectors otherwise will contain no spatial information. Additionally, only regulation areas that are visualized by color can be vectorized by this method. This limits what detailed plans can be digitized by the ADDP method, as the vectorization can only be performed on multicolor maps, and as older maps, in the case of Upplands Väsby - maps older than 2007, are generally not found in multicolor. Together these two factors effectively limited the number of detailed plans that could be vectorized to only 55%. Within the scope of this study it is not investigated how many of the detailed plans in sweden are in multicolor, but this would be relevant to know to estimate the total share of detailed plans that this method can be applied to.

The suggested method in Section 2.2.2 for calculating location uncertainty of detailed plans that have been digitized based on scanned maps could not be used. This is because the input data contained no information about the original surveying method for the plan map creation, which limits the insight in the accuracy of the overall digitizing process and how the quality of the plan map content impacts the result.

The second research question to be answered was whether the method had any weaknesses. The data from this study indicates that one weakness of the ADDP method is that it provides an output of excess polygons, and it does not identify what polygons belong to what regulation areas. For this reason it was challenging evaluating the result of Step 2, as it was necessary to begin by implementing an extensive matching procedure to determine which polygons should be compared to each other. As seen in Figure 12, some output polygons clearly resemble regulation areas, whereas other output polygons only resemble irregularities and impurities in the rasters. The evaluation of Step 2 is therefore also influenced by the quality of the matching procedure.

Furthermore, many regulation areas are marked with symbols, text, lines or dotted ground, which are features that the first two steps of this method do not consider. The third step of the method consisted of symbol and text orientation, but as that step was not evaluated within the scope of this study, there was no gained insight in what the possibilities are for digitizing regulations with other types of visual identificators.

Another identified weakness is that the ADDP method relies on the use of the detailed plan border vector set from Lantmäteriet as input. The quality of this vector set is not known, and has been digitized by each municipality using different methods and levels of accuracy. As the georeferencing procedure relies on the spatial data from this vector set, the accuracy of the result from the ADDP method is directly affected by the quality of this input data. This adds another error source to the digitization process which cannot be easily estimated. Another aspect of this is that it limits the possibilities for what detailed plans can be digitized, as the available input data needs to be not only the detailed plan map but also the corresponding plan border polygon. If this data is not available it first needs to be created, in itself a form of manual digitizing, and this raises the question of whether it will risk becoming double digitizing work.

The third research question to be answered was whether any of the factors (plan area, plan map resolution, color, coordinate reference system and map size relation) influence the result. Based on the standard deviations presented in Table 11 for the Step 1 evaluation, it can be observed that two factors have a large disparity between the statistical measures for each of its group. Plan area has the largest disparity, and in Table 10 it can be seen that this depends on Group 1, representing the smallest area, which has significantly worse results than the other groups. A similar relation can be observed for the map size relation-factor, where the group with the smallest map size relation has significantly worse results than the other groups. Coordinate reference system and plan map resolution factors display no major differences between the results for each group. For the color factor, a minor difference between monochrome and multicolor results can be seen, where the latter has worse outcome.

As the statistical measures are based on small groups with an average number of 12-17 detailed plans in each group, this increases the risk for the result to be influenced by outliers. Nevertheless, the above presented observations indicate that small plan areas and a small map size relation (in other words, a minimal illustration of the plan area) can have a negative impact on the result of the georeferencing. Especially for plan area category, where the Dice-coefficient and RMSE values improve for each group, a larger plan area is an indicator of better results.

6. Conclusions and future work

For this study, the first two steps of the ADDP method for automated digitization of detailed plans were evaluated. This was implemented by comparing the result output to manually digitized plan maps using a set of accuracy evaluation measures. Three research questions were formulated for this purpose: Was the method effective for digitizing detailed plans? Were there any weaknesses of the method? Did any of the factors that were identified in the previous study influence the quality of the result? In this section, the conclusions of the study and suggestions for future work are summarized.

6.1 Conclusions

The findings of this study indicates that the ADDP method is to some extent effective for digitizing detailed plans. It manages to digitize a majority of the detailed plans, but there is a significant share of the detailed plans that fail the digitization due to factors such as monochrome plan maps and failures of sub-procedures of the method. There is also a significant dispersion of the quality of the results, where some detailed plans are digitized with RMSE-values below 5-meter level, whereas others display RMSE-values on 100-meter level. The low accuracy in form of a large RMSE and area difference limits the areas of use for detailed plans digitized by the ADDP method. One unexpected finding was that the combination of the two steps of the method performed as effective as running Step 2 of with manually georeferenced input data, which indicates that the vectorization procedure is robust.

Four main weaknesses of the method were identified. The first one is that the georeferencing procedure requires extensive manual supervision and does not perform reliably. The second one is that the vectorization produces an excess of polygons, and the third that it does not identify what polygons belong to each regulation area. Finally, the method is only able to vectorize regulation areas that are visualized with color, which effectively limits the share of detailed plans that are possible to digitize. These results confirm previous findings, but also extends and deepens our knowledge regarding the effectiveness of the method.

Out of the evaluated factors, two displayed significant influence on the georeferencing output quality. The plan area had an apparent correlation with the accuracy measures, where the plans with the smallest area performed the worst. A similar relation was observed for map size relation, where detailed plan maps that had minor illustrations of the plan area compared to the total document size, demonstrated poor results in the evaluation.

One novel contribution of this study was to create a framework for evaluating the accuracy of digitization of detailed plans. As digital detailed plans will be increasingly common in the physical planning process, it will also be increasingly relevant to evaluate their accuracy. The methods for evaluation that have been used is a suggestion for what factors could be considered when digitizing detailed plans. The factors chosen for this were spatial overlap (measured by Dice-coefficient), offset (measured by RMSE), area difference and color offset (measured by RGB-value difference).

6. 2 Future work

Factors that were not considered, but that are also relevant for evaluating how well a digitization has been performed, are how well the correct shape of the polygon is translated and if the polygons fulfill Lantmäteriet's specification of maximum gap between neighboring polygons. Further research could include implementing additional measures for accuracy evaluation such as the ones mentioned, for evaluation of detailed plan digitization. Especially exploring how methods for evaluating shape can be applied to this issue could be a question for future research to make assessments of results more relevant. The suggested framework for evaluation of similarity between vector pairs could also be useful for other research fields, such as investigating urban- or land use change, verifying the quality of open source data for urban structures or digitizing historical maps.

To improve the performance of the georeferencing procedure, alternative methods for identifying detailed plan borders could be explored. This includes machine learning algorithms, to be able to utilize the homogenous visual characteristics of the detailed plan maps, but also manual identification if this can be done within an interface which minimizes the number of surrounding tasks that needs to be repeated for each matching.

The study was based on only 75 detailed plans because of the time consuming preprocessing of data and manual digitization process, and this, together with the fact that many detailed plans could not be digitized at all, limits the insight in the performance of the ADDP method. Especially when considering the results divided into groups by the factors such as resolution and plan area, where each group contained between 10 and 16 plans each. This is a comparatively minor sample for quantitative evaluations, and for this reason, it would be relevant to evaluate the method on a larger set of detailed plans in future research. Further on it could also be relevant to investigate the influence on quality of result of additional factors related to map characteristics, such as coordinate reference system, age of map, effect of "clutter" or whether the plan map contains multiple map illustrations.

7. References

Arriaga-Varela, E. J. & Takahashi, T. (2019). Automatic Georeferencing of Heterogeneous Historic and Illustrated Maps. *Abstracts of the ICA, 1*, s. 1–2. doi:10.5194/ica-abs-1-15-2019

Auffret, A., Kimberley, Ad., Plue, J.,Skånes, H., Jakobsson, S., Waldén, E., Wennbom, M., Wood, H., Bullock, J., Cousins, S., Gartz, M., Hooftman, D. & Tränk, L. (2017). HistMapR: Rapid digitization of historical land-use maps in R. *Methods in Ecology and Evolution, 8*. 10.1111/2041-210X.12788.

(2017). HistMapR: Rapid digitization of historical land-use maps in R. *Methods in ecology and evolution, 8* (11), s. 1453–1457. doi:10.1111/2041-210x.12788

Aurelie, L. & Jean, C. (2021). Segmentation of historical maps without annotated data. T*he 6th International Workshop on Historical Document Imaging and Processing. Association for Computing Machinery*, New York, NY, USA, 19–24. https://doi.org/10.1145/3476887.3476909

Bayer, T. & Kočandrlová, M. (2018). Reconstruction of Map Projection, its Inverse and Re-Projection. *Applications of Mathematics, 63*, 1-27. 10.21136/AM.2018.0096-18.

Bayer, T. 2016. Advanced methods for the estimation of an unknown projection from a map. *GeoInformatica, 20*, 241-284. doi 10.1007/s10707-015-0234-x

Boverket (2017). *Digitala detaljplaner*. https://www.boverket.se/sv/om-boverket/publicerat-av-boverket/publikationer/2017/digitala-detaljplner/ Retrieved 2022-05-01.

Boverket (2020). *Digitalisering av befintlig detaljplaneinformation*. https://www.boverket.se/sv/PBL-kunskapsbanken/planering/detaljplan/digitalisering/ Retrieved 2022-02-16.

Boverket (2021a). *Boverkets föreskrifter (2020:5) om detaljplan.* https://www.boverket.se/sv/lag--ratt/forfattningssamling/gallande/boverkets-foreskrifter-20205-om-detaljplan/ Retrieved 2022-02-16.

Boverket (2021b). *Planhandlingar för detaljplan*. https://www.boverket.se/sv/PBL-kunskapsbanken/planering/detaljplan/handlingar/ Retrieved 2022-02-16.

Boverket (2021c). *Begränsning av markens utnyttjande.* https://www.boverket.se/sv/PBL-kunskapsbanken/planering/detaljplan/planbestammelser/egenskapsbestammelser_/begransning/ Retrieved 2022-02-16.

Boverket (2021d). *Planbestämmelsekatalogen.* https://www.boverket.se/sv/PBL-kunskapsbanken/planering/detaljplan/planbestammelser/planbestammelsekatalogen/ Retrieved 2022-02-19.

Boverket (2022a). *Vad är en detaljplan.* https://www.boverket.se/sv/PBL-kunskapsbanken/planering/detaljplan/detaljplaneinstrumentet/vad-ar-detaljplan/ Retrieved 2022-02-19.

Boverket (2022b). *Planbestämmelser som ska och får användas.* https://www.boverket.se/sv/PBL-kunskapsbanken/planering/detaljplan/planbestammelser/att-reglera-med-planbestammelser/planbestammelser-som-ska-och-far-anvandas/ Retrieved 2022-02-10.

Brovelli, M. & Minghini, M. (2012). Georeferencing old maps: a polynomial-based approach for Como historical cadastres, *e-Perimetron, 7*, 97-110. http://www.e-perimetron.org/Vol_7_3/Brovelli_Minghini.pdf Retrieved 2022-02-10.

Chen, C.-C., Knoblock, C. A., Shahabi, C., Chiang, Y.-Y., Thakkar, S. (2004). Automatically and accurately conflating orthoimagery and street maps. *Proceedings of the 12th Annual ACM*

International Workshop on Geographic Information Systems (ACM-GIS), 47-56. https://doi.org/10.1145/1032222.1032231

Chiang, Y.-Y., Leyk, S., & Knoblock, C. A.. (2014). A Survey of Digital Map Processing Techniques. *ACM Computing Surveys, 47 (1)*, 1–44. https://doi.org/10.1145/2557423

Chiang, Y.-Y. (2017). Unlocking Textual Content from Historical Maps - Potentials and Applications, Trends, and Outlooks. *I: Artificial Intelligence Research* (111–124). Artificial Intelligence Research. doi:10.1007/978-981-10-4859-3_11

Costes, B. (2014). Matching Old Hydrographic Vector Data from Cassini's Maps. *e-Perimetron, 9* (2), 51-65. http://www.e-perimetron.org/Vol_9_2/Costes.pdf. Retrieved 2022-02-10.

Dhanachandra, N., Manglem, K., Chanu, Y., J. (2015) Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm, *Procedia Computer Science, 54*, 764-771. https://doi.org/10.1016/j.procs.2015.06.090.

Do, H.H., & Rahm, E. (2002). COMA - A System for Flexible Combination of Schema Matching Approaches. *VLDB.* 610-621 doi:10.1016/B978-155860869-6/50060-3

Fan, H., Yang, B., Zipf, A. & Rousell, A. (2016). A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. *International journal of geographical information science, 30* (4), s. 748–764. doi:10.1080/13658816.2015.1100732

Fonte, C. C. & Martinho, N. (2017). Assessing the applicability of OpenStreetMap data to assist the validation of land use/land cover maps. *International journal of geographical information science, 31* (12), s. 2382–2400. doi:10.1080/13658816.2017.1358814

Herrault, P.-A., Sheeren, D., Fauvel, M. & Paegelow, M. (2013). Automatic Extraction of Forests from Historical Maps Based on Unsupervised Classification in the CIELab Color

Space. *I: Intelligent Systems for Crisis Management*. 95–112. doi:10.1007/978-3-319-00615-4_6

Janata, T. & Cajthaml, J. (2020). Georeferencing of Multi-Sheet Maps Based on Least Squares with Constraints—First Military Mapping Survey Maps in the Area of Czechia. *Applied Sciences, 11* (299). doi:10.3390/app11010299.

Euzenat, J., Ferrara, A., Hollink, L., Isaac, A., Joslyn, C., et al. (2009) Results of the ontology alignment evaluation initiative. *In Proceedings of the 4th ISWC Workshop on Ontology Matching*. (p. 73–126).

Lantmäteriet. (n.d.). *About Us*. https://www.lantmateriet.se/en/about-lantmateriet/about-us/ Retrieved 2022-02-18.

Lantmäteriet. (1998). *HMK-Digitalisering*. https://www.lantmateriet.se/globalassets/om-lantmateriet/var-samverkan-med-andra/hmk/gamla-hmk-serien/hmk-digital.pdf

Lantmäteriet. (2022). *Nationell informationsspecifikation Detaljplan 4.0.* https://www.lantmateriet.se/globalassets/smartare-samhallsbyggnadsprocess/nationella-specifikationer/natspec-detaljplan-v4.0.pdf Retrieved 2022-02-18.

Lantmäteriet. (2019). *Nationellt tillgängliggörande av digitala detaljplaner.* https://www.regeringen.se/4ad5e8/contentassets/77c44592b3df48cdaecdac78d9644d3c/nationellt-tillgangliggorande-av-digitala-detaljplaner.pdf Retrieved 2022-02-12.

Lelo, K. & Baiocchi, V. (2014). Assessing the accuracy of historical maps of cities: Methods and problems. *Città e storia, IX*, 61-89. doi:10.17426/18984

Liu, J., & Yang, Y. H. (1994). Multiresolution color image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 16*, (7), 689-700, doi: 10.1109/34.297949.

Liu, T., Xu,P., Zhang, S., (2018). A review of recent advances in scanned topographic map processing, *Neurocomputing, 328*, 75-87, https://doi.org/10.1016/j.neucom.2018.02.102

Lucchese, L. & Mitra, S. (2001). Color Image Segmentation: A State-of-the-Art Survey. *Proceedings of Indian National Science Academy, 2* (2). https://www.ece.lsu.edu/gunturk/Topics/Segmentation-2.pdf Retrieved 2022-02-12.

Luft, J., & Schiewe, J.. (2021). Automatic content-based georeferencing of historical topographic maps. *Transactions in GIS*, 25 (6), 1–19. doi:10.1111/tgis.12794

Luft, J. (2020). Automatic Georeferencing of Historical Maps by Geocoding. *International Workshop on Automatic Vectorisation of Historical Maps*. doi:10.21862/avhm2020.10

Manzano-Agugliaro, F., San-Antonio-Gómez, C., López, S., Montoya, F. G., & Gil, C. (2013). Pareto-Based Evolutionary Algorithms for the Calculation of Transformation Parameters and Accuracy Assessment of Historical Maps*. Computers & Geosciences*, 57, 124–132. doi:10.1016/j.cageo.2013.04.010.

Molnár, G. (2010). Making a georeferenced mosaic of historical map series using constrained polynomial fit. *Acta Geodaetica et Geophysica Hungarica, 45*(1), 24-30. doi:10.1556/AGeod.45.2010.1.5

Örebro Kommun & Smart Built Environment. (2019). *Digital handbok för digitala detaljplaner.* https://www.smartbuilt.se/projekt/informationsinfrastruktur/digsam/digital-handbok-for-digitaliserade-och-standardiserade-detaljplaner/ Retrieved 2022-02-18.

Rijsbergen, C. J. V. (1979). *Information Retrieval*. Butterworth-Heinemann. http://www.dcs.gla.ac.uk/Keith/Preface.html Retrieved 2022-03-09.

Saalfeld, A. (1988). Conflation Automated map compilation. Int. J. Geogr. Inf. Sci., 2, 217-228. https://doi.org/10.1080/02693798808927897

Ståhl, N., Weimann, L. (2022). Identifying wetland areas in historical maps using deep convolutional neural networks, *Ecological Informatics, 68*, 101557, ISSN 1574-9541, https://doi.org/10.1016/j.ecoinf.2022.101557.

Sweco (2020). *Automatisk digitalisering av detaljplaner, Rapport 1.0*. Lantmäteriet.

Taha, A.A., Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging, 15,* (29) https://doi.org/10.1186/s12880-015-0068-x

OpenCV. (n.d.). *Template matching*. https://docs.opencv.org/3.4/d4/dc6/tutorial_py_template_matching.html Retrieved 2022-02-22

Thulin, C. (2022, April 11). *Frågor och svar, Nya Regler för detaljplaner*. SKR. https://skr.se/skr/samhallsplaneringinfrastruktur/planeringbyggandebostad/digitaliseringsamhallsbyggnad/digitaladetaljplaner/fragorochsvarnyareglerfordetaljplaner.62971.html Retrieved 2022-02-22

Thydén, B. (2021). *Undersökningsrapport - enkätundersökning om kommuners arbete med att digitalisera detaljplaner och översiktsplaner*. Lantmäteriet and Boverket.

Uhl, J. H., & Duan, W. (2021). Automating Information Extraction from Large Historical Topographic Map Archives: New Opportunities and Challenges. In M. Werner & Y.-Y. Chiang (Eds.), *Handbook of Big Geospatial Data* (pp. 509-522). Springer International Publishing. https://doi.org/10.1007/978-3-030-55462-0_20

Wang, Z., Wang, E. & Zhu, Y. (2020). Image segmentation evaluation: a survey of methods. *Artif Intell Rev, 53*, 5637–5674. https://doi.org/10.1007/s10462-020-09830-9

Wolter, D., Blank, D., & Henrich, A. (2017). Georeferencing River Networks Using Spatial Reasoning. *Proceedings of the 11th Workshop on Geographic Information Retrieval (GIR)*. https://doi.org/10.1145/3155902.3155907

Xavier, E. M. A., Ariza-López, F. J. & Ureña-Cámara, M. A. (2016). A Survey of Measures and Methods for Matching Geospatial Vector Datasets. *ACM computing surveys, 49,* (2), 1–34. doi:10.1145/2963147

Y.J. Zhang, (1996). A survey on evaluation methods for image segmentation, *Pattern Recognition, 29*, (8),1335-1346, https://doi.org/10.1016/0031-3203(95)00169-7.

Yan, W. Y., Easa, S. M. & Shaker, A. (2017). Polygon-based image registration: A new approach for geo-referencing historical maps. *Remote Sensing Letters, 8*, 703-712. 10.1080/2150704X.2017.1317928.

Zatelli, P., Gobbi, S., Tattoni, C., La Porta, N. & Ciolli, M. (2019). Object-based image analysis for historic maps classification. *The international archives of the photogrammetry, remote sensing and spatial information sciences, XLII-4/W14*, 247–254. doi:10.5194/isprs-archives-xlii-4-w14-247-2019

Zhang, W. & Ge, Y. & Leung, Y. & Zhou, Y. (2020). A georeferenced graph model for geospatial data matching by optimising measures of similarity across multiple scales. *International Journal of Geographical Information Science, 35*. 10.1080/13658816.2020.1858301.

Zhang, X., Ai, T., Stoter, J., Zhao, X. (2014). Data matching of building polygons at multiple map scales improved by contextual information and relaxation, *ISPRS Journal of Photogrammetry and Remote Sensing, 92*, 147-163. https://doi.org/10.1016/j.isprsjprs.2014.03.010.

Zitová, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing, 21*(11), 977–1000. https://doi.org/10.1016/s0262-8856(03)00137-9

TRITA – ABE-MBT- 22594