University of Zurich[UZH]

# Effect of Urban Area Definition on Urban Sprawl Measurement

GEO 511 Master's Thesis


**Author**
Patrick Schenker
18-733-493

**Supervised by**
Prof. Dr. Ross Purves
Yves Maurer (Yves.Maurer@are.admin.ch)

**Faculty representative**
Prof. Dr. Ross Purves


30.06.2023
Department of Geography, University of Zurich

*“We shape our buildings and afterwards our buildings shape us.”*

(Churchill, 1943, as cited by UK Parliament, 2023)

# Acknowledgments

I would like to express my sincere gratitude to Prof. Dr. Ross Purves, whose insightful inputs helped me to develop a solid foundation for my research. His willingness to share his expertise and constructive criticism were instrumental in shaping my ideas and refining my research methodology.

I am also deeply grateful to Yves Maurer for his invaluable support and expertise in the field. His indispensable guidance helped me to overcome many obstacles during the project.

Furthermore, I would like to thank all my faculty members and friends who commented on my research and helped me to see things from other angles.

Last but not least, I would like to thank my family, who supported me during my whole journey of writing this master's thesis.

# Abstract

Urban sprawl is a growing phenomenon that has gained attention due to its negative impacts on the environment, society, and economy. Due to the significance of urban sprawl, there has been a surge in research in this area in recent years, with the goal of quantifying and addressing this shift in human settlements. However, one aspect has been largely overlooked: the definition of the urban area. While urban sprawl measurements underwent thorough discussions and improvements, the delineation of urban areas is often treated as a peripheral and unimportant aspect of research in this field. This thesis analyzes this understudied phenomenon and aims to uncover: Are there differences in urban sprawl values, depending on the used definition of urban area? To answer this question, three different urban area definitions were elaborated and delineated. The study focused on five municipalities in Switzerland, namely Wetzikon, Winterthur, Ins, Celerina, and Grindelwald. The sprawl values of the years 2015 and the year 2021 as well as the difference in sprawl value between these two years were utilized to evaluate if differences between the three delineation methods exist. The urban areas were delineated using the $\alpha$-shape algorithm, an approach that has not been previously utilized in urban sprawl research to delineate urban areas. The urban sprawl measurements used to assess the impact of different urban area definitions on urban sprawl were the calculation methods of density, Moran's I, and *WUP* (Weighted Urban Permeation). The results of this thesis suggest that there are indeed differences in urban sprawl values depending on the employed definitions of urban area. This finding could have significant implications for urban sprawl research and highlights the importance of carefully delineating urban areas. Such a finding contradicts the current state of urban sprawl research, which is characterized by the presence of multiple independent and disparate delineations of urban areas. Utilizing well-defined and consistent urban areas in urban sprawl research could lead to more reliable and accurate assessments of urban sprawl, ultimately resulting in better-informed decision-making processes – thus helping to mitigate the negative impacts of urban sprawl. Furthermore, with an ascertainment of a homogenous urban area delineation, existing and future urban sprawl measurements can be compared more effectively, highlighting their respective advantages and disadvantages. This could lead to an upward spiral of improved urban sprawl assessment in urban sprawl research.

# Table of Contents

# List of Figures

# List of Tables

# Acronyms

| | |
|---|---|
| CORINE | Coordination of Information on the Environment |
| DIS | Degree of urban dispersion |
| HP | Horizon of Perception |
| NRI | National Resources Inventory |
| OSM | OpenStreetMap |
| SIUAFs | Small Isolated Urban Area Fragments: areas that meet the requirements to be classified as "Non-rural area" except for the 2000 $m^2$ size threshold |
| UD | Utilization Density (a component of WUP) |
| UP | Urban Permeation (a component of WUP) |
| UPU | Urban Permeation Units |
| WUP | Weighted Urban Proliferation: a calculation method to calculate urban sprawl |

# Chapter 1| Introduction

## 1.1 Foreword

As cities expand and populations grow, urban sprawl – the extension of the city into the countryside (Mayhew, 2015) – becomes an increasingly common sight in our rapidly changing world (Behnisch et al., 2022). This raises the question: How do we measure this sprawling phenomenon and how is it influenced by the way urban areas are defined?
Urban sprawl is on the rise, particularly in Europe (Behnisch et al., 2022). Switzerland is especially affected by high urban sprawl values (Behnisch et al., 2022).
Population increase and unplanned development of urban structures leads to issues like lack of availability of resources, infrastructures, services, and facilities (Rastogi & Jain, 2018). This can lead to the effect that the population is moving away from dense city centers towards the outer parts of settlements in a haphazardous manner, resulting in urban sprawl (European Environment Agency & Swiss Federal Office for the Environment, 2016; Rastogi & Jain, 2018). Urban sprawl is a phenomenon with a number of negative effects (Rogers et al., 2013). Not only does it add pressure on the environment, for example, by overbuilding areas worthy of protection, like marshes (Schwick et al., 2011), but is also results in decreased efficiencies in terms of energy, cost, and time due to the longer distances between residences and workspaces (Rogers et al., 2013). Urban sprawl has been also identified as a potential source for the loss of fertile lands, biodiversity, and open spaces; higher greenhouse gas emissions and pollution; lower water quality; an increased runoff and flood potential; excessive infrastructure and public service costs; the decline in public space; reduced social cohesion; the loss of a sense of community; traffic congestion; income inequality; and further issues (Mosammam et al., 2017; Swiss Federal Council, 2022). In order to prevent its detrimental effects, efforts have been undertaken to quantitatively identify urban sprawl. Accurately measuring urban sprawl is crucial to address its associated issues. Advancements in this field allow to not only tackle the issues of urban sprawl effectively but also help to ensure fairness regarding the policies that aim to do so.

## 1.2 Approach, Scope, and Objectives

In literature, a large number of urban sprawl research is based on a preceding urban area delineation. However, the input data used by urban sprawl researchers are totally different. Not only are there differences in the entities included in urban areas (for example if vacant parcels, city parks or greenhouses are to be seen as urban areas), but also how the selected entities should form an urban area.

For this thesis, three distinct approaches were utilized to delineate urban areas based on three different definitions. These definitions subsume the current views on urban areas for the purposes of urban sprawl measurement. The delineation of urban areas for this thesis was based on vector data of building footprints and other areas of the TLM3D dataset. It was based on the $\alpha$-shape algorithm, whose output had to be translated into polygon geometry. This presents a novel way to delineate urban areas for urban sprawl research. The resulting maps of urban areas were made single-handedly, not only because there are no existing publicly available maps of this sort for Switzerland that are sensible for a large-scale analysis, but also to ensure comparability between the sprawl measurements.

The three obtained urban area delineations were quantitatively analyzed using a small set of sprawl measurements. The used measurement methods were population density, Moran's I, and *WUP* (Weighted Urban Proliferation), including its components. Only calculation methods that are strongly linked to space were utilized, as opposed to methods that rely on factors such as the time required for commuters to reach their workplaces.

The research question to be answered in this thesis is: Are there differences in urban sprawl values, depending on the used definition of urban areas? This aims to show that the current heterogeneity of urban area definitions in urban sprawl research, and thus the variable delineations of it, have an impact on the outcomes of existing urban sprawl measurements. Until now, this question has remained mostly unaddressed, as is apparent from the very limited research quantifying and highlighting the existing disparities in urban area delineation for urban sprawl research. Large differences in urban sprawl values stemming solely from the urban area definition used could have important and significant implications. This highlights the needs for a consistent and accurate method of defining urban areas, especially as the failure to do so could result in misallocation of resources, as decisions about urban planning are often based on these measurements. Otherwise, an appropriate and fair legislation regarding urban sprawl is difficult – if not impossible. Moreover, the lack of consistent urban area delineations could limit the comparability of existing and future studies. Apart from the goal of answering the question if diverging urban areas definitions lead to a measurable effect on urban sprawl values, it also shows the consistency or lack thereof that can be found when comparing the different urban sprawl calculation methods with each other.

This master's thesis could help to verify or challenge the understanding of the term "urban area" within the urban sprawl research community and raise awareness for the possible inconsistencies in the approaches currently used. With the insights gained from this study, urban sprawl can possibly be assessed in a more appropriate way.

In sum, the objective of this study is to examine the relationship between urban area delineations and urban sprawl measurements. The delineation was based on polygons constructed from the geometry resulting from the $\alpha$-shape algorithm. The results of three different ways to delineate urban area were analyzed in this thesis. Measuring urban sprawl through density, Moran's I, and WUP, this study aims to contribute to the understanding of the complex interplay between urban form and sprawl.

## 1.3 Structure of this Thesis

*Chapter 2* presents the state-of-the-art approaches to delineate urban areas, including theoretical considerations and the practical implementations, all in context of providing a basis for urban sprawl calculation. This chapter also presents the calculations employed to measure urban sprawl, establishing a foundation for comprehending the findings of this thesis. In addition, the limited previous research is discussed, and existing approaches and thresholds to delimit urban areas are presented. Furthermore, the research question and research gaps are addressed.

*Chapter 3* details the methods used to delineate the urban areas constructed for this thesis, which includes a description of the data used, the algorithm employed for the delineation, and the process used to create the polygons. Additionally, the steps necessary between the creation of the polygons and the calculation of the urban sprawl are explained.

*Chapter 4* visualizes and explains the outputs of the urban area delineations and the final result: the urban sprawl measurements.

*Chapter 5* further discusses the findings from the previous sections and provides possible explanations for the results. Additionally, the limitations of the chosen approaches are presented in this chapter. Furthermore, the urban areas delineated for this thesis are compared with already existing ones. It also states the implications of this research for future studies and for urban sprawl research as a whole.

*Chapter 6* summarizes the main findings and highlights their importance.

# Chapter 2 | Theoretical Background

## 2.1 Preface

Urban sprawl, although an extensively analyzed phenomenon, lacks universally accepted characteristics and a generally agreed upon definition (Manesha et al., 2021; Siedentop, 2005; Torrens & Alberti, 2000; Tsai, 2005). Although it can be described as the extension of the city into the countryside (Mayhew, 2015), different researchers measure sprawl in a very distinct manner (Tsai, 2005). The motivations of urban sprawl detection and the negative effects attributed to this phenomenon are also very distinct (Rubiera-Morollón & Garrido-Yserte, 2020; Siedentop, 2005).

The discussed issues of urban sprawl also are, at least partially, dependent on when the discussion took place: Whereas urban sprawl in the past was rather seen as a problem of urban mobility, an economic issue or a social issue, more recently the thought of urban sprawl as an environmental problem has gained attention (Rubiera-Morollón & Garrido-Yserte, 2020). To this day, however, no issue of urban sprawl has gained the upper hand in literature as the clear main problem of sprawl (Siedentop, 2005). The exact issues associated with sprawl are also subject to a great variety of opinions (Siedentop, 2005). For example, whereas some researchers which focus on the environmental detriments discuss sprawl's effects on the soil stemming from soil sealing and the loss of prime farmland, others point to the displacement of native flora and fauna – especially those species that cannot cope with their changing habitat (Siedentop, 2005).

This lack of agreement also affects the indicators used to detect sprawl: To measure urban sprawl, a lot of different indicators have been used over time. However, no consensus has been achieved until now, regarding which measurable properties of an area characterize sprawl best (Tsai, 2005). Furthermore, even if researchers agree on the characteristic used to detect sprawl, for example to use compactness to assess urban sprawl, the way to measure this characteristic can still be distinct (Rubiera Morollón et al., 2015; Steurer & Bayr, 2020).

In this chapter, the similarities and differences in urban sprawl research will be illustrated – specifically, the indicators used, the way to measure the ones applied in this thesis, and the delineation of the urban area.

## 2.2 Sprawl Characteristics

The selection of indicators to evaluate urban sprawl remains a matter of debate among researchers and varies based on the individual perspectives of urban sprawl scientists (Nazarnia et al., 2019; Steurer & Bayr, 2020; Torrens, 2008; Tsai, 2005). Still, researchers have attempted to grasp and quantify the concept of urban sprawl by identifying specific characteristics that are indicative of urban sprawl. An impactful description of development patterns associated with urban sprawl was proposed by Altshuler et al. (1993, p. 67): "Continuous low – density residential development on the metropolitan fringe, 'ribbon' low density development along major suburban highways, and development that 'leapfrogs' past undeveloped land to leave a patchwork of developed and undeveloped tracts" (Galster et al., 2001; Siedentop, 2005).

The phenomena of leapfrogging and ribbon development are still a frequently discussed characteristic of sprawl and are depicted in Figure 1 (Rubiera-Morollón & Garrido-Yserte, 2020; Siedentop, 2005). Further forms of development frequently used to characterize sprawl are areas with large-lot single-family residential housing, or containing commercial buildings either being built in the form of a commercial strip, or in a widespread form (Galster et al., 2001). By and large, leapfrog and scattered development, commercial strip development, and large expanses of low-density development could be considered main characteristics of urban sprawl (Musakwa & Van Niekerk, 2014).

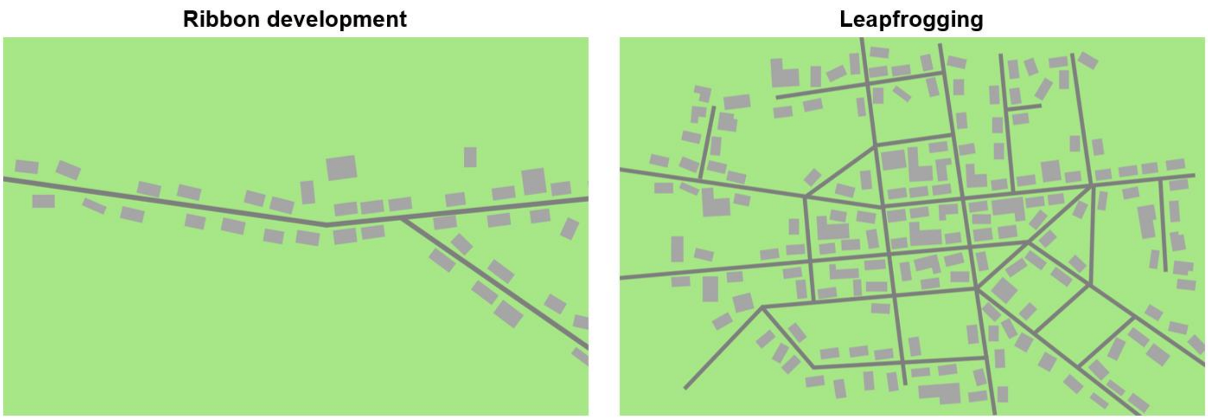



Figure 1: Example of a ribbon development (left) and leapfrogging (right). Pay attention to the linear arrangement of buildings along roads in the ribbon development illustration. In the leapfrogging image, take note of the alternating pattern of developed and undeveloped areas that would be experienced while walking from the urban center outwards.

To measure urban sprawl, it is also necessary to find a way to quantify it. To do that, over the years, a vast number of sprawl indices were employed (Torrens, 2008). These metrics have to

be based on a measurable real-world phenomenon, for which a quantitative assessment is possible.
Amongst them, density measurements can be seen as the most important (Lopez & Hynes, 2003; Torrens, 2008). This metric fits directly to the concept of urban sprawl as a low-density phenomenon (Torrens, 2008).
There are several other metrics used to evaluate sprawl, among which measures of clustering are the most extensively studied and significant, aside from density measurements (Zhou et al., 2019). The measure of clustering aims to quantify the decentralization and compactness of the city, which includes the detection of non-compact developments, like leapfrog or ribbon developments (Torrens, 2008; Tsai, 2005).
However, measuring density and clustering, like other characteristics of urban sprawl, is not straightforward and the various methods employed have faced criticism (Steurer & Bayr, 2020). These measurement methods will be discussed in the coming chapter.

The various characterizations of urban sprawl discussed earlier in this thesis imply that it is a definitive condition that can affect a region. While this perspective on sprawl exists in the urban sprawl literature, it is also widely recognized as a process of development over time (Galster et al., 2001). This creates further division within the community of urban sprawl researchers, with some viewing it as a quantifiable phenomenon that can be measured on absolute scales (Steurer & Bayr, 2020) while others consider it a dynamic process that occurs between two or more points in time (European Environment Agency & Swiss Federal Office for the Environment, 2016).

## 2.3 Sprawl Measurements

Similar to the characteristics of sprawl, the methods used to quantify it are also diverse. (Steurer & Bayr, 2020; Tsai, 2005). This is not only due to the fact, that the different characteristics of urban sprawl require distinct calculation methods, but also due to the existence of different calculation methods aiming to quantify the same characteristic of urban sprawl (Torrens, 2008).
The discussion of methods to measure sprawl will be limited to:

- Density measurements, which capture the most important characteristic of sprawl (Lopez & Hynes, 2003; Torrens, 2008).
- Moran's I, an indicator of clustering, which over the years has become a widely used method to determine urban sprawl (Zhou et al., 2019).
- Weighted Urban Permeation (WUP) including its components. The WUP measurement underwent validation through sets of sustainability criteria (Behnisch et al., 2022).

- Entropy measures, a common approach to quantify sprawl (Nazarnia et al., 2019; Steurer & Bayr, 2020).
- Fractal analysis, another common method for analyzing sprawl (Jiang & Liu, 2012).

The density measurements, along with the calculation of clustering using Moran's I and WUP and its components, were employed to assess the sprawl in the urban areas. The methodology and results of this analysis will be discussed in the Chapters 3 and 4.

It is important to mention that sprawl characteristics, sprawl measurements and the urban area delineations are interconnected. Whereas sprawl measurements quantify the urban sprawl characteristics, urban area delineations are a necessary input for many urban sprawl measurements. Which urban area delineations are common in urban sprawl research and their characteristics are discussed in Section 2.4 of this thesis. The relationship between sprawl characteristics, sprawl measurements, urban area delineations and the calculated urban sprawl value is visualized in Figure 2. The presented urban sprawl measurements are all dependent on an urban area delineation. In this thesis, three different urban area inputs are used to evaluate the urban sprawl values for their differences. To understand the presented urban sprawl measurements, the precise delineation of urban areas is irrelevant. Thus, until the presentation of the different common definitions of urban areas in Section 2.4, the urban areas can be imagined as all the areas characterized by the presence of buildings, similar to the delineation visible in Figure 2.

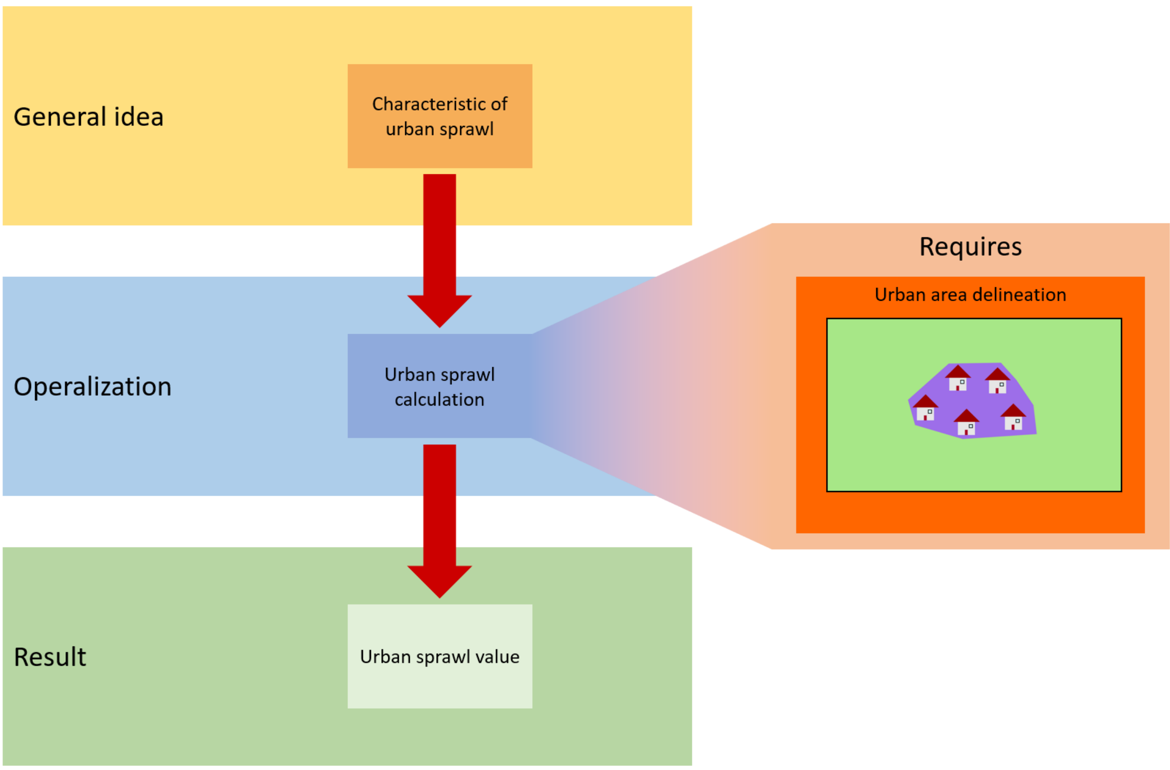



Figure 2: Relationship between sprawl characteristics, sprawl measurements, urban area delineations and the calculated urban sprawl value.

### 2.3.1 Density

Although in literature reviews the density measure is denoted as the most important sprawl measurement (Lopez & Hynes, 2003; Torrens, 2008), there is no agreed-upon method for its precise calculation (Steurer & Bayr, 2020). The general consensus is that the term "density" in the context of urban sprawl refers to the calculation of the number of people divided by a specific area (Galster et al., 2001). However, there is controversy, which group of people should be used in the calculation, as well as which area should be employed (Galster et al., 2001; Schwick et al., 2018; Steurer & Bayr, 2020; Torrens, 2008). The general formula of density in the context of urban sprawl is as follows:

$$d = \frac{N}{A}$$

where $d$ stands for the density, $N$ for the number of people, and $A$ for the chosen area. The group of people used to calculate the density can be residents only (Galster et al., 2001; Steurer & Bayr, 2020) or can additionally include employees (Schwick et al., 2018; Torrens, 2008). In the general discussion of urban sprawl, density is understood as residential density, partly because the number of employed people varies with the business cycle and because non-residential areas tend to cluster, distorting the measure of sprawl (Galster et al., 2001). However, there are still numerous and notable examples of urban sprawl researchers not only relying on population, but also employment as input for their density calculations or other calculations related to density (Glaeser, Kahn, Arnott, et al., 2001; Glaeser, Kahn, & Chu, 2001; Lopez & Hynes, 2003; Schwick et al., 2018; Tsai, 2005).

The discussion regarding the area used for the density calculation is more complex and is directly linked to the discussion of what defines an urban area, which is the central part of this thesis and will be presented in detail in Section 2.4.

The calculation of density, despite its importance in measuring sprawl, should not be used as the only indicator when calculating sprawl (Tsai, 2005). This, as density does not take into account the urban form, which is important too, when trying to measure urban sprawl (Tsai, 2005). Specifically, phenomena like ribbon development can remain undetected using this sole measure. For this reason, the use of multi-dimensional indices (i.e., indices that are composed of different measured characteristics of urban sprawl) is a popular way to measure sprawl (Zhou et al., 2019). In combination with other quantified characteristics of urban sprawl, important aspects of urban sprawl can be detected that could remain undetected utilizing unidimensional measures, such as using density alone (Zhou et al., 2019). However, such multi-dimensional indices do not only raise the unanswered question, which characteristics to quantify (Torrens, 2008), but can also spark debate over the weights put on the different components of the multi-dimensional indicator (Rubiera Morollón et al., 2015).

### 2.3.2 Moran's I

Moran's I is a widely used method to calculate clustering in order to calculate urban sprawl (Zhou et al., 2019), with the aim of assessing the compactness of an area and quantifying phenomena like leapfrogging and ribbon developments (Torrens, 2008; Tsai, 2005). This measure can be used to reveal the spatial autocorrelation between a set of areas (Anselin, 1995). It can be differentiated between the local Moran's I and the global Moran's I. Whereas the local Moran's I calculates the autocorrelation locally, i.e., of sub-areas, the global Moran's I yields a sprawl value on an entire dataset (Musakwa & Van Niekerk, 2014). Although the local Moran's I is useful to determine sprawl hot spots and cold spots (Musakwa & Van Niekerk, 2014), this thesis focuses on the calculation and implementation of the global version of Moran's I, which provides information on the overall pattern of spatial autocorrelation of the data (Musakwa & Van Niekerk, 2014). From here on, unless specified differently, the term Moran's I will be used to denote the global version, acknowledging that certain concepts may also hold true for the local Moran's I. This (global) Moran's I, is calculated as:

$$Moran's\ I = \frac{N \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}(X_i - X)(X_j - X)}{\left(\sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}\right)(X_i - X)^2}$$

where $N$ is the number of sub-areas; $X_i$ is the variable of interest in sub-area $i$; $X_j$ is the variable of interest in sub-area $j$; $X$ is the mean; and $W_{ij}$ is the weighting between sub-areas $i$ and $j$ (Zhou et al., 2019). The "variable of interest", in the context of urban sprawl, can constitute the pervasiveness of different phenomena, such as population, employment, or developed land (Zhou et al., 2019). The result of the shown formula yields a result lying between -1 and 1, where -1 indicates strong negative spatial autocorrelation (i.e., resembling a checkerboard pattern), 0 indicates random spatial ordering and 1 indicates a strong positive spatial autocorrelation (i.e., clustering of similar values) (Musakwa & Van Niekerk, 2014). A depiction of this can be seen in Figure 3.

**Moran's I**

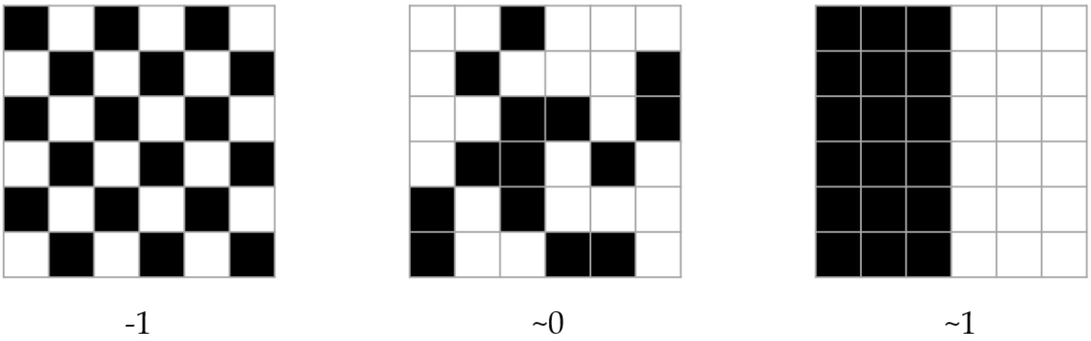



Figure 3: Example of the Moran's I values for different configurations using a black and white lattice depicting distinct values of the variable of interest. (Image source: own image, based on Sandoval Félix and Castañón-Puga (2019) and Böck et al. (2017))

The calculation of Moran's I requires several decisions, such as determining the study area, selecting the sub-areas, and choosing an appropriate weighting scheme (Zhou et al., 2019). The challenge of defining the study area is not exclusive to the calculation of Moran's I. The definition of the study area also has to be done when calculating sprawl using other methods (García-Álvarez & Camacho Olmedo, 2017; Zhou et al., 2019).

Urban sprawl can be calculated at various scales, including the scale of neighborhoods, municipalities, metropolitan areas, federal states, megaregions, or countries (Schwick et al., 2018; Zhou et al., 2019). The choice of the study area also affects the data that can be used for the calculation (Schwick et al., 2018). The sub-areas can vary in nature, they can for example be based on political subdivisions, cadastral subdivisions, or a square grid (Musakwa & Van Niekerk, 2014; Salvati & Carlucci, 2014; Steurer & Bayr, 2020).

If a square grid is selected, a decision must be made regarding the size of the cells (Zhou et al., 2019). Literature shows a wide range of square edge lengths used for calculating urban sprawl, ranging from the 10 m scale to the km scale (Rubiera Morollón et al., 2015; Schwarzak et al., 2014; Steurer & Bayr, 2020; Zhou et al., 2019). The choice of too large edges can result in the failure to detect some localized phenomena (Terzi & Kaya, 2011), such as ribbon developments.

Furthermore, the calculation of Moran's I requires the choice of weighting between the sub-areas (Zhou et al., 2019). There are various methods available for this (Musakwa & Van Niekerk, 2014; Tsai, 2005). Using inverse distance leads to a more sensitive and accurate characterization of the urban form, compared with assigning values of 0 to non-neighboring sub-areas and 1 to bordering ones (Tsai, 2005).

As previously explained, the variable of interest for Moran's I can encompass various phenomena. Some examples of specific values used in literature are those of population density, employment density, or the percentage of developed land in each sub-area (Tsai, 2005;

Zhou et al., 2019). This thesis will largely disregard the calculation of Moran's I using population or employment density. This is due to the reason that using population or employment density to calculate the Moran's I circumvents the necessity for a discussion of urban area, as it does not rely on any border between urban and non-urban area (Tsai, 2005). Thus, the approach using population or employment density does not provide any insight on the effects of urban area definition on urban sprawl measurements. Furthermore, the reliance on the totality of population/employment without considering urban borders is not exempt from criticism. Using population or employment density as input for the Moran's I calculation leads to the reliance of them to assess urban sprawl, without even considering whether all population/employment is urban or if urbanity is solely defined by the presence of population or employment (Fulton et al., 2001).

Moran's I can be used in combination with other measures to assess urban sprawl from different angles (Zhou et al., 2019). As it already powerfully characterizes different components of compactness and sprawl within one index, it was also proposed as an independent sprawl index (Tsai, 2005). It has been demonstrated that the Moran's I index can detect changes in urban form, such as leapfrog and ribbon development. Additionally, in comparison to multi-dimensional indices, it does not bear the risk of being an arbitrarily weighted amalgam of different sprawl indices (Tsai, 2005). However, the view that Moran's I can be used to assess urban sprawl on its own is not universally accepted due to certain limitations associated with its use (Steurer & Bayr, 2020). For example, issues can arise if the variable of interest consists of values on a continuous scale (e.g., percentage of developed area). Then, it can be the case, that the change in values occurs gradually (e.g., there is a gradual transition from very low to very high development). Under these circumstances, this smooth transition in values leads to a high detected spatial autocorrelation, as the value in each sub-area is similar to its neighboring ones. To achieve a low spatial autocorrelation, the distribution of values should be characterized by an interplay of neighboring areas with very high and very low values (Steurer & Bayr, 2020). Furthermore, the result of Moran's I is a one-way street: Various effects affect the Moran's I, but from the result alone, it is not possible to derive the urban form (Tsai, 2005). One of the reasons for a rather high Moran's I can also be a skew in the distribution of values (Steurer & Bayr, 2020). For example, in an area where large tracks of land are undeveloped, the Moran's I will skew towards a higher value, as it detects the high autocorrelation between these undeveloped areas (Steurer & Bayr, 2020).

### 2.3.3 WUP

WUP (Weighted Urban Proliferation) is a metric constructed out of several components (Jaeger & Schwick, 2014). These components are the degree of urban dispersion (DIS), the urban permeation (UP), and the utilization density (UD) (Jaeger & Schwick, 2014). The relationship between these components can be seen in Figure 4.

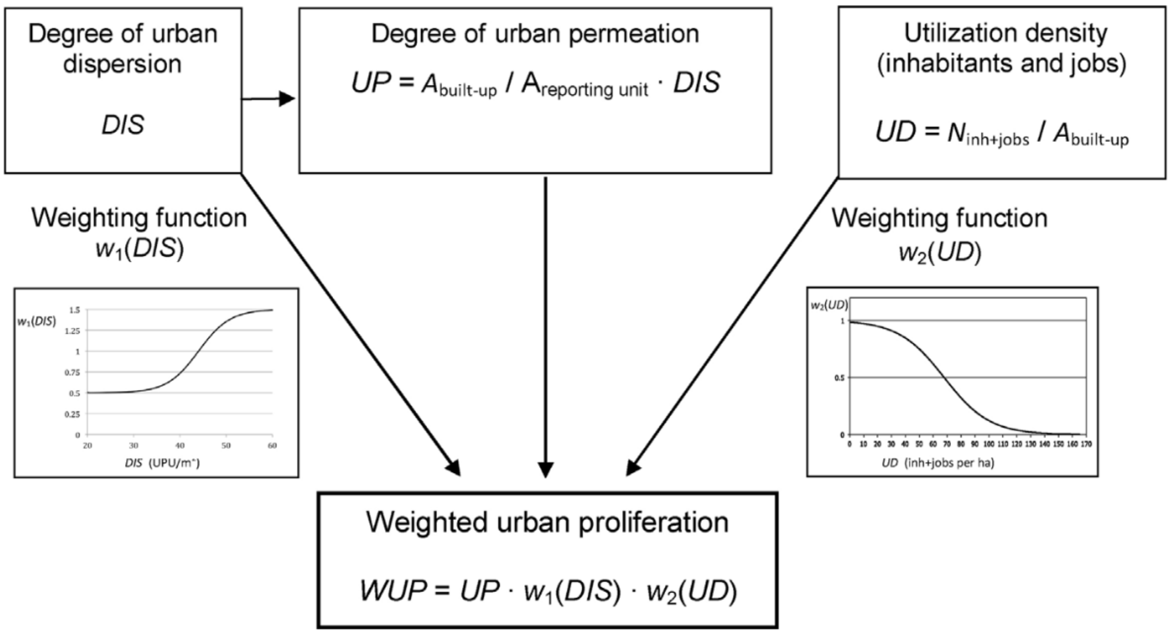



Figure 4: Relationship between the different components of WUP (Jaeger & Schwick, 2014).

To calculate the degree of dispersion (DIS), it is necessary to have an urban area, partitioned into two classes (Jaeger, Bertiller, Schwick, Cavens, et al., 2010). A proposed way to do this, is to use a square lattice with the cells having an edge length of 15 m (Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Schwarzak et al., 2014). Such an edge length was preferred over larger ones, as small cell sizes are less of a generalization of the defined urban area, and lead to more accurate results (Jaeger, Bertiller, Schwick, Cavens, et al., 2010). Like Moran's I, DIS is employed to capture the spatial relationship of the cells of the urban area. In the calculation of WUP there is also the use of a so-called "horizon of perception" (Jaeger, Bertiller, Schwick, Cavens, et al., 2010). It is utilized to ensure that this method can capture the capabilities of human perception, as the definition used for urban sprawl during the creation of WUP was: "Urban sprawl is visually perceptible. A landscape suffers from urban sprawl if it is permeated by urban development or solitary buildings. The more urban area present in a landscape and the more dispersed the urban patches, the higher the degree of urban sprawl" (Jaeger, Bertiller, Schwick, Cavens, et al., 2010, p. 428). The suggested value for the horizon of perception used by the authors is 2 km (Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Schwarzak et al., 2014).

Based on this raster and the set horizon of perception, DIS can be calculated, which is the average weighted distance between any two sub-areas chosen randomly within the urban area in the study area (Jaeger, Bertiller, Schwick, Cavens, et al., 2010). The formula of DIS is:

$$DIS = \frac{1}{A_{built-up}} \int_{\vec{x} \in urban\ areas} \frac{1}{\int_{\vec{y} \in urban\ areas\ and\ |\vec{x}-\vec{y}|<HP} d\vec{y}} \int_{\vec{y} \in urban\ areas\ and\ |\vec{x}-\vec{y}|<HP} \sqrt{\frac{2 \cdot |\vec{x} - \vec{y}|}{1\ \mathrm{m}} + 1} - 1 \frac{\mathrm{UPU}}{\mathrm{m}^2}$$

where $A_{built-up}$ denotes the total amount of urban area within the area of interest, $\vec{x}$ and $\vec{y}$ are sub-areas within the area of interest, $HP$ is the horizon of perception, and m and UPU are units (meters and urban permeation units) (Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Jaeger & Schwick, 2014; Schwick et al., 2018). Using the calculated DIS value, it is then possible to calculate UP (degree of urban permeation), as shown in Figure 4. In the formula of UP, $A_{built-up}$ and $A_{reporting\ unit}$ refer to the size of the urban area inside the area of interest and the size of the area of interest itself, respectively (Jaeger, Bertiller, Schwick, Cavens, et al., 2010). With DIS, UP and UD, calculated using the size of the built-up area inside the area of interest and its inhabitants and employees, the WUP value can be calculated (Schwick et al., 2018). The formula for WUP can be seen in Figure 4. As the formula shows, there is a weighting applied to DIS and UD (Schwick et al., 2018). Weighted DIS and UD are called $w_1(DIS)$ and $w_2(UD)$, respectively (Schwick et al., 2018). These weighted values are calculated like this:

$$w_1(DIS) = 0.5 + \frac{e^{0.294432\ \mathrm{m}^2/\mathrm{UPU} \cdot DIS - 12.955}}{1 + e^{0.294432\ \mathrm{m}^2/\mathrm{UPU} \cdot DIS - 12.955}}$$

$$w_2(UD) = \frac{e^{4.159 - 0.000613125\ \mathrm{km}^2/(\mathrm{INH+J}) \cdot AD}}{1 + e^{4.159 - 0.000613125\ \mathrm{km}^2/(\mathrm{INH+J}) \cdot AD}}$$

where m (meters), UPU (urban permeation units) and $\mathrm{INH} + \mathrm{J}$ (inhabitants and employees) are units (Schwick et al., 2018).

The aptness of the WUP and UP measures was assessed by checking if they fulfill a set of suitability criteria (Jaeger, Bertiller, Schwick, Cavens, et al., 2010). These are (Jaeger, Bertiller, Schwick, Cavens, et al., 2010):

1) intuitive interpretation
2) mathematical simplicity
3) modest data requirements
4) low sensitivity to very small patches of urban area
5) monotonous reaction to increases in urban area
6) monotonous reaction to increasing distance between two urban patches when within the scale of analysis
7) monotonous reaction to increased spreading of three urban patches
8) same direction of the metric's responses to the processes in criteria 5, 6, and 7
9) continuous reaction to the merging of two urban patches

10) independence of the metric from the location of the pattern of urban patches within the reporting unit
11) continuous reaction to increasing distance between two urban patches when they move beyond the scale of analyses
12) mathematical homogeneity
13) additivity (i.e., the combination of multiple reporting units produces an equivalent result to the computation of them all together from the outset)

UP meets all thirteen of the mentioned criteria, whereas WUP meets criteria 1-12, and, if the dispersion of the built-up area and the utilization density do not differ between the reporting units that are combined, it also meets criterion 13 (Nazarnia et al., 2019). As WUP already measures multiple dimensions of sprawl, it was also proposed that this measure could be used in isolation, i.e., without the need of additional measures, to assess urban sprawl (Nazarnia et al., 2019). In addition, WUP also satisfies the 34 requirements proposed by Niemeijer and de Groot (2008) to select environmental indicators (Behnisch et al., 2022). Furthermore, various tests have shown that this method captures urban sprawl well (Behnisch et al., 2022). Besides that, this sprawl measurement has already been used by the Swiss Federal Office for the Environment and the European Environment Agency (Behnisch et al., 2022).

### 2.3.4 Entropy Measures

Entropy measures are a widespread way to calculate urban sprawl (Nazarnia et al., 2019; Steurer & Bayr, 2020). Of the entropy measures used to calculate urban sprawl, the Shannon's index is the most used one (Steurer & Bayr, 2020). This entropy measure can be applied to calculate the equality of distribution of the urban area (Tsai, 2005). For this, the area of interest is divided into sub-areas using a set of concentric rings around the city center (see Figure 5) (Nazarnia et al., 2019). Despite its frequent appearance in urban sprawl literature, several issues were detected when using the Shannon's index to calculate sprawl (Steurer & Bayr, 2020). For instance, it cannot be applied to data containing sub-areas with a density of zero (Tsai, 2005). This is an issue, as zero-density areas are not unheard-of (for example parks constitute such areas) (Tsai, 2005). Moreover, interpreting the Shannon's index is challenging, as it consists of two separate indices, resulting in similar entropy values for very different settlement arrangements (Steurer & Bayr, 2020). In addition, the use of the Shannon's Index relies on the choice of the city center, which is necessary for the calculation (Nazarnia et al., 2019). This affects the outcome of the calculation – an effect that should not be present in a good sprawl measurement (Nazarnia et al., 2019). Furthermore, it was found, when comparing the Shannon's index approach to the 13 suitability criteria mentioned before, that it only meets five of the 13 criteria (Nazarnia et al., 2019). Due to the mentioned issues associated with entropy measures, they will not be further explained and discussed in this thesis.

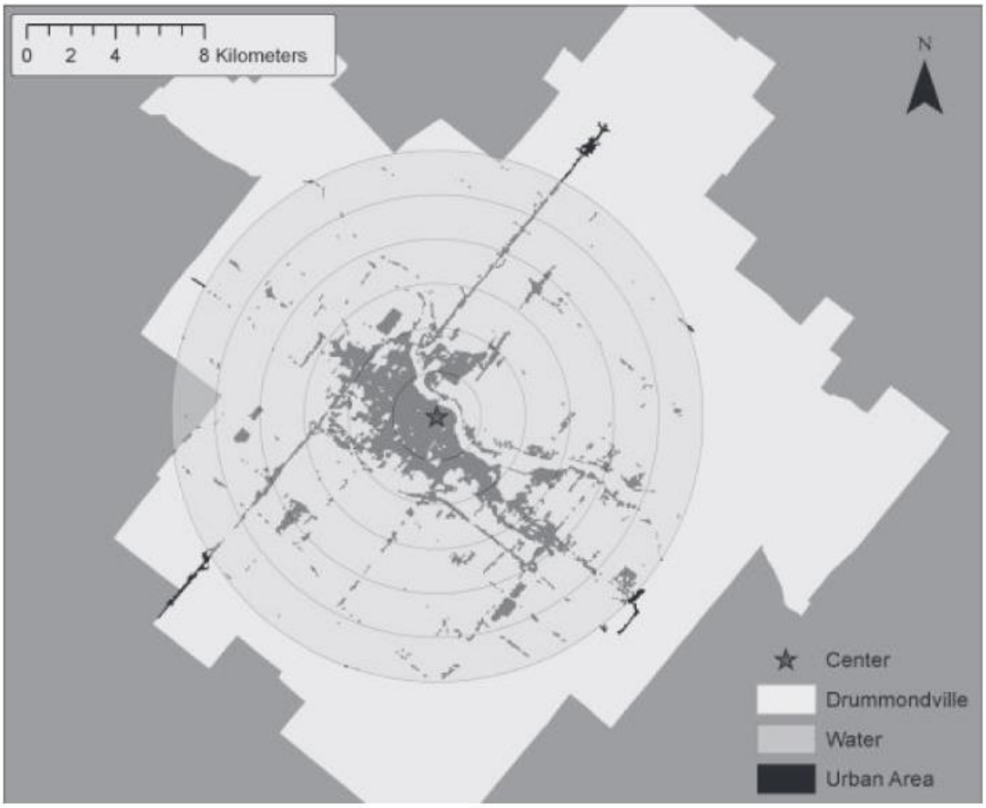



Figure 5: Example of the concentric rings used to calculate Shannon's entropy on a city (in this case Drummondville) (Nazarnia et al., 2019).

### 2.3.5 Fractal Analysis

Another frequently used method to quantify urban sprawl is the box-counting method, which is a specific form of fractal analysis (Jiang & Liu, 2012). It can be used to assess the complexity of the shape of urban areas (Terzi & Kaya, 2011) and is used to assess the compactness of an urban form (Jiang & Liu, 2012). Although there is usefulness in this measure, it has several drawbacks (Jiang & Liu, 2012). Especially when using differently sized study areas, it was observed that the results of the box-counting methods of different areas were incomparable (Jiang & Liu, 2012). Furthermore, the choice of parameters, like the prerequisite to set the box-size, are not obvious and are disputed (Jiang & Liu, 2012). In addition, the relationship between fractal dimension and urban sprawl is not straightforward, as fractal dimension was observed to rise (indicating more sprawl) when already sprawling areas became more compact and to shrink (indicating less sprawl) when the urban area grew in a more dispersed, semi-linear form (Terzi & Kaya, 2011).

Given the limitations of using fractal analysis to measure urban sprawl, it will not be explored further in this thesis.

## 2.4 Urban Area Definitions

### 2.4.1 Preamble

Many of the presented calculations to quantify urban sprawl, but also others, depend on some sort of definition of urban areas or urbanity (Ewing & Hamidi, 2014; Galster et al., 2001; Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Jiang & Liu, 2012; Musakwa & Van Niekerk, 2014; Nazarnia et al., 2019; Petrescu, 2019; Steurer & Bayr, 2020; Terzi & Kaya, 2011). This makes sense, as the presence and distribution of urban land cover directly relates to the definition of urban sprawl (Galster et al., 2001).

However, in the search for literature for this thesis, it became apparent that the literature on what constitutes an urban area in the context of urban sprawl is very scarce, and a discourse on how to best define urban areas to calculate urban sprawl is almost non-existent.

Instead, urban areas are often subject to individual discretion. If the urban areas were delineated by the urban sprawl researchers themselves, the way urban areas were delineated or urbanity as a concept was defined was found to mostly be based on individual opinions.

In this section, the state-of-the art approaches for defining urban areas for the purpose of measuring urban sprawl will be discussed. Three distinct overarching concepts of urban area delineation are introduced. These concepts form the foundation for assessing the differences among the various definitions of urban areas in the context of identifying urban sprawl in this thesis. Given the scarcity of theoretical discussions on the optimal way to define urban areas for urban sprawl research, this section will mainly focus on the practical approaches employed in obtaining urban areas for such measurements. The selection of urban area delineations used by various researchers in the field of urban sprawl is often poorly justified, with limited or no support from literature and insufficient explanations for the reasoning behind their choices. Typically, only information on the chosen datasets, areas, or entities included in the delineations is provided. To better understand the reasoning behind the choice of urban area delineation in previous research, it was necessary to analyze the included and excluded areas in order to form a coherent understanding of the concepts behind the presented pattern. For this reason, which areas were included or excluded in the respective urban area delineations in many cases was the main lead to paint a picture of the concept behind the delineations of urban areas used by the urban sprawl researchers.

A common approach to defining urban areas is to delegate the task to an external party that provides the delineation of urban areas. However, these delineations were in general not made for use in urban sprawl research only. In such cases, it is also possible that the original mapmaker did not define any areas as “urban” or as any other area directly applicable for urban sprawl calculation. In such instances, the urban sprawl researchers then defined the urban area by deciding which different classes of the dataset they used to depict the urban area.

The diverse perspectives from these approaches are consolidated into three groups, namely "Built-up area", "Area with strong anthropogenic influence", and "Non-rural area". The concepts behind these three viewpoints on urban area were used to delineate the urban areas for this thesis, which is presented in Chapter 3 of this thesis. The three main concepts of urban areas presented here can be explained like this:

- The concept that urbanity is equivalent to the presence of built-up area
- The point of view that urbanity refers to areas strongly influenced by anthropogenic activity
- The viewpoint that areas strongly influenced by non-rural anthropogenic activity exhibit a high urbanity

It has to be mentioned that the approaches used in practice to assign urbanity to certain areas were hardly ever created to be in accordance with specifically one of the three viewpoints on urbanity. They often represent a mixture of these archetypical three approaches. The characterizations of urban areas used as foundation for the three main definitions of urban areas in this thesis, which will be presented in the following, are always either such ones made for urban sprawl research or maps for general use which were utilized as input for urban sprawl calculations.

## 2.4.2 Built-Up Area

### 2.4.2.1 Overview and Definition

The class "Built-up area" presented in this thesis encompasses all the areas that could be identified as built-up. For example, areas characterized by residences are part of this class. In this class, vegetation cover is allowed, providing that it is in the context of specific built-up entities (e.g., gardens). Parks or other green spaces are excluded. In practice, no urban area delineation strictly adheres to this definition. In all cases, small areas of urban greenery or other uninhabited areas can be incorporated into urban area delineation, as long as they are situated near to built-up areas. In addition, small areas adhering to the given definition are not always detected in the existing urban area delineations. Despite this, the adherence to built-up area as a primary criterion is used in a good number of urban area delineations. In this section, literature that endorses this method of urban area delineation are referenced, and examples of approaches that followed this definition in practice are presented to shed a light on this viewpoint on urban area delineation.

### 2.4.2.2 Literature Review

The notion that urban area is best characterized as "Built-up area" can be found in the work of Schwick et al. (2018), for instance. For their urban sprawl calculation, the urban area was delineated based on the presence of buildings (Schwick et al., 2018). The exact delineation of the urban area, however, is subject to individual discretion (Schwick et al., 2018). This is the case, as the delineation of the urban area is based on the visual inspection of the scene and a subsequent delineation by hand of the urban area (Schwick et al., 2018). Individual buildings that do not appear to be grouped with other structures are considered part of the urban area and are assigned a surrounding area (Schwick et al., 2018). In a next step, areas larger than 2-4 ha without any buildings were cut out from the urban area (Schwick et al., 2018). In sum: The method used by Schwick et al. (2018) mostly excludes areas that are not characterized by the presence of buildings, at least if they surpass the threshold of 2-4 ha (Schwick et al., 2018). The reason for this choice of urban definition over other approaches is not stated – but is likely in relation to the chosen method to calculate the urban area: the WUP. Namely, the urban area-input necessary for the WUP calculation is sometimes also referred to as "built-up area" (Jaeger & Schwick, 2014).
Instead of using self-made maps to calculate sprawl, Steurer and Bayr (2020) decided to use pre-made delineations of urban area. In this case CORINE land cover data (Environment Agency Austria, 2019) were used. Despite the fact that some maps, like the CORINE land cover, do not contain any layers directly reflecting any areas that constitute urban area itself (Environment Agency Austria, 2019), this does not stop urban sprawl researchers from using CORINE data (Salvati & Carlucci, 2014; Steurer & Bayr, 2020). This is feasible because these maps feature delineations of specific areas that can be combined to create a map of urban regions (Steurer & Bayr, 2020). If and to what extent the used urban area delineation is compatible with the "Built-up area" definition depends on the combination of areas of the CORINE land cover used. In the case of Steurer and Bayr (2020), the classes "continuous urban fabric" and "discontinuous urban fabric" were used (European Environment Agency, 2011; Steurer & Bayr, 2020). These two classes are characterized by their relatively high amount of coverage with impermeable features, like buildings and roads (Environment Agency Austria, 2019). Consequently, the approach by Steurer and Bayr (2020) is also largely based on the concept of "Built-up area" to delineate urban areas for the purpose of urban sprawl detection. It has to be mentioned here, that approaches like the one of the CORINE land cover do not prohibit the inclusion of some vegetation in its "urban fabric" class, as the percentage of soil sealing necessary to be considered as "continuous urban fabric" and "discontinuous urban fabric" is below 100% (Environment Agency Austria, 2019). This is attributed not only to the limitations inherent in oversimplifying vector data, but also to the notion that areas like residential gardens are considered a crucial component of built-up areas, despite not

contributing to soil sealing in their region (Copernicus, 2020; Environment Agency Austria, 2019).
Some urban sprawl researchers opted for data from the "Urban atlas" (Copernicus, 2020). Like the CORINE land cover map, this dataset contains different classes (Copernicus, 2020). This dataset is used for example by Petrescu (2019), where the classes for example include residential buildings, ports, airports, mineral extraction sites, dump sites, isolated structures, commercial and industrial sites, and all mapped roads and railways. However, vegetated areas like forests or parks are not included in this definition of urban areas (Petrescu, 2019) and an adherence to the notion of urban areas as purely built-up areas is apparent. This is even more the case in Prastacos and Lagarias (2016), who also utilized data from the "Urban Atlas" and confined their examination of urban sprawl to areas primarily consisting of residential buildings and isolated structures.
The perspective that urban areas are defined by the presence of artificial structures is commonly seen in literature on urban sprawl that utilizes remote sensing data, such as in the works of Ewing and Hamidi (2014), Zhou et al. (2019) and Manesha et al. (2021). This is linked to the fact that these approaches often rely solely on the material coverage of an area.
To what degree urban area delineations based on remote sensing imagery overlap with the characterization of urban areas based on the concept of "Built-up area" varies. However, it is not always just the presence of buildings that is considered, but rather the combination of building-covered and road-covered areas, which together depict the regions impacted by soil sealing (Schwick et al., 2018). The use of soil sealing to describe urban areas is prevalent in research on urban sprawl, especially when it is based on classified imagery derived from processed remote sensing data (Schwick et al., 2018; Zhou et al., 2019). In these cases, the crucial land cover for the detection of urbanity are areas characterized by constructed materials like asphalt or concrete, including residential, commercial, industrial and transportation lands (Zhou et al., 2019). Relying solely on the material covering the area using raster data can lead to two opposite effects: depending on the situation, the context in which a real-life object, like a road or a building, is located, is either irrelevant for the classification or is decisive for it (Ewing & Hamidi, 2014; Soulard et al., 2018). This depends on the spatial resolution available to do the analysis. When the classification of the area is done with large cells, phenomena like single buildings surrounded by natural vegetation easily remain undetected. However, if the cells are sufficiently small, the spectral signature of these objects is strong enough to be detected. Thus, the areas that can be assigned urbanity using raster data have to be both:

- Covered by a land cover specific for urban areas (like asphalt or concrete)
- Be sufficiently large in comparison to the cell size used to detect the land cover.

It can be concluded that using smaller cells leads to the delineation of urban areas in alignment with the concept of "Built-up area", while larger cells tend to encompass larger swaths of areas such as greenery, they overlook smaller areas covered by artificial structures. The urban area delineation in those cases would resemble more the two urban area definitions discussed in the following sections of this thesis.

To detect built-up areas, the direct use of classified land cover images is a prevalent method in urban sprawl literature (Ewing & Hamidi, 2014; Manesha et al., 2021; Zhou et al., 2019).

The use of automatically classified remote sensing data based on reflectance values is an important part of many definitions of urban areas used by urban sprawl researchers, like the urban areas classified by the United States Census Bureau (2022), for example used by Nelson (1999), the CORINE land cover map, for instance used by Steurer and Bayr (2020) and Salvati and Carlucci (2014), or the Florida land use and land cover classification used by Sim and Mesev (2011), and many more.

When defining urban areas using soil sealing thresholds, areas like roads outside settlements may or may not be classified as urban – merely depending on how detailed the map is meant to be (Environment Agency Austria, 2019; Soulard et al., 2018). Maps that may contain roads outside settlements are used in urban sprawl literature (Manesha et al., 2021; Zhou et al., 2019). However, Soulard et al. (2018) argue that areas covered by roads outside settlements should not be classified as part of the urban area, albeit their argument is not limited to urban sprawl measurement. Rather, urban land is generally distinguished from other forms of development based on its higher population density, higher building density, higher land use intensity, and/or more impervious cover (Soulard et al., 2018).

In contrast to this, roads inside settlements are still to be viewed as urban (Soulard et al., 2018). This point of view coincides with the stance taken by the map of the Urban Morphological Zones, a map created based on classes of the CORINE land cover map, intended to be especially useful for urban sprawl calculation (Copernicus, 2021).

The layer "Ortslage" of the ATKIS dataset (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 2021), used to calculate urban sprawl by Schwarzak et al. (2014) also excludes roads remote from the settlement, but includes the ones in close functional context to them. The effect of the decision to either include or exclude roads outside settlements as part of the urban area can be seen in Figure 6.

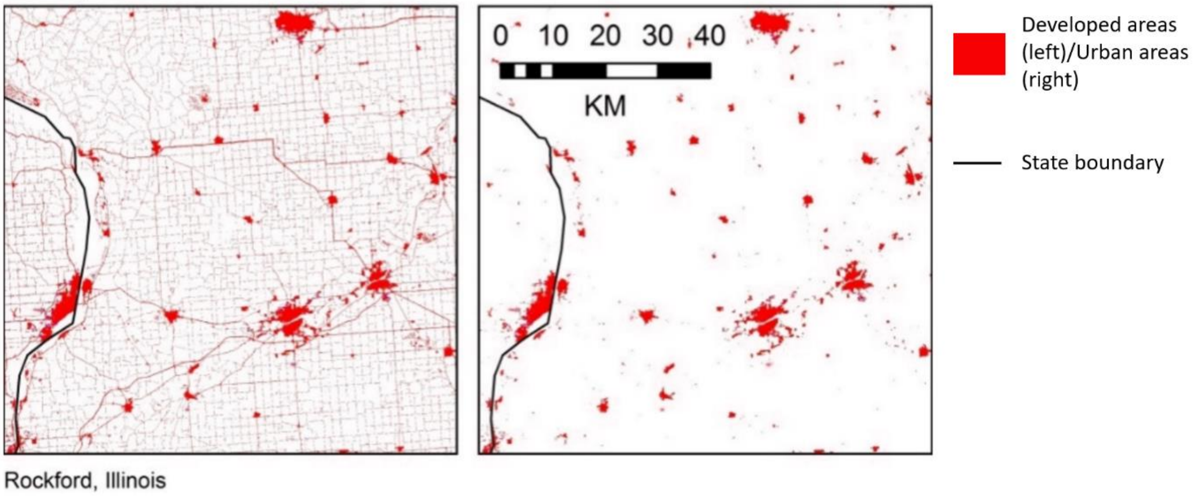



Figure 6: The effects of rural roads on the size of developed areas: excerpt of the results of Soulard et al. (2018). On the left are the developed areas (which includes rural roads), on the right are the urban areas obtained after the removal of rural roads (and other small built-up areas).

This thesis defines the term "Built-up area" as excluding areas covered by roads (or railways, too) that traverse rural or natural regions. Although the practice of including roads in urban area delineations exists, it is primarily restricted to specific classifications based on remote sensing imagery.

## 2.4.3 Area With Strong Anthropogenic Influence

### 2.4.3.1 Overview and Definition

Although many urban area definitions use the built-up area as a base for their urban area delineations, it is very common to complement this area with certain other areas. This "certain other areas" are very diverse and mostly provided as a list by urban sprawl researchers. These lists are quite different to each other when looking at them in detail, but it is in general possible to subsume these areas, in combination with the built-up area, as the area with strong anthropogenic influence.

This perspective on urban areas is prevalent in pre-constructed land cover maps utilized by researchers studying urban sprawl. Although these maps often feature a complex classification system with various rules and exceptions, they are not specifically designed for assessing urban sprawl, but are meant to serve as a basis for many analyses on land cover and its change. Nevertheless, the use of such maps is very commonplace in urban sprawl research and thus is strongly influential on the outcomes of urban sprawl calculations.

The viewpoint of urban areas currently under discussion is referred to as "Area with strong anthropogenic influence", given it encompasses not only built-up areas, but also certain other territories impacted by human activity. The term "Area with strong anthropogenic influence" used to describe this viewpoint on urban areas should not be rigidly interpreted, as some territories, such as agricultural fields, which are unambiguously shaped by human activity, are typically not included in the urban area. Other areas, such as certain forests, also belong to this category, even though they may experience less human impact in some aspects compared with agricultural fields. It encompasses not only built-up areas, but also includes green areas with different uses that accompany human settlements, for example golf courses, zoos, or urban green spaces. Natural and agricultural areas will generally be excluded from this class, unless they are in context with urban areas as defined in the previous sentences. If they are, such areas will be considered as a type of urban green space and thus are a part of the "Area with strong anthropogenic influence".

In urban sprawl literature, a clear boundary between the concept of "Built-up area" and "Area with strong anthropogenic influence" does not exist. Rather, these two classes represent two extremes of a continuous spectrum. In the following, several approaches will be mentioned, that are somewhere on this spectrum, highlighting the range of perspectives on urban area delineation and not just the extreme cases. The approaches will be mostly arranged in a sequence, starting with those closest in nature to the concept of "Built-up area" and progressing towards those that align more with the idea of "Area with strong anthropogenic influence". In the Chapters 3, 4, and 5 of this thesis, the focus will be on the extreme cases, disregarding intermediate variations as it is probable that similar results will arise when comparing similar urban area definitions.

In the course of the following literature review, it will become clear that the method chosen to calculate sprawl also has an influence on the perception of urban area. This means, that for example an urban sprawl researcher using the Moran's I indicator as his main way to assess sprawl may have other opinions on where an urban area should begin and end than one focused on doing the calculation using mainly density measures.

#### 2.4.3.2 Literature Review

In theoretical discussions regarding urban sprawl, it becomes apparent, that urban area can be defined by more than only built-up areas. Galster et al. (2001), for example, opine that the presence of certain areas interrupting the continuous development patterns of urban areas should not result in a higher detection of sprawl in continuity measures. They refer to water bodies, wetlands protected by conservation laws, forests, parks, hilly or rocky terrain, freeway exits, and public spaces as examples of such areas. This effectively means that those areas should be included in the urban area. Galster et al. (2001) argue, however, that the way to handle green belts and open spaces is not self-evident, as their contribution to urban sprawl is disputed.

The fact that various urban sprawl researchers use different definitions of urbanity, specifically when calculating density, either excluding or including non-residential areas like cemeteries, disposal sites and industrialized areas, parks, water, bits of wetlands, or deserts was recognized as a disagreed topic by Tsai (2005) who, however, refrained from judging the different approaches.

When studying urban sprawl literature, it becomes apparent that discussions about what areas to include in the urban area are not widespread, and that the different opinions on the delineation of urban areas must instead be directly extracted from the approaches taken in urban sprawl literature. Thus, in the following, different practical approaches to delineate urban areas will be presented, that deviate from the narrow perspective of characterizing urban areas solely as built-up areas.

Some approaches largely adhere to the principle of defining an urban area based on the built-up area. An example is the one used by Schwick et al. (2018). This approach, however, also incorporates small parks as part of the definition. This is influenced by the thought that otherwise, infill would come with a larger area, which would lead to a higher sprawl value when using WUP to calculate urban sprawl (Schwick et al., 2018).

The layer "Ortslage" of the ATKIS dataset, used by Schwarzak et al. (2014) is more generous: Apart from areas with residential, commercial, and institutional character, certain areas in close functional context to the mentioned areas are included. Such areas can consist of vegetation, areas used for traffic, areas of waterbodies, areas that contain buildings/installations for sport, recreation or relaxation, stadiums, sport fields, swimming pools, inruns of ski jumping hills, shooting ranges, and game enclosures (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 2018).

An influential classification scheme, serving as basis for different land cover maps (Soulard et al., 2018) is the classification by Anderson et al. (1976). Its class "urban and built-up land" contains residential areas (ranging from residential strip developments to villages and cities), areas used for transportation (like highways, railways and airports), commercial areas, extensive parts of recreational areas (like golf courses and ski areas), zoos, surface structures associated with mining operations, waste dumps, urban parks, cemeteries, the residential parts of farmsteads and more (Anderson et al., 1976). In general, the "urban and built-up land" could be explained by the land strongly influenced by human activity, whenever this activity is not directly related to agriculture (Anderson et al., 1976). This represents a departure from the notion of urban areas being limited to the built-up area, instead requiring the consideration of a range of other areas. If this characterization is the right one for urban sprawl analysis cannot be clearly answered. In practice, on the one hand, urban sprawl research was conducted based on the classification scheme of Anderson et al. (1976), for example by Ewing and Hamidi (2014) or Musakwa and Van Niekerk (2013). On the other hand, its characterization of urban area was only partly followed by other researchers, as in the case of Schwarzak et al. (2014) or Steurer and Bayr (2020), which did not include all the mentioned areas into their urban area delineations.

The practical implementation of Anderson et al.'s (1976) classification by Musakwa and Van Niekerk (2013) illustrates the deviation from the definition of urban areas as merely built-up areas: They included areas with cluster housing, commercial facilities, community establishments, residential areas, educational institutions, governmental buildings, industrial sites, informal settlements, areas with mixed use, office parks, recreational areas, smallholdings, transportation hubs, and vacant land as urban areas and other areas, if they consume agricultural land (Musakwa and Van Niekerk, 2013).

The viewpoint that built-up areas are insufficient for defining urban areas is also reflected in the "Urban Morphological Zones" map, specifically created for detecting urban sprawl (European Environment Agency, 2011). This map is created using polygon shapes stemming from the CORINE land cover dataset. Instead of just relying on the classes of "continuous urban fabric" and "discontinuous urban fabric" like Steurer and Bayr (2020), the map of the "Urban Morphological zones" includes urban fabric, industrial units, commercial units, and green urban areas (European Environment Agency, 2011). Existing holes in the formed area get filled up if their land cover consists of certain vegetation, for example forests (whereby the tree cover should make up at least 30% of the area), scrublands, moors, or heathland (European Environment Agency, 2011). Additionally, the urban area gets enlarged by adjacent ports, airports, and sports and leisure facilities (European Environment Agency, 2011). On top of that, certain parts of water streams (i.e., flowing water – either natural or artificial) are added to the urban area. This, as water courses can serve as links between urban areas (European Environment Agency, 2011). For the creation of the map of the "Urban Morphological zones", rivers lying between urban areas that are separated by not more than 300 m are added to the urban area. The same procedure is applied for parts of road or railway networks (European Environment Agency, 2011). The decision to use 300 m was made after several trials in different European cities with predominant rivers, roads, or railway tracks (European Environment Agency, 2011). For rivers with a width of more than 200 m, issues with the 300 m-threshold were observed (European Environment Agency, 2011). In such cases it should be decided if the river should be seen as a part of the urban area or not (European Environment Agency, 2011). The approach used for the map of the "Urban Morphological zones" (European Environment Agency, 2011) as well as the one by Steurer and Bayr (2020) connects urban areas lying less than 200 m apart. In the approach used to create the "Urban Morphological Zones" this was achieved by applying a buffer of 100 m and followed by a second buffer of -100 m (European Environment Agency, 2011). In conclusion, it can be stated, that the map of the "Urban Morphological Zones" (European Environment Agency, 2011), and to a lesser degree the delineation by Steurer and Bayr (2020), reject a delineation of urban areas based solely on the presence of built-up areas and opt to include additional areas.

Olazabal and Bellet (2017) deem the class "artificial areas" of the CORINE land cover classification appropriate to assess urban expansion, as it does not only view residential areas as urban entities, but also includes business/industrial areas, ports, airports, road and railroad infrastructures, mines, dumps and landfills, construction areas, urban green spaces, and infrastructure for recreation or sport (Olazabal & Bellet, 2017). The definition of green areas included in the urban area is broad and encompasses not only publicly accessible greenery, but also the inner spaces of city blocks and vegetated areas that could potentially be used for recreation, despite not being their primary function (Environment Agency Austria, 2019). Thus, not just residential use is seen as urban, but also other areas in the urban context that allow the daily functioning and internally connect the city (Olazabal & Bellet, 2017).

The urban area can also be defined as those areas characterized by buildings of residential, public and administrative use, and certain parks (Environment Agency Austria, 2019; Salvati & Carlucci, 2014; Salvati & Sabbi, 2011). This definition, apart from being used in studies regarding urban sprawl, for example by Salvati and Carlucci (2014) and urban expansion, for example by Salvati and Sabbi (2011), coincides with the classification scheme in land cover maps directly. Maps like these often contain classes such as "artificial areas" or "urban areas", which are further sub-classified into more specific areas, e.g., residential areas. As is the case in urban sprawl research, the term "urban areas" is used for a variety of different areas. In a land cover or land use map, for an area with low soil-sealing to be incorporated into a class of artificial or urban areas, no visible intensive use is always necessary. For instance, according to the Swiss Arealstatistik classification, open forests in urban areas are considered as public parks and therewith classified as a sub-category of urban areas (Bundesamt für Statistik, 2006). The same goes for other publicly accessible areas that can be recreationally used, like zoos and meadows (Bundesamt für Statistik, 2006). The classification of areas as urban or artificial can vary greatly depending on the land cover/land use map. Some maps may classify areas such as mines, landfills, parks, zoos or other recreational areas – including those with buildings, like holiday cottages – as urban, while others may place them in a separate category of artificial areas. These differences in classification become apparent when comparing the land cover and land use maps of the ATKIS of Germany (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 2018), the Arealstatistik of Switzerland (Bundesamt für Statistik, 2006), the SIOSE AR of Spain (Equipo Técnico Nacional SIOSE, 2022), the COS of Portugal (Direção-Geral do Território, 2019) and the Cooperative Land Cover map of Florida (Redner & Srinivasan, 2014).

Returning to urban area delineations utilized in urban sprawl research, Sim and Mesev (2011) used the map provided by the Florida Fish and Wildlife Conservation Commission (Gilbert & Stys, 2004). This land cover map and its updated versions were designed primarily to assess the habitats of wildlife (Gilbert & Stys, 2004; Kawula & Redner, 2018; Knight et al., 2010). Generally speaking, urban areas are defined here as those areas that are not suitable for rural or natural wildlife habitats, including for example parks (Kawula & Redner, 2018). Thus, it is apparent, that here, too, built-up areas do not fully characterize the delineated urban areas.

In addition to Petrescu (2019) and Prastacos and Lagarias (2016), whose approaches are explained in Section 2.4.2.2, Lagarias and Sayas (2018) also utilized "Urban atlas" data. However, in their urban sprawl research, Lagarias and Sayas (2018) utilized a different set of areas for their urban area delineation. In addition to areas including residential buildings and isolated structures, they also utilized the classes "industrial, commercial, public, military and transport units", "green urban areas", and "sports and leisure facilities". Thus, their definition of urban areas extends beyond just built-up areas, too, and is more in line with the definition of "Area with strong anthropogenic influence".

It has to be mentioned that many maps have minimum mapping units, i.e., areas smaller than a specific threshold are classified based on their surrounding area. This means that it is very well possible, that small areas, like small parks are not classified as such, but rather as a part of a residential area. Therefore, even definitions that only rely on built-up areas in practice often also include additional areas, solely due to the scale at which they are delineated.

As evident from the various approaches adopted by urban sprawl researchers, a significant number of them move away from the narrow definition of urban areas as just built-up areas. The extent to which additional areas are included in the urban area varies greatly. Whereas some approaches are similar to the approach of using built-up areas and including only few additional areas, other approaches have a much more generous conception of urban areas. This entire spectrum of viewpoints can be supplemented with additional perspectives regarding the exclusion of specific areas, as elaborated in the subsequent section of this thesis.

## 2.4.4 Non-Rural Area

### 2.4.4.1 Overview and Definition

In addition to the question if urban areas are best described by the built-up area or rather should be extended to incorporate further areas, there are also restrictive viewpoints regarding the delineation of the urban area.

As discussed in the previous section of this thesis, the notion of urban areas as areas with strong anthropogenic influence can be seen as an extension of the concept of urban areas as built-up areas. However, various researchers also use maps of urban areas that exclude certain areas that would be present using a narrow approach of these two ways to delineate urban areas.

The decision to exclude certain areas can be based on the perspective of defining urban areas as either "Built-up area", or "Area with strong anthropogenic influence", or anything in between. That is to say: Regardless of whether an urban sprawl researcher considers urban areas to be solely built-up areas or to encompass additional areas, they may hold the belief that certain areas with these characteristics should not be classified as urban. Thus, the term "Non-rural area" is probably best described as an additional dimension, that exists in addition to the spectrum between "Built-up area" and "Area with strong anthropogenic influence".

The now presented concept is called "Non-rural area", as it excludes those areas from the urban area delineation that are isolated from larger groups of buildings or other urban areas. Despite the name that was given to this perspective on urban areas, it not only excludes entities like farmsteads and barns but also other solitary buildings such as holiday cottages. It is also irrelevant if these structures are in a rural or in a natural context.

Just like for the previous perspectives on urban areas, there is a spectrum of how closely this viewpoint is adhered to in practical urban sprawl research.

In the following literature review, the concept of "Non-rural area" will be explored regardless of whether the remainder of the urban area delineation is based on the idea of "Built-up area" or "Area with strong anthropogenic influence". This is in contrast to the definition used in the rest of this thesis, where "Non-rural area" will always be based on the concept of "Area with strong anthropogenic influence". This has four reasons:

*First*, in theory, the exclusion of certain built-up areas from the definition of urban areas based on built-up areas appears incongruous, although in practice, many approaches only use a selection of built-up areas.

*Second*, the concept of "Area with strong anthropogenic influence" does not per se clash with a viewpoint of excluding certain areas, as depending on how one might define "strong", individual buildings could be considered as having minimal impact on the landscape.

*Third*, in this study, the impact of excluding certain areas was only planned to be analyzed once, requiring a decision to be made regarding the starting point for the exclusion of the selected areas.
*Fourth*, the exclusion of certain small areas based on the concept of "Area with strong anthropogenic influence" leads to especially compact urban areas, whose effect on urban sprawl measurements could serve as a good comparison to the other two concepts of urban areas presented in this thesis.
The frequent perception of urban areas as something opposed to rural territory was also noted by Salvati and Carlucci (2014), which state that in land cover change literature "… urban land uses have often been lumped together in such a way that 'urban' itself means little more than 'non-rural'." (Salvati & Carlucci, 2014, p. 2).

### 2.4.4.2 Literature Review

In urban sprawl literature, theoretical arguments exist, that call for the restrictions mentioned in Section 2.4.4.1. According to Lopez and Hynes (2003), housing- and recreation-oriented developments that exhibit densities of one dwelling unit per five (or more) acres (corresponding to one dwelling per 2.023 ha) should be excluded from urban sprawl calculations. Such developments, deemed as rural by Lopez and Hynes (2003), should instead be assessed separately, as the sprawl in rural areas is a distinct phenomenon. The exclusion of dwellings in less densely populated areas is also discussed by Ewing and Hamidi (2014), which favor the focus on areas with population densities of suburban areas and higher (≥ 100 residents per square mile, corresponding to ≥ 0.386 residents per ha) (Ewing et al., 2003). In this case, dwellings at lower densities were excluded, as they would result in a distorted average of the urban sprawl metrics (Ewing et al., 2003; Ewing & Hamidi, 2014). Apart from proponents of the idea that urban areas shall include all areas which are built-up or have strong anthropogenic influence, the opinions vary on the extent to which the urban area should be reduced. For example, Galster et al. (2001) warn from the exclusion of semirural development at the urban fringe that some consider the epitome of sprawl.

In practice, a large number of urban sprawl scientists base their calculations on areas which are exempt from low-density development. An example is the urban sprawl research by Fulton et al. (2001), which used the *NRI* (National Resources Inventory) dataset, that does not categorize isolated groups of less than four dwellings as urban (P. E. Flanagan, personal communication, September 9, 2022). Another one is the study of Sim and Mesev (2011), that used the map of the Florida Fish and Wildlife Conservation Commission (Gilbert & Stys, 2004), which does not regard agricultural buildings as urban. A further example is the urban sprawl measurement by Schwarzak et al. (2014), that is based on the ATKIS data of "Ortslage", which exclude isolated structures (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 2021). In addition, due to the spatial resolution of remote sensing imagery or the minimum mapping units in land use and land cover maps, isolated structures very often are systematically excluded from the urban area (Copernicus, 2021). For instance, due to the minimum mapping unit of the CORINE land cover map being at 25 ha, a lot of details, including individual structures, are lost (Copernicus, 2021). Subsequent sprawl calculations thus systematically exclude those areas.

### 2.4.5 Concluding Remarks

The extent to which the different urban area delineations in literature adhere to the three presented archetypes of urban area delineations: "Built-up area", "Area with strong anthropogenic influence", and "Non-rural area" varies greatly. Whereas methods using remote sensing imagery or the approach by Schwick et al. (2018) can be said to adhere quite strictly to a delineation sticking to the principle of "Built-up area", the areas used for the calculations of Musakwa and Van Niekerk (2013) and Sim and Mesev (2011) are better characterized as areas with strong anthropogenic influence. At the same time, urban areas could be viewed as only those territories where the prevalence of built-up areas or human influence reaches a certain threshold. Approaches using remote sensing imagery to delineate urban areas or the land cover map used by Sim and Mesev (2011) show, that this restriction of urban areas does not necessarily contradict the proposed archetypes of "Built-up area" and "Area with strong anthropogenic influence", but rather present a different viewpoint on the scale of urban sprawl. For instance, maps based on low spatial resolution remote sensing imagery assume that low-density development falling below a certain threshold are not built-up areas. Similarly, the urban area delineation used by Sim and Mesev (2011) for their urban sprawl calculation excludes buildings at such low densities, that they basically pose no threat to nature. Otherwise, the urban area delineation utilized by them closely adheres to the concept of "Area with strong anthropogenic influence" (Stys et al., 2004).
The presence of different perspectives on urban area among urban sprawl scientists becomes especially apparent, when several of them use the same land cover or land use maps that are not tailored for the use for urban sprawl research. In such cases, urban sprawl researchers often try to correct for misclassifications in the original land cover or land use maps, as they are aware of the limitations these maps present for urban sprawl research, doing so from an urban sprawl perspective. That often leads to different urban area delineations. This effect can be seen in the example of the use of the CORINE land cover dataset to calculate sprawl. Its land cover classes representing urban fabric refrain from including much more vegetation than the one present due to gardens around residential buildings (Environment Agency Austria, 2019). Whereas Steurer and Bayr (2020) do not make an effort to include vegetated areas into

the urban area, but rather hold the opinion that these areas should be excluded from the urban area, Salvati and Carlucci (2014) enrich the urban fabric areas with territories of the class "green urban areas" of the CORINE land cover map, which contains areas like inner spaces of city blocks, green grounds of mansions or forests inside the city area. The approach by the Urban morphological zones (European Environment Agency, 2011) is different again, including a step where certain natural areas, like forests or moors are incorporated into the urban area.

In sum, the presented literature review shows, that a consistent, generally accepted definition of the urban area for sprawl measurement is missing. The absence of a clear definition what constitutes an urban area does not only hinder urban sprawl researchers to measure urban sprawl using a homogeneous definition, but also inhibits the creation of generally accepted urban area maps, that can be used for urban sprawl measurement. Thus, this thesis aims to present the effects of different urban area definitions, specifically testing the delineations of the "Built-up area", the "Area with strong anthropogenic influence" and the "Non-rural area" for their effects on the measurements of WUP, Moran's I, and density.

## 2.5 Previous Research on the Effects of Urban Area Delineation on Urban Sprawl Measurements

The literature on the effects of different urban area delineations on urban sprawl is very sparse. Quantitative observations of the effects of urban area definitions were not found in the literature search made for this thesis, besides the one by García-Álvarez and Camacho Olmedo (2017). They noticed differences in the results of urban sprawl measurements using two different maps of urban areas, namely the CORINE Land cover of Spain and a generalization of the SIOSE map (García-Álvarez & Camacho Olmedo, 2017). However, these two different vector maps were not only constructed using different methods but also at different scales (García-Álvarez & Camacho Olmedo, 2017). In addition, the employed urban sprawl measure was fractal dimension (García-Álvarez & Camacho Olmedo, 2017). Together, this would lead to the statement that spatial objects delineated at a larger scale, are more complex – a statement that makes sense logically as the details of vector data grow with a larger map scale. However, it does not answer the question if different definitions of urban area affect the sprawl measurement.

In order to give more insight into the impact of different perspectives on urban area delineations on measurements of urban sprawl, this thesis aims to show the discrepancies that arise from these differing stances.

## 2.6 Techniques to Delineate Urban Areas

Delineations of urban areas that are used for urban sprawl research are mostly based on one of two techniques: Either pixels are classified, for example using artificial intelligence (see, e.g., the approach by Manesha et al. (2021) or Zhou et al. (2019)), or the delineations are human-made upon visual inspection of photo interpreters (see, e.g., the approach by Schwick et al. (2018) or by the CORINE land cover map (Environment Agency Austria, 2019). The analysis of the approaches used to delineate these maps gives many insights on which areas to potentially include into the urban area, but contain very limited information that is useful for an exact delineation of the urban area. For example, the approach using pixels classified with artificial intelligence sets the boundary between urban area and non-urban area at the borders of the pixels and is limited in their capability to detect any anthropogenically manipulated non-built-up area. The visual interpretation approach, in contrast, relies on human interpretation and the ability to distinguish between areas with varying levels of human influence, such as gardens, meadows, fields, orchards, forests, or buildings, without being subject to pixel geometry.

One map that was used for urban sprawl research contains especially helpful information when trying to delineate urban areas on the basis of building footprints, like in this thesis: the NRI dataset (United States Department of Agriculture, 2022), which was used by Fulton et al. (2001) and Ewing and Hamidi (2014) in their research. Although primarily produced using human visual interpretation, it contains some quantitative information useful for automatic urban area delineation. First, a specific threshold is used to distinguish urban buildings from non-urban ones: If less than four otherwise individual buildings are near to each other, they do not contribute to the urban area. What "near" means is also specified: the houses need to be connectable using a distance of maximally 480 m (P. E. Flanagan, personal communication, September 9, 2022). Second, the size of the areas surrounding the houses that should be included in the urban area is specified. The houses of residential areas are given a different surrounding area depending on the region they are in. However, in each case, hexagons are used to delineate the surroundings of a house. In the east of the USA, a hexagon with 2.481 acres is used (which is 1.004 ha), in the western states (Alaska included) 2.422 acres is used (which is 0.988 ha), 2.740 acres in central states (which is 1.109 ha) and 0.880 acres (which is 0.356 ha) in the Caribbean region (P. E. Flanagan, personal communication, September 9, 2022). The use of a circle buffer preserving these area thresholds would result in a radius of 56.532 m for western states, 56.086 m for eastern states, 59.410 m for central states und 33.669 m for the Caribbean.

In an unpublished report of an approach in Switzerland by the Federal Statical Office (Bundesamt für Statistik, Sektion GEO, 2007), it is proposed to use building centroids to create the delineation of urban areas, with distances between buildings lying maximally 60 m apart in one version and lying maximally 110 m apart in another version.

Another approach for urban area delineation is presented by Kalinina et al. (2018) in their research. In this study, the $\alpha$-shape algorithm was utilized to delineate urban areas, although not for the purpose of calculating urban sprawl. For this, they used point-data of OpenStreetMap to create polygons, using an $\alpha$-shape algorithm distinct to the one used in this thesis. In that study, this approach was used to delineate what they deemed as urban space, excluding for example waterbodies, green zones and industrial zones (Kalinina et al., 2018). The algorithm used by Kalinina et al. (2018) works with a parameter on the degree of concavity to construct their polygon. Therefore, they do not use an input value representing the radius of a circle to construct the connections between the points, as is the case in this thesis, whose approach will be presented in Chapter 3. This makes it more difficult to find the appropriate input value, as the chosen distance between buildings or other urban areas cannot be directly used as input values for the algorithm.

The $\alpha$-shape algorithm was also used by Arribas-Bel et al. (2021) to delineate urban areas, although in this case, again, this was not done in relation to urban sprawl research. In the study of Arribas-Bel et al. (2021), the DBCAN algorithm was employed to filter out those areas with a low building density, and in a second step the $\alpha$-shape algorithm was employed to connect these buildings to each other to construct an urban area. The $\alpha$-shape algorithm was based on fitting a circle between points (Arribas-Bel et al., 2021; Arribas-Bel & Wolf, 2022). More precise information on this concept can be found in Chapter 3. In this case, the radius of the circle was not fixed, but varied in order to construct the smallest single polygon based on all the points in a DBSCAN-cluster, which was performed using a distance threshold of 2000 m (Arribas-Bel et al., 2021; Arribas-Bel & Wolf, 2022).

Previous research has also employed the use of concave hull algorithms to quantitatively assess urban expansion patterns (Wang et al., 2021).

## 2.7 Research Question and Research Gaps

As presented in this chapter, a wide variety of urban areas are employed to measure urban sprawl. The choice of the urban area is generally treated as an unimportant aspect of urban sprawl research, and the rationale behind the choice is often non-existent. The research question of this thesis aims to uncover potential differences in urban sprawl values stemming from varying urban area delineations:

**RQ: Are there differences in urban sprawl values, depending on the used definition of urban areas?**

In addition to addressing this overlooked aspect of urban sprawl research, this thesis will also shed light on the variations in urban sprawl results that arise due to the use of different calculation methods. However, since it is already recognized that the choice of the urban sprawl calculation method influences the measured urban sprawl, this thesis cannot offer novel perspectives on this matter. It can only quantify and illuminate the discrepancies stemming from the use of different sprawl methods, specifically from a Swiss perspective.

# Chapter 3 | Methods

## 3.1 Opening

This chapter explains the methods employed in delineating the urban areas, detailing the required data, the decisions taken, and the other necessary steps. Furthermore, the methods employed to obtain the urban sprawl results will be explained, including the process for obtaining and preparing the data required for the calculations. This shall pave the way to answer the research question if there are differences to be found in urban sprawl values, depending on the used definition of urban areas.

For a clearer understanding of the explanations that follow, it is important to repeat here that in this thesis, the three different definitions of urban areas mentioned before were considered: "Built-up area", "Area with strong anthropogenic influence", and "Non-rural area". These definitions were then applied to delineate the urban area and compared with one another using three different calculation methods.

As already mentioned in Chapter 2 of this thesis, in practical applications, the border between the archetypical viewpoints of "Built-up area", "Area with strong anthropogenic influence", and "Non-rural area" are blurred. For this thesis, the extremes of these takes of urban areas will be compared with each other. For example, the urban area delineation "Area with strong anthropogenic influence" will be based on its strict sense. Thus, the differences between the three viewpoints shall become clearer, as this thesis aims to show the differences stemming from the different existing opinions on how to delineate urban areas.

The common techniques to delineate urban areas for urban sprawl calculation cannot be directly implemented in an automated approach, unless one is ready to accept certain reductions in quality, as can be seen in Chapter 2. Thus, it was decided to delineate the urban areas using a different method than the common approaches in urban sprawl literature. The choice fell on primarily using an algorithm connecting nearby points to delineate the urban areas. For this, processed pre-existing vector data on building footprints and certain areas commonly associated with urban areas were used, based on the concepts of urbanity discussed in Chapter 2. A general overview of the delineation and measurement approach in this thesis can be seen in Figure 7. Prior to that, several decisions and validations had to be made, which will also be explained in this chapter.

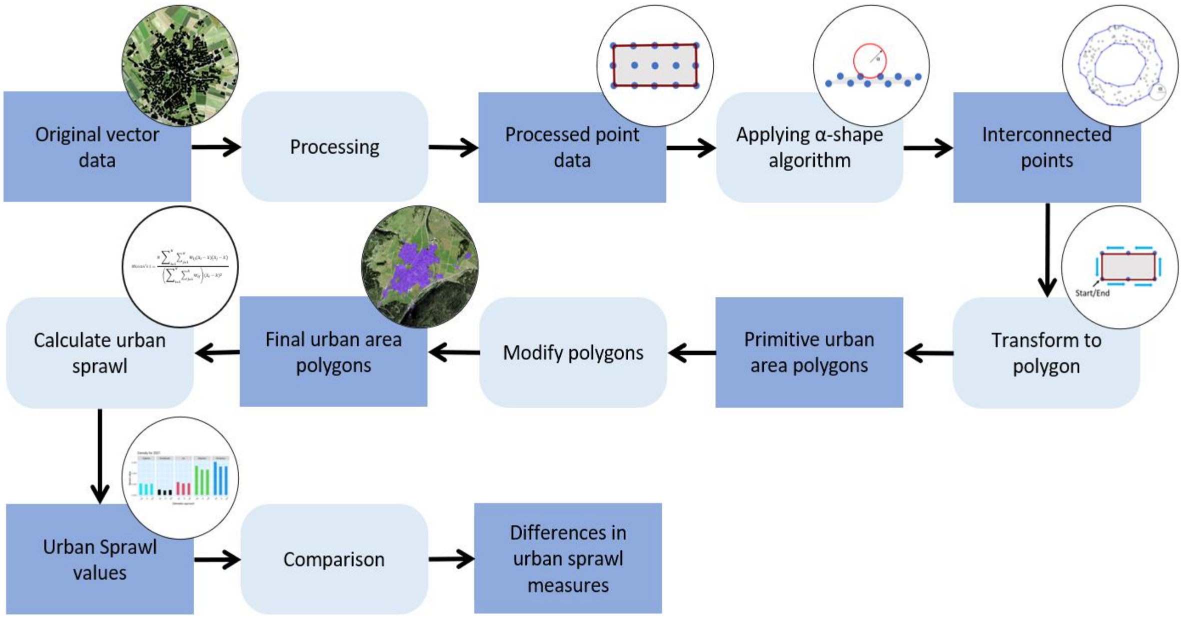



Figure 7: General approach of delineating and measuring urban sprawl in this thesis. (Image source: own work, containing imagery from Pateiro-López and Rodríguez-Casal (2010) and Environmental Systems Research Institute (2022))

The choice to create the urban areas anew, instead of relying on already existing urban area delineations for Switzerland, like the one of the CORINE land cover (Environment Agency Austria, 2019) or the Arealstatistik (Bundesamt für Statistik, 2006) has multiple reasons:

*First*, the comparison of a set of different land cover and land use maps would mean that it would be necessary to rely on the definition of the respective land cover/land use map. Thus, it would not be possible to freely choose the definition that is to be examined, but one would have to rely on the definition of the land cover/land use map. In addition, comparing different land cover/land use maps does not only lead to the comparison of the different definitions of urban areas, which is the aim of this thesis, but aspects like the different values of accuracy and precision of the respective maps would inevitably influence the result. This could raise fears that, instead of measuring the effects of different urban area definitions, the effects of using different accuracies when delineation of urban areas would be shown.

*Second*, the already existing land cover and land use maps were found to be too imprecise, i.e., omitting or inexactly depicting certain small-sized phenomena, like individual buildings.

*Third*, the border between urban and non-urban areas was deemed too inaccurate in the already existing delineations.

*Fourth*, the a priori grouping of multiple different phenomena into one single class present in the already existing maps forces the urban area to be made up by these multi-component classes, instead of enabling a delineation of urban areas based on self-chosen territory.

## 3.2 Data

The data utilized for the urban area definition in this thesis originate from the TLM3D (Bundesamt für Landestopografie, 2022) of Switzerland. Although a new version of this data is published yearly (Bundesamt für Landestopografie, 2022), a full revision of the data is only achieved every six years (Bundesamt für Landestopografie, 2023). The last two times such 6-year cycles were accomplished, were in the years 2015 and 2021 (Y. Maurer, personal communication, December 21, 2022). The data of the year 2021 could be downloaded from the official download site (Bundesamt für Landestopografie, 2022), whereas the data for the year 2015 were provided by Yves Maurer from the Swiss Federal Office for Spatial Development, as the official download site did not provide this outdated version anymore. For the urban area delineation, the data on building footprints, railway tracks, and other area deemed necessary to delineate the urban area were utilized.

The specific areas used for the urban area delineation depend on the different versions of the map and will be explained when describing the methods for the different versions of urban area delineation used in this thesis. In addition to the material of the TLM3D, the data on the uninhabitable areas, created and provided by Yves Maurer, were used. These data were necessary for the urban sprawl calculation using WUP and Moran's I and represent areas that are uninhabitable and thus should be excluded from the urban sprawl calculation. The delineation of these areas was based on the characterization of uninhabitable areas by Schwick et al. (2018): flowing and stagnant waters, glaciers, firn, unvegetated areas, areas with a gradient of more than 45°, and protected areas, i.e., forests, scrub forests, woods, national parks, floodplains, moors, dry meadows, dry pastures, and World Natural Heritage sites. The population data used to measure the density were taken from official statistics of the areas of interest in the respective years (Bundesamt für Statistik, 2022f) whereas the data of the population and employment used for the WUP calculation consist of unpublished data of the Federal Statistical Office (Bundesamt für Statistik, 2022a, 2022b). Specifics on the population and employment data utilized to calculate the different urban sprawl measurements will be given in the sections of the methodology.

Initially, based on the newest classification of the TLM3D-data, certain areas and building types were planned to be excluded from the urban area. For example, greenhouses were planned to be excluded from certain or all delineation methods in this thesis. However, greenhouses and other entities were not delineated as a separate class in the year 2015. Additionally, several areas that could be considered as "urban" were newly delineated in 2021. Certain classes were newly introduced, resulting in previously unmapped entities being newly delineated. Other classes were merely formed by splitting up an already existing class. The discrepancies in classification meant that the classification from 2015 had to be used as a reference to determine which building types or areas could be considered part of the urban area. Otherwise, the observed changes in the urban area could be attributed to the method of classification used in the different years or to newly added features, instead of actual changes in the region.

During the analysis, some data quality issues were detected which are described in section 5.5.5.

## 3.3 Choice of Study Areas

The study areas consist of five municipalities in the country of Switzerland: Celerina, Ins, Grindelwald, Wetzikon, and Winterthur. These municipalities have a population of around 1600 inhabitants, 3700 inhabitants, 4000 inhabitants, 25'000 inhabitants and 120'300 inhabitants, respectively (Gemeindeverwaltung Celerina/Schlarigna, 2023a; Gemeindeverwaltung Grindelwald, 2023; Gemeindeverwaltung Ins, 2023; Stadt Wetzikon, 2021; Stadt Winterthur, 2023). To detect the effects of urban area definitions, different municipalities with distinct characteristics were used, in order to take into account different urban forms. The presented set of municipalities comprise three villages, one town, and a city. Ins, Wetzikon, and Winterthur are located in the densely populated Swiss plateau. Celerina and Grindelwald are located in the less populated alpine belt and their economies are, to a large part, dependent on tourism (Baur et al., 2014; Gemeindeverwaltung Celerina/Schlarigna, 2023a). Upon examining the form of urban development in these two alpine villages, it becomes evident that the arrangement of their buildings are distinct: Whereas Celerina is built in a very compact way – containing a center with very limited green areas and where the overwhelming number of buildings is situated and large tracts of rural and natural undisturbed of human settlement can be found – Grindelwald is characterized by its large tracts containing a mixture of grassland and scattered buildings. The five analyzed municipalities and their location are portrayed in Figure 8. All of them are located in the German-speaking part of the country. In Celerina, however, Romansh is spoken in addition to German (Gemeindeverwaltung Celerina/Schlarigna, 2023b).

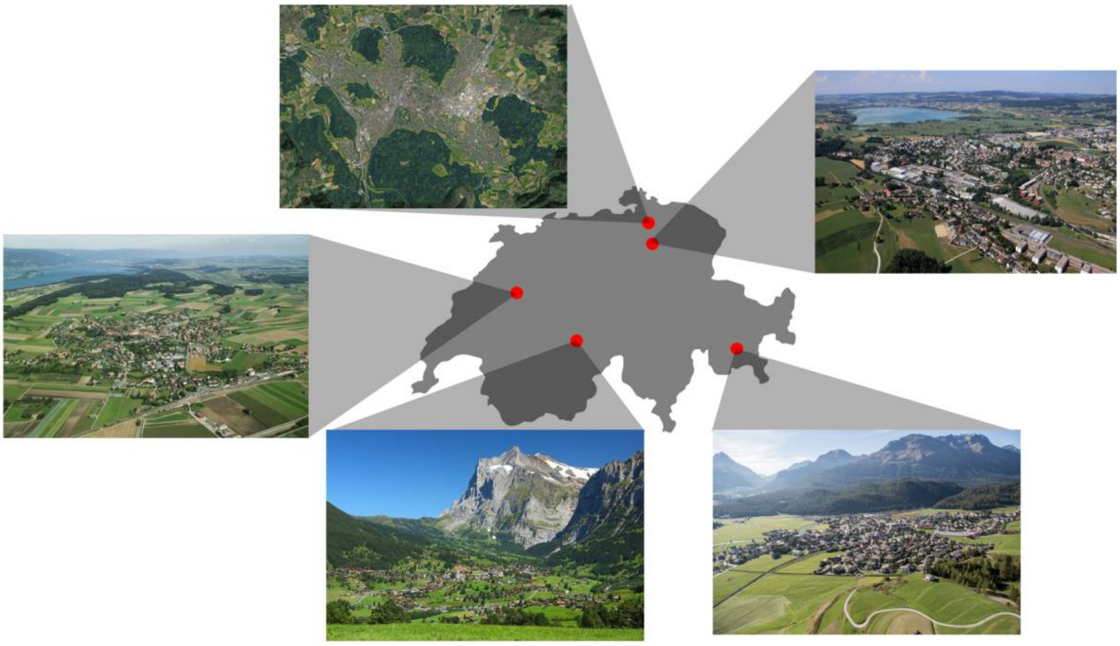


Figure 8: The five selected municipalities and their locations in Switzerland. (Image source: own image based on Fabian Gattlen (2023), Gemeinde Ins (2023), Google Earth (2021), Stadt Wetzikon (2023), Vemaps (2019), Viator (2023))

Urban sprawl calculations are possible at a wide range of different scales, municipalities only being one of the options to choose from. The reasons to choose municipalities as the study areas are the following:

1) Municipalities are subject to political decision processes, meaning that the results of urban sprawl calculations can directly be translated into political decisions. This is for example not the case for the totality of urban areas lying in specific biogeographical regions.
2) They are relatively small. The time needed to run the automatized approach to delineate urban areas used in this thesis depends on the number of features inside each area. Larger areas usually contain more of these features, meaning that the time to run the urban area delineation would be longer.
3) Using the borders of municipalities as the limits of the areas of interest usually leads to the effect that the edges of the area of interest lie in natural or rural areas. This prevents that the borders of the area of interest lie inside urban areas. However, there are certain municipalities that do not end in natural or rural areas, but instead directly border on other urban areas. This is for example the case for the city of Zurich, where the urban areas of the municipality of Zurich touch the urban areas of surrounding municipalities. Thus, to prevent cutting the urban area into pieces it would be necessary to include the urban areas of the surrounding municipalities that are part of the urban agglomeration. This effect is also visible in other big cities of Switzerland like Berne, Basle, or Geneva. Of the big cities of Switzerland, Winterthur is the only one

where the municipality border did not intersect with large tracts of urban area. For this reason, the big city chosen for the calculation in this thesis was Winterthur.

4) Calculations on municipalities are fairly reproducible and do not require lengthy additional explanations on how the area of interest was chosen. They are also fairly stable over time or at least their borders are mostly not subject to personal interpretation.

## 3.4 Validation of Existing Parameters to Construct Urban Areas

In order to form a contiguous urban area, it had to be determined which exact areas to use for the urban area delineation and the proximity required for them to be considered connected. The areas necessary for the urban area delineation do not only contain the areas described previously in the "Data" section of this chapter, but also certain additions to these areas. Specifically, the TLM3D-dataset does not contain any data on gardens and does not provide polygon data on large numbers of parallel railway tracks, like they are typical at or near railway stations. However, after the literature review, it became apparent that the inclusion of such areas into the urban area would help to delineate urban areas more in line with previous studies. Thus, to paint a more complete picture of the urban areas, it was necessary to correct for these missing data.

To determine the boundaries of residential gardens, a similar approach to that of the previously discussed NRI was considered. This involved utilizing one of the specified area sizes to assign surrounding areas to the houses. However, after visualizing the scenario of enlarging residences using the area size specified by the NRI, it was evident that the NRI's suggested surrounding area sizes for residences were excessive. Upon examination, in Swiss settlements, the sizes proposed by the NRI for the surroundings of a residence typically include not just the residence and its garden, but also large tracts of other adjacent areas. While the NRI's proposed values for the construction of residential surroundings may not be suitable for use in Switzerland, this does not discredit the NRI's approach, which could still be appropriate for use in the United States. Although still too large, the value provided for the Caribbean came closest to providing a reasonable threshold for the use in Switzerland. To find a more reasonable value to approximate the area of residential gardens, calculations were conducted with the approximate plot size of Swiss single-family dwellings, the mean number of inhabitants of such houses, the average living space per person in such houses, and the mean number of floors in this dwelling type. The footprint of the residential building $r$ is calculated like this:

$$r = \sqrt{p - \frac{i \cdot a}{f}}$$

where $p$ stands for the average plot size of residential houses, $i$ is the mean number of inhabitants in such dwellings, $a$ represents the average floor-space per inhabitant in residential houses, and $f$ depicts the average number of floors in such buildings. Assuming a square plot size and a square residential building in the center of the plot, the minimum distance between the parallel lines of the square plot and the square dwelling footprint, i.e., the minimum garden width $g$, can be calculated as follows:

$$g = \frac{\sqrt{p} - \sqrt{p - \frac{i \cdot a}{f}}}{2}$$

where the variables represent the same values as explained for the calculation of the footprint of the residential building. An illustration of some of the values can be found in Figure 9.

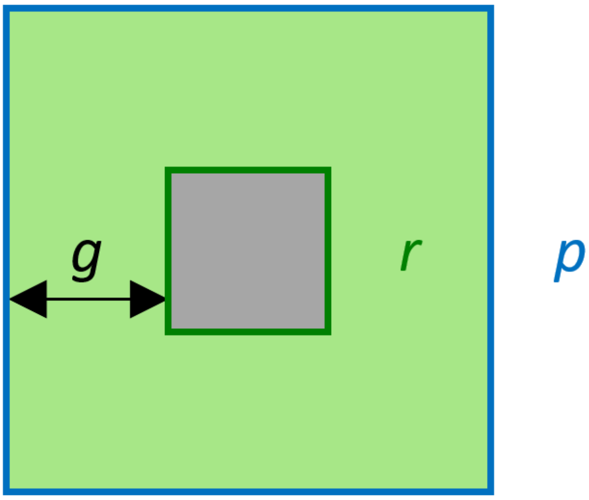



Figure 9: Illustration of the minimum garden width $g$, the footprint of the residential building $r$, calculated as $\sqrt{p - \frac{i \cdot a}{f}}$, and the average plot size of residential houses $p$.

The input values used in this calculation are 500 $m^2$ for $p$, an $i$ of 2.7 people, a value of $a$ of 55 $m^2$, all based on different federal sources (Bundesamt für Raumentwicklung & Fahrländer Partner AG, 2008; Bundesamt für Statistik, 2022c, 2022e), and an $f$-value of 2.21 floors, calculated using the prevalence of single-family dwellings with a certain number of floors (Bundesamt für Statistik, 2022d). The result of the calculation was a value of 7.08 m for $g$, which was rounded down to 7 m for the purposes of this thesis.

In the urban area calculation, also data from the TLM3D on railway tracks were employed. Underground railway tracks were excluded. The distance used to buffer railway lines was 4.5 m, which is the usual minimal spacing between railway tracks in Switzerland (Bundesamt für Verkehr, 2020). This distance, however, was not sufficient to create contiguous areas out of parallel railway lines. This contiguous area was created after the construction of the largest parts of the urban area. The method to include them into the urban area will be explained in this thesis during the discussion of the construction of the urban area in Chapter 3.5.

The distance used to connect nearby buildings and areas to build the urban area had to be validated, too. For this, the effects of different distances to connect buildings were visualized in a region interspersed with a number of hamlets in an otherwise largely rural region. Specifically, the area between the Swiss villages of Gossau and Hinwil in the canton of Zurich was used for this analysis. This region was used, as the several hamlets in this area lead to a relatively large border zone between the rural areas and the urban areas, at the same time containing a relatively small number of buildings in a reasonably small area, which makes a visual assessment less challenging. The impact of varying distance thresholds is particularly evident in the extensive boundary region between the urban and non-urban area. This is because there are more opportunities for the cluster of buildings that constitute the main part of the respective hamlet to connect to isolated structures, such as standalone farmsteads, located at a greater distance. To detect the differences arising from distinct distance thresholds, an algorithm was necessary that shows which buildings could be connected to each other using the respective threshold distance. For this, in a first step, the centroid points of the buildings were used to represent the building footprints, and in a second step the polygon coordinates of the building footprints were used as the points utilized in this validation process. The visualization of the impact of different distance thresholds was achieved using the DBSCAN algorithm (Hahsler et al., 2019), setting the minimum number of neighboring points necessary for a point to be regarded as a core point to 1, which includes the respective point itself. To compare the results of differing threshold distances, the threshold distances were incrementally increased by 10 m, starting with 10 m and reaching a maximum of 150 m. The results of the threshold distances of 10 m, 60 m, and 120 m using the coordinates of the polygons of the building-footprints as points are shown in Figure 10. Employing both the building centroids and the polygon coordinates of the building points as input for the DBSCAN algorithm resulted in similar outcomes: A threshold between 40 m and 70 m seemed most sensible to delineate urban areas, as this allowed for the maximum contiguity of urban areas while avoiding the inclusion of too many green areas that should be better classified as non-urban. Based on this result, the value of 60 m found in literature (Bundesamt für Statistik, Sektion GEO, 2007) to connect the components of the urban area was chosen to serve as input for the urban area construction.

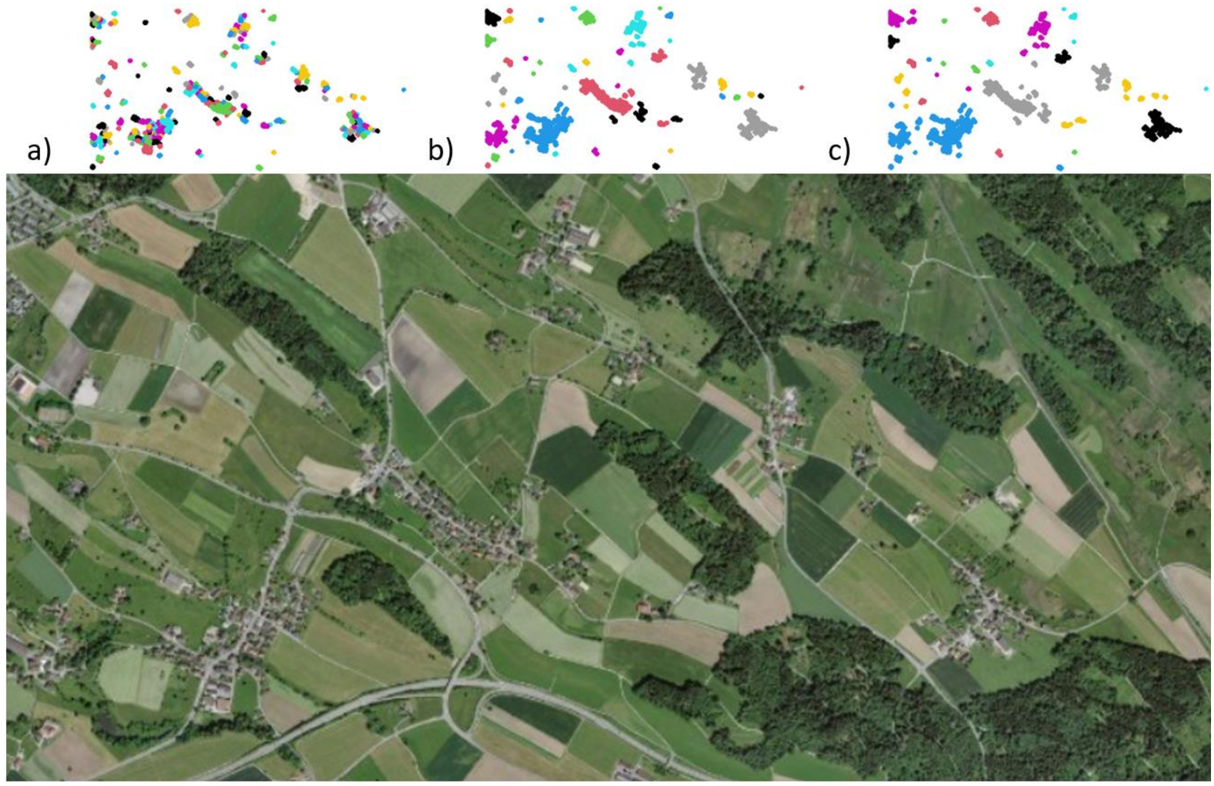



Figure 10: Image of the area used to validate the distance envisaged to connect houses and results of the DBSCAN algorithm on the coordinates of the building footprint polygons using different threshold distances of a) 10 m, b) 60 m, and c) 120 m. The different clusters created by points that can be connected with each other using a straight line maximally as long as the threshold distance are colored in different colors. Please note that due to the limited number of colors different clusters have sometimes the same color. Whereas a) would lead to many separate urban areas close to each other in settlements that are actually made up by further buildings, version c) would give rise to urban areas that incorporate a lot of green areas like agricultural fields. (Source of the remote sensing image: map.geo.admin.ch (Geoportal of the Confederation, 2023))

## 3.5 Constructing the Urban Areas

To construct the urban areas, an $\alpha$-shape algorithm implemented in R was used (Pateiro-López & Rodríguez-Casal, 2010). This approach provides a basis to automatically construct urban areas based on the available input data, without relying on time-consuming and costly manual delineation. The way this algorithm works in practice can be described as trying to fit a circle with the radius $\alpha$ between a pair of points of a point dataset. At the same time, this circle must not contain other points (Pateiro-López & Rodríguez-Casal, 2010). If it is possible for a circle to touch both points at the same time, a straight line is drawn between these two points (Pateiro-López & Rodríguez-Casal, 2010). A visualization of this concept can be found in Figure 11.

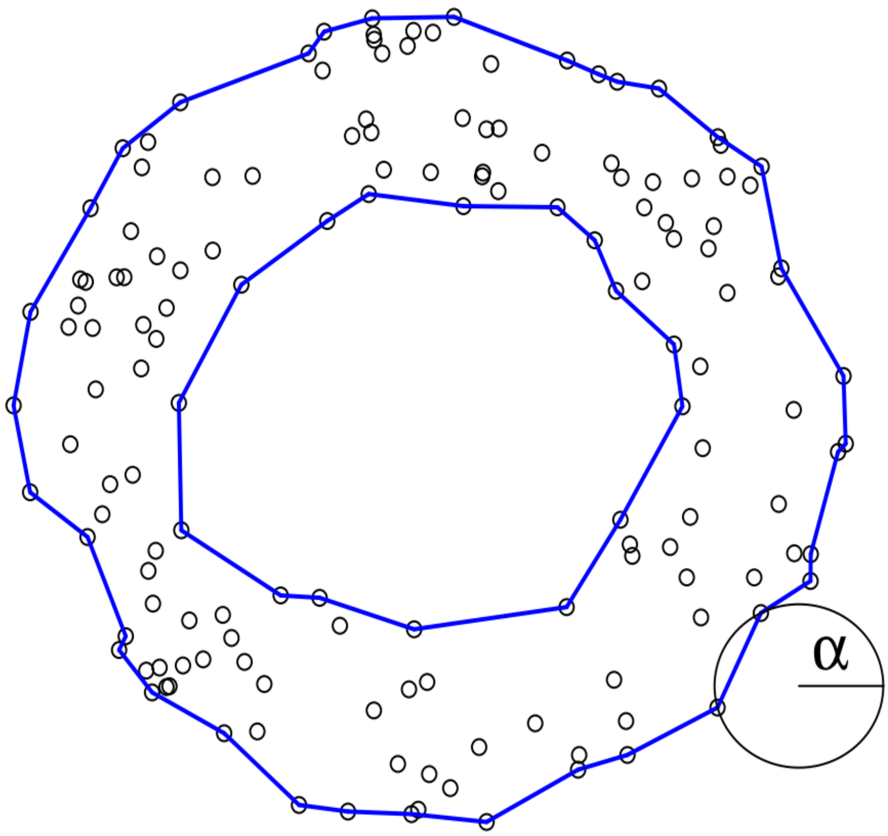



Figure 11: Approach used to construct an $\alpha$-shape, $\alpha$ being the chosen radius. (Image source: Pateiro-López and Rodríguez-Casal (2010), slightly modified)

Instead of using the $\alpha$-shape algorithm to delineate urban areas, it is also conceivable to use algorithms constructing a concave or convex hull. However, algorithms creating a concave hull lack the ability to detect holes, and algorithms creating a convex hull have the same issue, with the additional problem of not being able to exclude concave areas (Barber et al., 2022; Gombin et al., 2022). For this thesis, the $\alpha$-shape algorithm by Pateiro-López and Rodríguez-

Casal (2010) was utilized, specifically written for implementation in the R programming language. Unfortunately, the output of this algorithm is not a polygon or any other geometry that could be easily converted into one. How this problem was tackled, will be explained later in this section. An algorithm that produces polygons is the one by Bellock (2021) called "alphashape" written for the use in Python. Unfortunately, upon testing, it was discovered that it lacked the ability to cut out holes.

As already mentioned, the $\alpha$-shape algorithm needs point-data as input. However, as previously described, the data used for this thesis mostly consist, at least originally, of polygons (or lines in the case of railway tracks). The buffer of 7 m around the footprints used to recreate gardens did not change the fact that the data were present in polygon format. Applying the $\alpha$-shape algorithm directly on the coordinate points of the polygons works for many buffered buildings and areas, but issues can arise. The areas of the TLM3D and the buffered buildings are comprised of a large number of coordinates, which are used to describe their polygonal shapes. This leads to a long calculation time for the algorithm and the subsequent steps. To tackle this problem, the polygon geometries were simplified in this thesis, reducing the number of coordinates describing the areas and buffered building footprints. For this, the function "st_simplify" of the library "sf" in R was utilized (Pebesma, 2018). The simplification of geometry was performed with the constraint that it must not alter the border of the polygon by more than 1 m. This 1 m threshold was selected to effectively reduce the number of points for further calculations while minimizing any significant changes to the geometry. However, this simplification brought about another issue: Namely, polygons of large areas and buffered buildings with simple geometries are then represented by a relatively small number of coordinate points. Specifically, it results in the loss of points along the long edges of the buffered building footprints or areas. This can lead to issues when applying the $\alpha$-shape algorithm directly to the simplified coordinate points of the buffered footprint or area polygons: The distances between the points may now be too large to connect the points, although the points should be certainly connected, as they are part of the same area or buffered building. To address this issue, newly created points were introduced to interrupt large edge sections without any points. Adopting this approach, the maximum length of edges in the buffered building footprints and areas was set to 11.5 m. This length is equivalent to half the radius used for the $\alpha$-shape algorithm in this thesis. The reason behind the length of the radius will be explained later in this section. The value of half the length of the radius used for the $\alpha$-shape algorithm was used after a visual assessment of the possible effects of the edge length. It was discovered that, in addition to the issue of edge lengths larger than double the $\alpha$-shape algorithm's radius causing disconnected points, edge lengths not much smaller than double the radius can result in problems, too. This is the case, as when the edge is sufficiently long, i.e., the distance between the points is sufficiently large, the circle used in the $\alpha$-shape algorithm may touch points lying beyond the straight line connecting the two points on the edge. These points can be lying on the opposite side of the area or buffered building. This concept can be seen in Figure 12. By setting the distance between points on the edge of the area/buffered building footprint to half the radius used in the $\alpha$-shape algorithm, the extent to which the circle can extend beyond the points on the same side of the area/buffered building footprint is greatly reduced. While reducing the distance between the points generally has positive effects on the delineation accuracy, distances that are too small can lead to an increase in the number of points that must be considered during subsequent calculations, causing them to slow down.

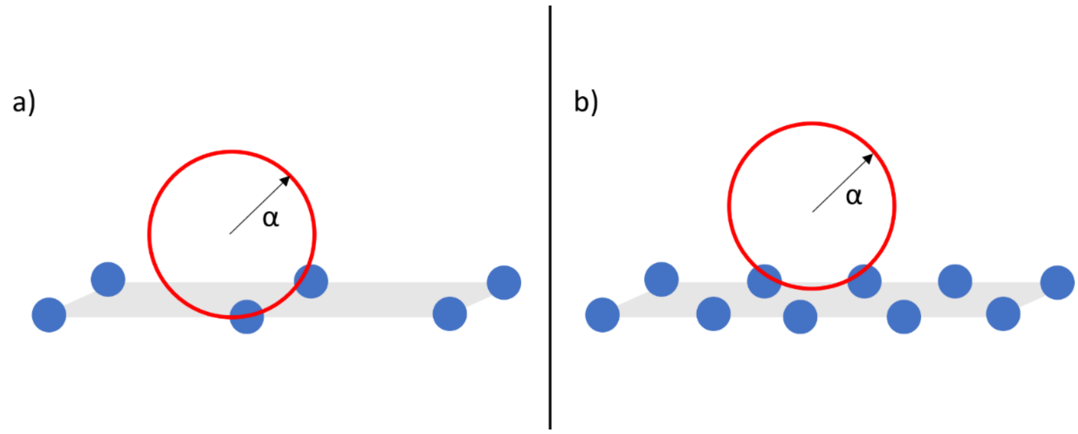



Figure 12: Finding the optimal distance between points: a) shows an example of a case where the distance between the points is too large. The polygon representing the building footprint/area is in grey, its simplified coordinate points of it in blue, and the circle used in the $\alpha$-shape algorithm is shown in red. Although the distance between the points on the edges of the building is smaller than the double radius ($\alpha$) used in the $\alpha$-shape algorithm, the connection between the points will be false, as the circle can touch a point beyond the straight line that can be drawn between the points on the same side of the building/area. b) shows a scenario where the mentioned issue is not present, due to a smaller distance between the vertices that are used to represent the polygon.

After solving the issues related to the edges of the building footprints and areas, it was observed that large buildings and areas posed an additional problem: Since the polygon points were only placed along their edge, the geometries of large areas or buffered buildings could allow the $\alpha$-shape algorithm circles to fit inside them. This would again lead to errors. This problem was solved by creating a point grid inside the areas and buffered buildings. The chosen distance between these grid-points was 11.5 m (half the radius used for the $\alpha$-shape algorithm), which is sufficient to prevent the accommodation of circles during the running of $\alpha$-shape algorithm inside the geometries of the areas and buffered buildings. The mentioned issue and its solution are visualized in Figure 13.

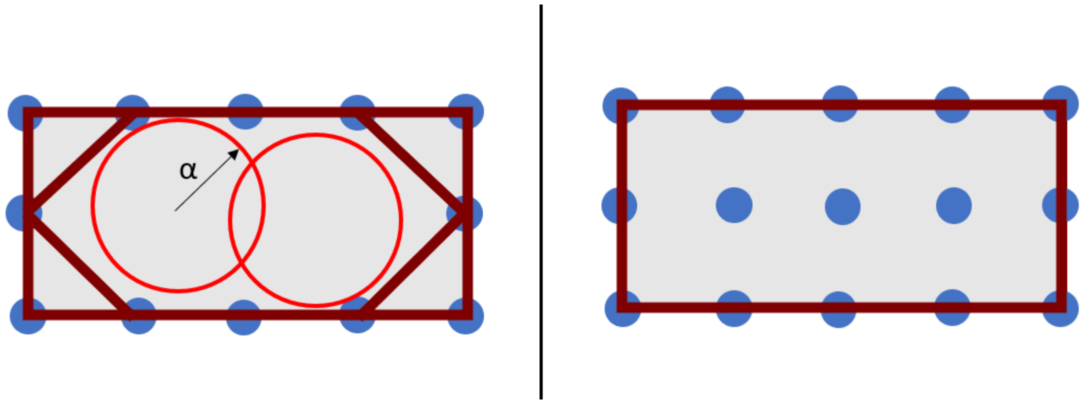



Figure 13: Creating a point grid inside the areas and buffered buildings: On the left, the $\alpha$-shape algorithm applied to a large buffered building or area without points inside it. The connections that result from the $\alpha$-shape algorithm are in dark red. Note the wrong connections that ensue on the left. On the right, the connections take place as planned, due to the newly introduced points inside the area.

Before applying the $\alpha$-shape algorithm, it was necessary to delete closely spaced points, as otherwise the connections between them could lead to strange forms that impede the creation of polygons in the later stages. To address this problem, the coordinates of the points were rounded to three decimal places (i.e., at the millimeter scale) and duplicates were removed. Like this, a manipulated set of polygons was created, which was different for the distinct urban area delineations of this thesis. In contrast, the building footprints that were used for the $\alpha$-shape algorithm were the same in every version. The exclusion of certain buildings in the version "Non-rural area" is an effect of data manipulation after the $\alpha$-shape algorithm. The delineation of the "Area with strong anthropogenic influence" and the "Non-rural area" share the same input data as that used for the delineation of the "Built-up area", but with additional data included. The kinds of areas used for the calculations for the years 2015 and 2021 are the same. However, as mentioned in Section 3.2, between 2015 and 2021 some new classes for areas were created. To prevent any distortion of the results due to a larger number of classified entities, newly delineated entities in 2021 that were not merely formed by splitting up an already existing class were excluded from the calculation.

As already stated, the entities utilized for the delineation of the urban areas were different between the distinct versions of urban area delineation used in this thesis. Apart from the buffered building footprints, the urban area delineation using the definition of "Built-up area" used other, mostly unvegetated areas with strong visible anthropogenic influence. Specifically, all recreational areas (e. g. sports grounds) of the TLM3D were used, except the areas of golf courses and zoos, due to the fact that they are often vegetated. Roads and railway tracks were excluded, the latter being incorporated in a later step. Other areas used for transportation (like airports) were included straightaway. Further areas with limited vegetation were also included into the urban area, like mines, quarries, landfills, allotment gardens, and cemeteries. The mentioned areas were also utilized to demarcate the urban areas in the versions "Area with strong anthropogenic influence" and "Non-rural area", which in addition incorporated golf courses and zoos.

Having increased the space taken up by buildings using the 7 m buffer, it was decided that the 60 m threshold to connect the different components of the urban area had to be adapted. The new threshold was set to a distance of 46 m, accounting for the 14 m lost by the 7 m gardens of two buildings. As this 46 m would be the diameter of the circle, the equivalent radius measures 23 m. Thus, the radius used for the $\alpha$-shape algorithm was 23 m.

The result of the $\alpha$-shape algorithm was, like already mentioned, a set of points that may or may not be connected to each other, depending on their distance to each other. The information about the connections between the points was then used to create polygons. In general, the self-made algorithm that creates polygons from the result of the $\alpha$-shape algorithm works in the following manner: The process of creating a polygon involved selecting a starting point, determining its connections to other points, saving its coordinates, and repeating this process for each interconnected point until the algorithm returned to the first point. Finally, the collected coordinates can be utilized to construct a polygon. To avoid looping back and forth between a set of points, the algorithm excluded points that had already had their neighborhoods assessed when selecting the next point to check. The way this algorithm traverses the geometry is visualized in Figure 14.

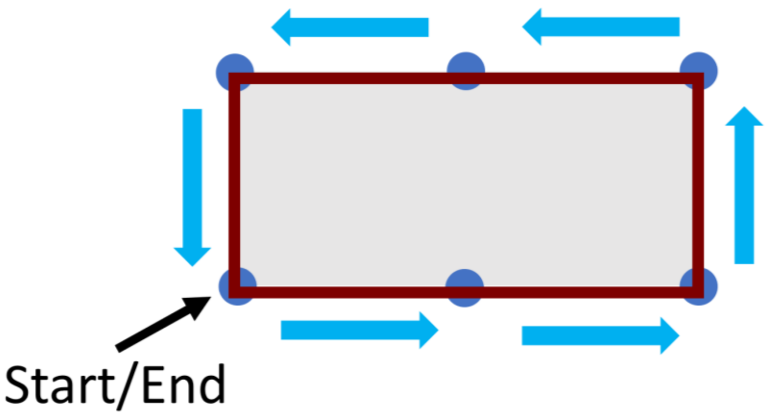



Figure 14: Illustration of the presented algorithm progressing through the connected nodes of the $\alpha$-shape algorithm (indicated by the blue arrows) in preparation for the creation of the polygon. In this case, the algorithm starts and ends at the point on the lower left.

Please note that points without any connections exist. The reason for this are never points that are too far away from other points to establish a connection, as such scenarios were made impossible in the preparation of the data. Rather, it is due to the fact that the circles used for the $\alpha$-shape algorithm do not permit the connection of points. For reference, see the points lacking any connection to any other point in Figure 11. The presented principle would work very well in simple settings, as long as the urban area is only made up by one single polygon. However, there is another issue apart from the fact that many urban areas are disconnected and thus comprise multiple polygons: For now, the algorithm can only handle situations where points are connected to exactly two other points (or none at all). However, when using the output from the $\alpha$-shape algorithm, even in the simplest real-life scenarios, there was at least one instance where this rule was broken, although here too, the points generally are only connected to two other points. To address the issue posed by disconnected urban areas in the $\alpha$-shape algorithm's output, a rule was implemented where the algorithm would restart from a randomly selected point in the group of unchecked points after completing a run on the edges of one polygon.

The problem posed by points connected to more than two other points is a more challenging one to address. In the process of this thesis, the most common problems stemming from this issue were eradicated, but there can still be configurations of points that lead to issues and need human intervention. One scenario that occasionally can be observed, is that points serve as the sole connection between two otherwise isolated urban areas. These points have four connections instead of two. An illustration of this issue can be seen in Figure 15.

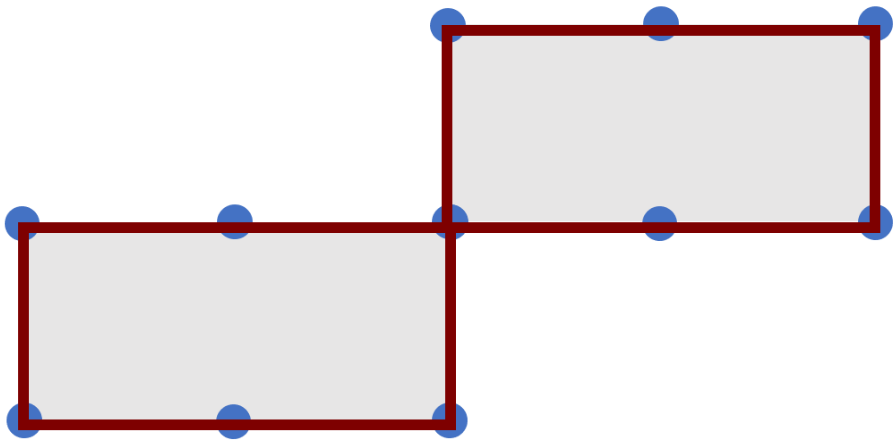
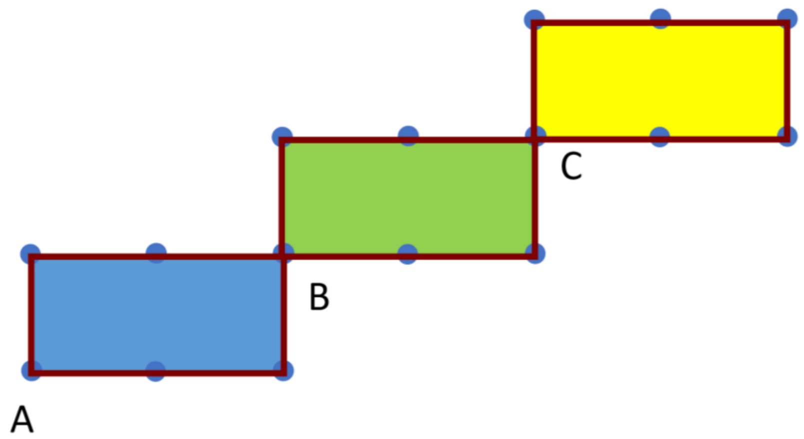


Figure 15: Illustration of a challenge when delineating urban areas using the $\alpha$-shape algorithm: One of the points is connected (dark red) with four other points.

The challenge posed by points with four connections is that the point with the four connections may have to be passed twice in order to finish the delineation of the shape correctly. This contradicts the earlier established rule of not passing a point more than once. To resolve this conflict, it was decided that points can be passed twice, as long as they are not the point just previously analyzed. It has to be mentioned that points are only allowed to be passed multiple times, if they are connected to points that were not passed before. Therefore, points with only two connections are not traversed multiple times. However, the fact that a point is represented more than once inhibits a direct construction of a polygon, as the algorithm used to finalize the polygon construction does not accept this. Therefore, the sub-areas are split into individual polygons before polygon creation. The solution to address points with four connections was to cut out all the points in the list that were situated between points that appeared twice in the list. For this, it is important to only use the first point and its duplicate points between the first pair of duplicated points and not all duplicated points. Otherwise, when presented with a form with a series of points that have four connections, false results occur. The mentioned procedure then is run multiple times, until the whole series of areas connected by points with four connections is resolved. An illustration of this can be seen in Figure 16. Please note that, unlike what this simplified illustration suggests, the points with four edge connections (i.e., point "B" and "C") will not be assigned to only one of the polygons it is a part of, but both. Therefore, the polygons are more correctly characterized as "A … B …. A", "B … C … B", and "C … C", respectively.




| Point |
| --- |
| A |
| … |
| B |
| … |
| C |
| … |
| C |
| … |
| B |
| … |
| A |

Figure 16: Approach used to divide shapes with a series of points having four connections. In this case, the algorithm employed to split the form starts at point A. Upon reaching point B for the first time, the algorithm snaps the list of points and then again at the second instance of point B. Therefore, the parts of the table and the graphic colored in blue are separated from the ones in green and yellow. In a second step, the same principle is applied for the green and yellow parts of the list and the graphic– i.e., the parts of the list containing the points B and C – separating the green parts of the list and graphic from the yellow ones. Consequently, three individual polygons are created.

Another issue that needed resolution were cases where points had connections to three other points. The simplest of such scenarios can be seen in Figure 17. This case could be solved by simply removing the link lying between two points that have three connections.

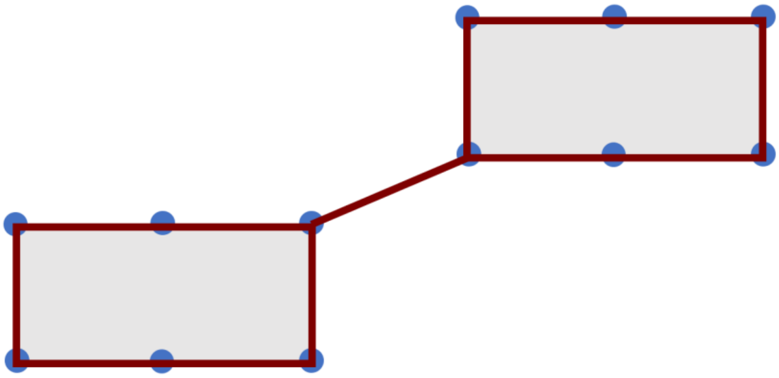


Figure 17: Simple case of two parts of urban area being linked by a single connection.

More complex challenges are depicted in Figure 18. A version of this problem could be a scenario where the shaded area is delimited by four edges connecting four points on the corners of the shaded rectangle (situation "a)"). In this case, deleting all edges between points with three connections in one step results in an arrangement of edges that would fail to provide the structure to construct a polygon. Therefore, edges connecting points with three connections need to be deleted one by one. Consequently, as a result of removing the first edge connecting points with three connections, the two remaining edges connected to the points that now lost a connection and other points with three connections are now only connected to two other points, thus avoiding further deletion. The edge at the opposite side of the shaded area, however, still gets deleted. This leads to a configuration of edges that makes the construction of a polygon possible. In the case where there is an additional point between the mentioned points (situation "b)"), a similar approach is possible, but instead of deleting the edges connecting points with three connections one has to delete the edges connecting points with three connections with points that are connected to other points that have three connections. Situation "c)" depicts the same concept, but with even more points lying between the points with three connections. It is also possible to have hybrid scenarios between the shown situations, for example if scenario "a)" would contain an additional point separating its upper edge. All mentioned scenarios are very seldom, and the frequency at which these scenarios occur gets increasingly smaller the more points are involved. However, even if one of these scenarios only happens once over the whole urban area, this impedes the possibility to construct a polygon. Thus, the approaches to fix the geometries of versions "a)" and "b)" explained before were implemented in the code. Thus, a geometry with more than one point lying between points with three connections would still lead to issues in the algorithm, as no fix for such scenarios was coded. Therefore, issues arising from such relationships between points have to be repaired manually by deleting the edges causing the problem. Implementing an approach similar to the one of situations "a)" or "b)" in such settings could lead to new problems. For example, the edges connecting sets of problematic triple points could be deleted, as long as there are multiple occurrences of this type of problem in close proximity.

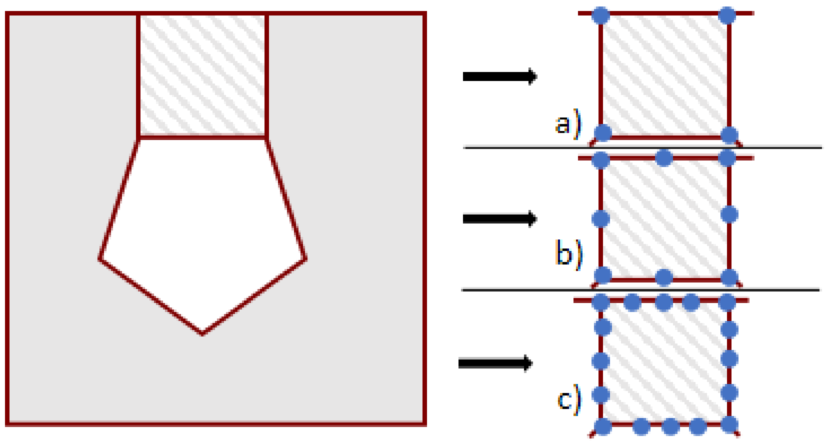



Figure 18: Tricky situations requiring resolution. The red lines depict the lines created by the $\alpha$-shape algorithm. They grey area depicts the area that finally was classified as urban. The shaded area is the area that was included into the urban area by chance. On the right are the three scenarios discussed in the text.

With the implementation of the aforementioned modifications, it was finally possible to construct polygons. These polygons were nevertheless subject to some modifications. In the "Built-up area" delineation for both years, areas characterized by a large number of parallel railway lines were included. Such areas were included into the urban area only if they actually could be described as lying inside of the urban area. This approach is based on the viewpoint often represented in literature that roads and railway tracks inside the urban area are a part of the urban area itself. The following approach was employed in order to address this matter: To only use the railway tracks inside the urban area, a dataset was constructed consisting solely of the buffered railway tracks (created using the aforementioned 4.5 m buffer) located within holes of the urban area. Then, the polygons in this dataset describing the railway tracks and their immediate surroundings in the holes of the urban area were enlarged by an additional buffer of 23 m, the size of the radius used for the $\alpha$-shape algorithm. Using this buffer, nearby railway lines should connect and create a new, connected polygon. This area was then shrunk by 46 m, the double radius used for the $\alpha$-shape algorithm, in order to only include large areas that are characterized by parallel railways. Then, another buffer was used to enlarge the areas with the multiple parallel railways: this time with a size of 69 m, the triple size of the radius used in the $\alpha$-shape algorithm. This buffer was intended to reverse the shrinkage of the net 23 m applied to the area in all the remaining areas, and to apply an additional 46 m buffer to connect the area of the multiple railway lines with the surrounding urban area. As for now, the railway area may abut green urban area, meaning the railway area would be 46 m too large at those places, as no connection with the surrounding urban area took place. To correct for this, the newly delineated urban area (i.e., including the area with parallel railway lines) is shrunk by 46 m, and this area is then merged with the original urban area (i.e., the area before the adding of the parallel railway lines).

Further processing was conducted on the urban areas categorized as "Area with strong anthropogenic influence" and "Non-rural area". In both cases, all holes in the urban area were filled and classified as urban area. Public parks, urban greenery, and areas like forests should therefore be included into the urban area, if they were surrounded by it. For the creation of the "Non-rural area", additionally, urban area bits that were smaller than a certain threshold were excluded from the urban area. This threshold was set to be the equivalent of four residential houses and their surroundings. This value was based on the process used for the creation of the NRI. The area of four residential houses and their surroundings were determined to be 2000 $m^2$, based on the assumption that each residential house covers an area of 500 $m^2$. This same assumption was already used in calculating the buffer size applied to the building footprints. Throughout this thesis, areas that meet the requirements to be classified as "Non-rural area" except for the 2000 $m^2$ size threshold will be referred to as *SIUAFs* (small isolated urban area fragments). The presence or absence of this areas constitutes the differences between the delineations of "Area with strong anthropogenic influence" and "Non-rural area". The general differences between the delineation of "Built-up area", "Area with strong anthropogenic influence", and "Non-rural area" areas can be seen in Figure 19. This simplified image illustrates the general differences between the three mentioned classes, ignoring the differences in area due to different input data (e.g., golf courses are not part of all delineations).

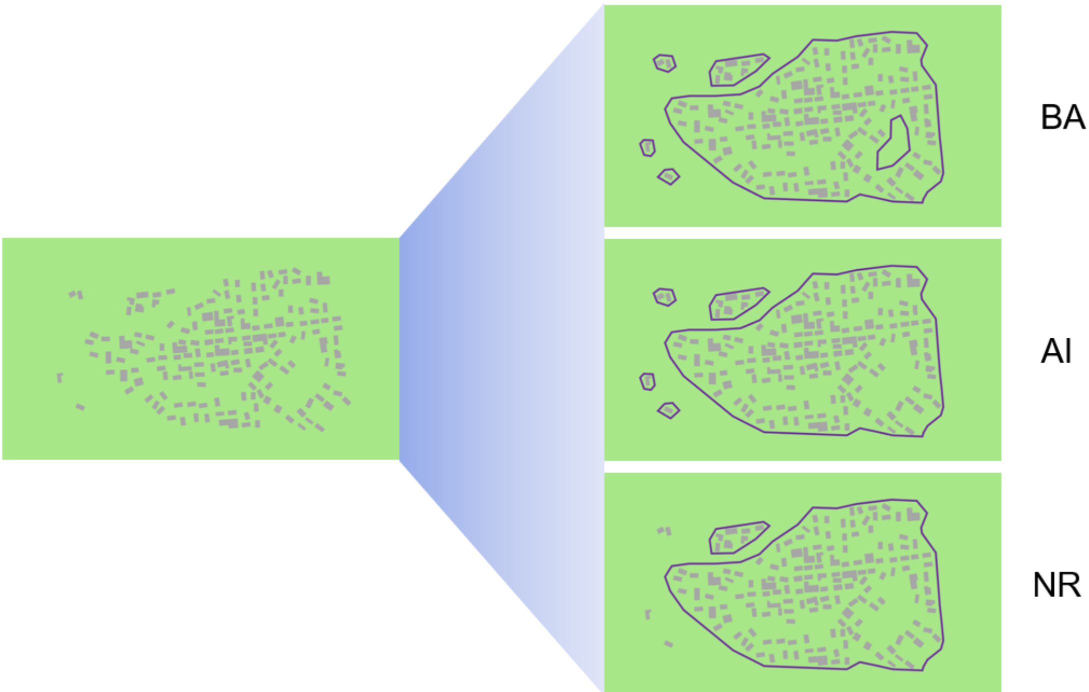



Figure 19: Basic illustration of the three urban area definitions as applied in this thesis (outlined), based on a fictional village. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

## 3.6 Preparation of the Data for Urban Sprawl Measurements

To limit edge effects stemming from relying solely on municipal data instead of data encompassing the entire country, the aforementioned techniques were applied not only to the municipal data, but also to all buildings and areas that intersected with the territory within 100 m of the municipality boundaries. After the creation of the urban area polygons, this had to be adjusted, only considering the territory within the municipality boundaries. In the case of the ensuing density calculation, these newly created polygons could be directly used. However, for the calculations of WUP and Moran's I, certain adjustments were still necessary. Both approaches need sub-areas. In the case of WUP, the urban area was rasterized by Yves Maurer, using a cell size of 15 m (Y. Maurer, personal communication, December 7, 2022), the recommended threshold (Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Schwarzak et al., 2014). For the Moran's I calculation, the same threshold was used to create grid cells representing the sub-areas. In order to perform the calculation of Moran's I, this grid needed to contain values pertaining to the variable of interest in the area of the municipality. To exclude uninhabitable areas, it was determined that they should be removed from the municipality's territory. As a result, the effect of uninhabitable areas, like lakes, on the urban sprawl calculations should be mitigated. This uninhabitable area was automatically corrected if it was overlapping with the urban areas delineated for this thesis. Thus, the data used for the Moran's I calculation consisted of a set of square subareas containing values lying in the inhabitable part of the respective municipalities. The threshold of 15 m was selected for the calculations of WUP and Moran's I due to various reasons:

1) Cells with the size of 15 m are recommended in the literature for the WUP measurement (Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Schwarzak et al., 2014).
2) It appeared reasonable to assume that this threshold is also valid for the Moran's I calculation.
3) Upon theoretical considerations and evaluating the impact of different thresholds, it was determined that a cell size of 15 m strikes a good balance between the number of cells and the level of accuracy. In urban sprawl literature, usually higher cell sizes were used. For example, Zhou et al. (2019) used cells with an edge length of 1 km, and Salvati and Carlucci (2014) used different edge lengths ranging from 2 km to 50 km. However, larger thresholds than 15 m increasingly led to the loss of finer details in the urban area delineation and the disappearance of SIUAFs. This effect can be seen in Figure 20. This can be especially problematic when this results in the failure to detect key features of urban sprawl, such as ribbon development or leapfrogging. Furthermore, larger thresholds would lead to the erosion of differences between the versions depicting

urban areas of "Area with strong anthropogenic influence" and "Non-rural area". In addition, holes would tend to disappear, eroding the differences between the version "Built-up area" and the other two versions.

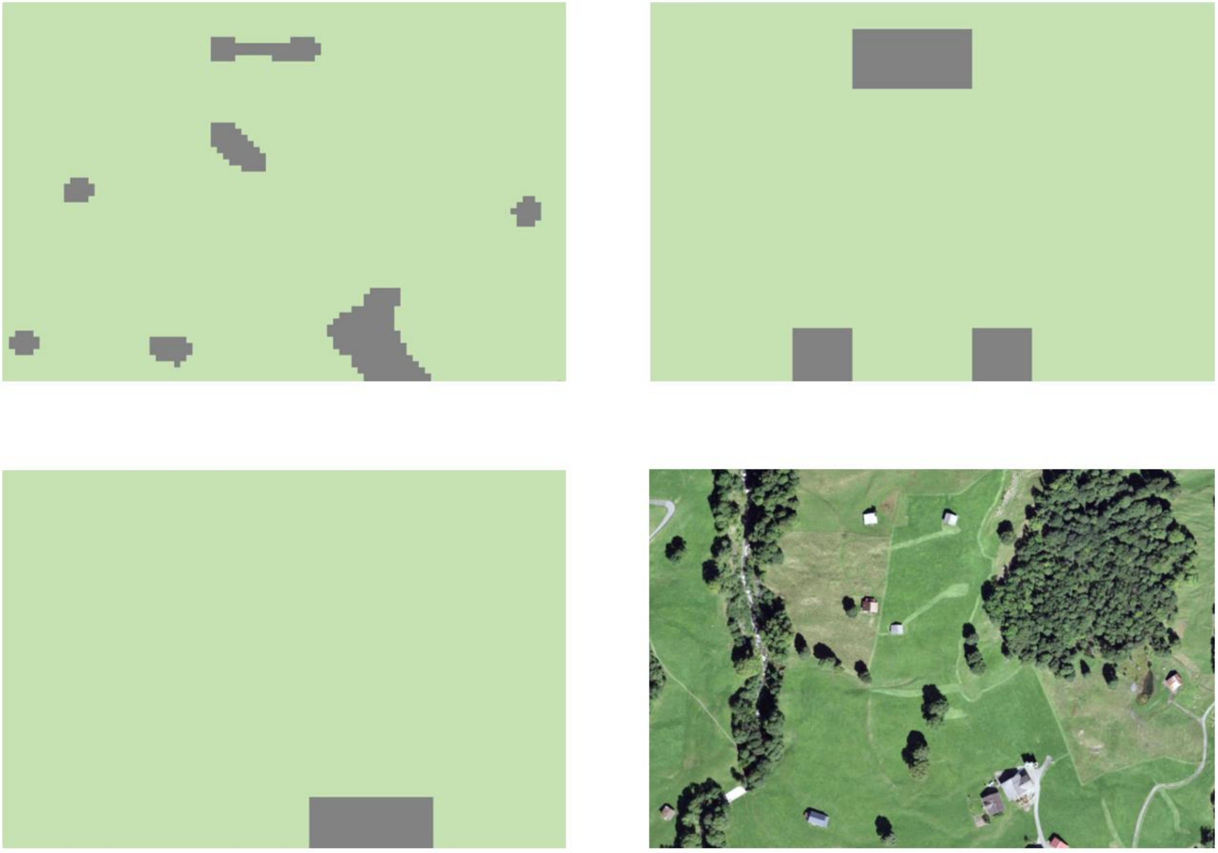


Figure 20: Example of the effect of different cells sizes used for the sub-areas on the version "Built-up area". On the top left, the used threshold is 5 m, on the top right it is 50 m and on the lower left it is 100 m. On the lower right is an image of the area. (Image source of the lower right image: Environmental Systems Research Institute (2022))

## 3.7 Calculating Urban Sprawl

The measures of density, WUP, and Moran's I were calculated for all three delineation versions for the years 2015 and 2021. In the following, the approaches to calculate these measurements will be discussed.

### 3.7.1 Density

Density was measured using the urban areas delineated for this thesis directly and the population in the municipalities in the respective years. This way of calculating density is well represented in urban sprawl literature, as was explained in Chapter 2 of this thesis. The density calculation was done in R.

### 3.7.2 WUP

WUP was calculated by Yves Maurer using the QGIS-plugin by Schwab and Horiguchi (2020). This plugin uses another formula for the DIS-component than the one presented in Chapter 2 of this thesis. However, it still serves as a good approximation for the originally presented formula (Schwab & Horiguchi, 2020; Schwick et al., 2018). This formula is as follows:

$$DIS = \frac{1}{n}\Sigma_{i=1}^{n}\frac{1}{n_i}\left(\sum_{j=1}^{n_i}\left(\sqrt{\frac{2\cdot d_{i_j}}{1\ \mathrm{m}}+1}-1\right)\frac{\mathrm{UPU}}{\mathrm{m}^2}+\sqrt{0.97428\cdot\frac{b}{1\ \mathrm{m}}+1.046}-0.99629\right)\frac{\mathrm{UPU}}{\mathrm{m}^2}$$

where $n$ describes the number of cells, $n_i$ the number of cells of the urban area inside the horizon of perception around the cell $i$, $d_{i_j}$ being the distance between the cells of urban area $i$ and $j$, and $b$ describing the edge length of the single cells, where $\mathrm{UPU}$ and m (meters) represent the units or parts of the units of the respective values. Apart from the DIS formula, the other calculations regarding the WUP were the same as described in Chapter 2 of this thesis. The edge length of the cells was set at 15 m and the horizon of perception at 2 km, as proposed in literature (Jaeger, Bertiller, Schwick, Cavens, et al., 2010; Schwarzak et al., 2014).

In addition to determining WUP and its components as mentioned, yet another version of WUP was also calculated, which takes into consideration the uninhabitable area. This measure, which will here be called *Adjusted WUP*, is calculated multiplying WUP with the product of the division of the total area of the municipality by the inhabitable area (Schwick et al., 2018).

As a formula, the calculation of Adjusted WUP looks like this:

$$Adjusted\ WUP = WUP \cdot \frac{A_{reporting\ unit}}{A_{inhabitable}}$$

where $A_{reporting\ unit}$ stands for the whole study area, in this case the municipality, and $A_{inhabitable}$ for the inhabitable part of it.

The version of the WUP that does not adjust for the uninhabitable area will be referred to as *Unadjusted WUP*.

The calculation of UD and, therefore, of WUP, too, was performed using the data on inhabitants and employment of the respective areas, following the approach proposed by Schwick et al. (2018). This calculation was performed using data on the inhabitants and employment within the urban area defined for the thesis.

### 3.7.3 Moran's I

The calculation of Moran's I was done using the set of cells previously described. Each cell was assigned one of two values: 0 (non-urban area) or 1 (urban area). The decision to employ a binary classification was made primarily for a single reason: Before rasterizing the map, the values were also binary, as the areas were either classified as urban, or they were not. Assigning values on a continuous scale would consider the fact that cells at the edge of the urban area were typically not fully covered by urban area polygons. However, it would artificially lead to a less abrupt transition between the value assigned to non-urban areas (0) and the one assigned to urban areas (1). This would result in a smoother transition of values on the border between urban and non-urban areas, leading to a higher value of Moran's I, indicating more homogeneity. This effect was already criticized by Steurer and Bayr (2020).

There exist various ways to calculate Moran's I. For this thesis, inverse distance was used to weight the spatial relationship between the cells, for the reasons explained in Chapter 2 of this thesis. The Moore neighborhood was used. The calculation of Moran's I was done with ArcGIS Pro.

# Chapter 4 | Results

The final results of this thesis are the differences on sprawl measurements based on the utilization of the three different urban area delineations mentioned previously. However, as elaborated in Figure 7 in Section 3.1, the different urban area polygons had to be created as an intermediate result. In this section, the intermediate and final results will be presented. The presentation of these results will be chronological, meaning the final results are at the end of this section.

## 4.1 Urban Area Delineations

Using the previously described approaches, a total of 30 different urban areas were delineated. This, as there is an urban area delineation for every one of the five municipalities, using three different urban area definitions, at two different points in time. The maps based on the three definitions are visually different, as will become apparent from the subsequent images. In the following, the delineated urban areas based on the three definitions will be discussed in further detail.

### 4.1.1 Built-Up Area

The delineation of urban areas based on the notion that they should fundamentally be delineated as built-up areas, led to the creation of visually very distinct urban forms. Some of their differences can be directly seen when comparing the urban areas of the different municipalities in Figure 21. Looking at the results of using the definition of “Built-up area” to delineate the urban area, several observations can be made: Celerina, a quite compact village, has a delineated urban area which can be described as a large blob with a small number of holes. Only few zones marked as urban area lay outside the largest polygon, which delineates the main settlement and contains most of the buildings in this municipality. In contrast, the urban area exhibits a very complex form in Grindelwald. The settlement seems to have no clear center and contains holes of massive dimensions. Many buildings are not contained in the largest polygon, but rather in smaller ones, which contain few or only one house at a time. Holes can be found in other municipalities, like Wetzikon or Winterthur, too. In these cases, certain holes are also rather large. Still, Grindelwald is the uncontested winner in terms of complexity of holes and outlying urban area bits.

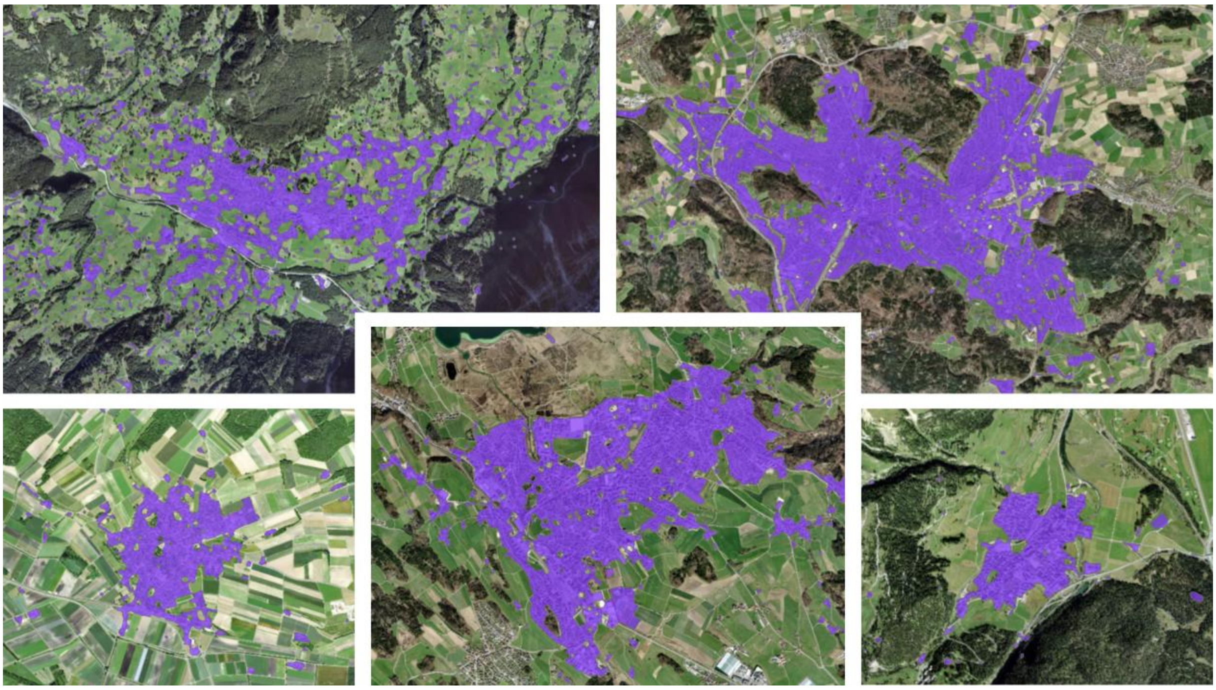


Figure 21: Delineations of the “Built-up area”: Shown are the urban areas (violet) of 2021 of the five analyzed municipalities – Grindelwald (upper left), Winterthur (upper right), Ins (lower left), Wetzikon (in the middle), and Celerina (lower right) – defining urban area as the “Built-up area”. Please note that the images are not of equal size. The urban areas are only delineated for the areas inside the municipality. For better visualization, some images may not cover the whole area of the municipality. (Source of the background image: Environmental Systems Research Institute (2022))

Areas that do not meet the criteria of “Built-up area” do not have to be very large to be excluded from the urban area, as can be seen by the size of the holes in Figure 22. Using this delineation approach, single urban-area polygons can be made up by only one building and its assigned garden. In the case of Grindelwald this is especially common and can be the norm in certain areas of the municipality. An area dominated by such single buildings can also be seen in Figure 22. In this figure it is also possible to view the delineation of the “Built-up area” at a rather large scale.

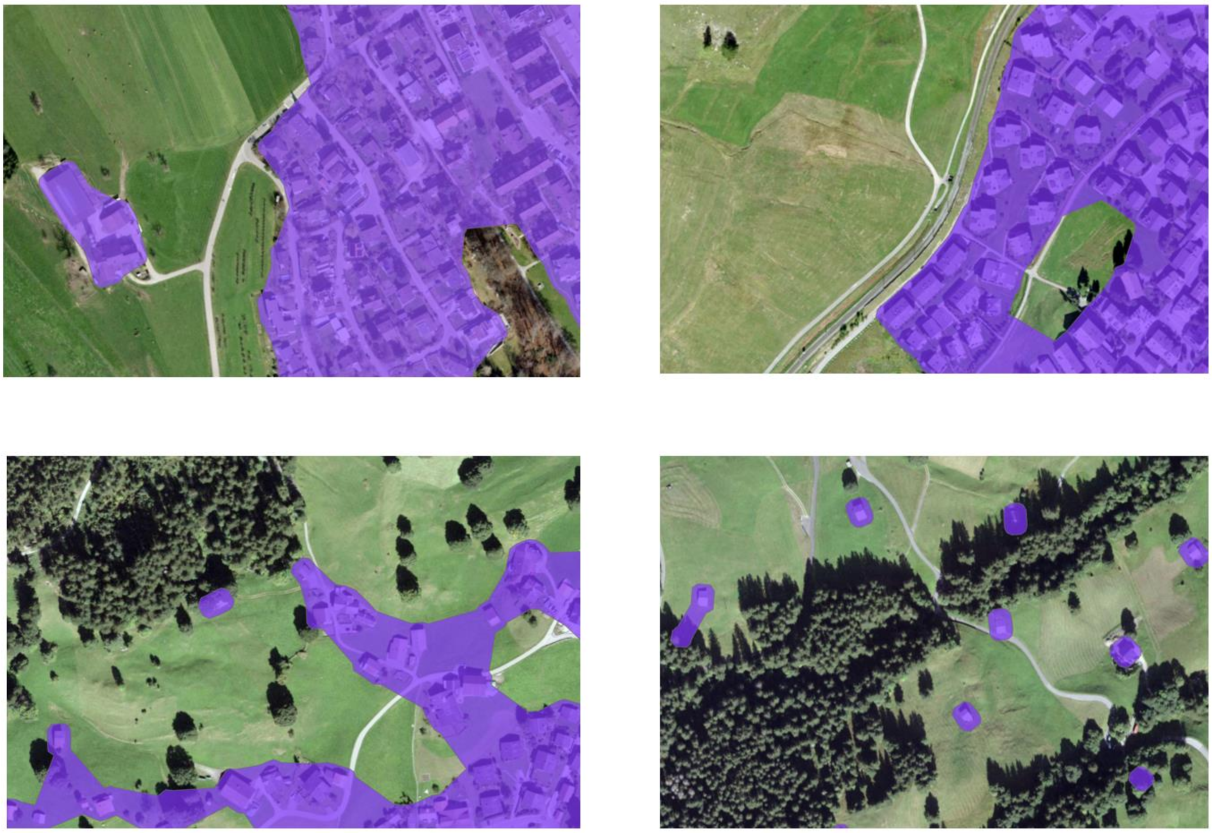


Figure 22: Urban area delineations of 2021 on the border to rural and natural areas in different municipalities. On the upper left, an excerpt of the urban area delineation of Wetzikon can be seen, on the upper right of Celerina, and at the bottom are two excerpts from the delineation of Grindelwald. (Source of the background image: Environmental Systems Research Institute (2022))

### 4.1.2 Area With Strong Anthropogenic Influence

The delineation of urban areas based on the concept of "Area with strong anthropogenic influence" is visually similar to the one using the "Built-up area", but with a striking difference: the lack of holes. This can be seen in Figure 23. In settlements with few holes or mostly small ones, this effect is less obvious than in municipalities like Grindelwald, where the whole intricate interplay between urban and non-urban area accompanying the largest polygon is lost to a large degree. Implementing the definition of "Area with strong anthropogenic influence", the urban area of Grindelwald now contains a great amount of greenery at certain locations. However, the urban area of Grindelwald is still visually distinguishable from the one of the other municipalities due to the large number of SIUAFs outside the main core of the settlement. The incorporation of large areas with greenery can also be prominently seen in the case of Winterthur. In this case, a forest is now part of the urban area, which previously was classified as non-urban.

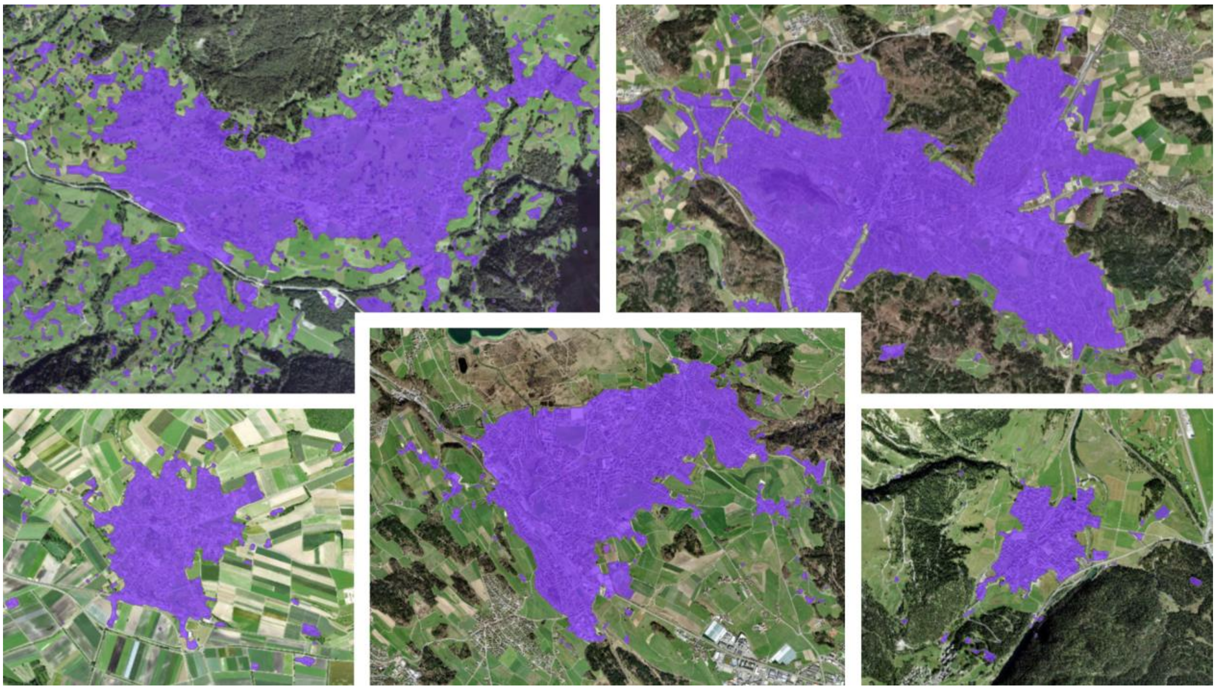


Figure 23: Delineations of the "Area with strong anthropogenic influence": Shown are the urban areas (violet) of 2021 of the five analyzed municipalities – Grindelwald (upper left), Winterthur (upper right), Ins, (lower left), Wetzikon (in the middle), and Celerina (lower right) – defining urban area as the "Area with strong anthropogenic influence". (Source of the background image: Environmental Systems Research Institute (2022))

In addition to the areas gained from greenery in the holes of the urban area, they were also enlarged by the incorporation of other largely vegetated areas: golf courses and zoos. Such areas can notably increase the urban area, as can be seen in Figure 24.

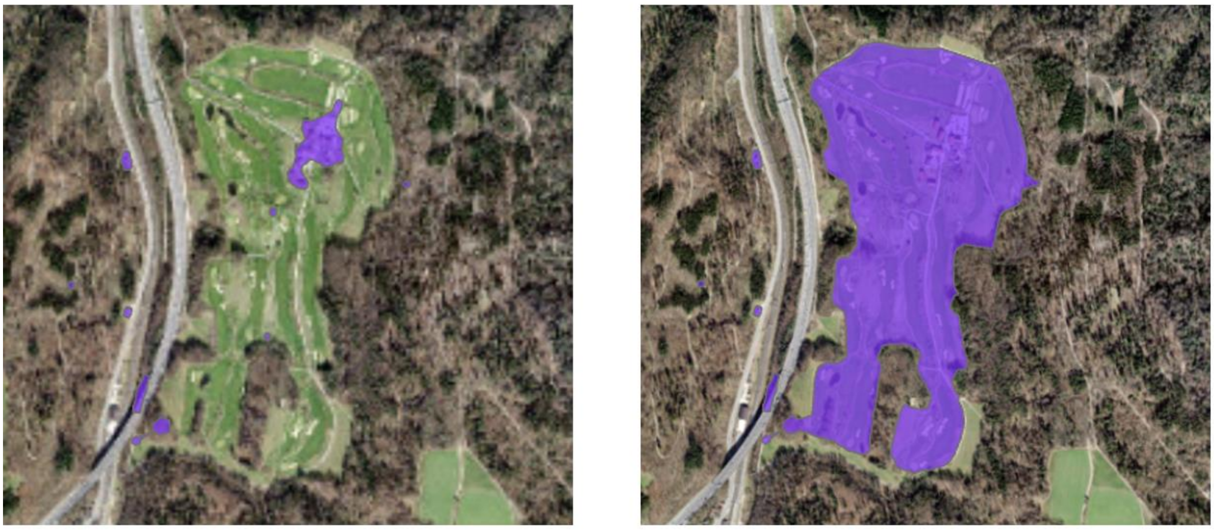


Figure 24: Golf course in the south of Winterthur. The result of using the definition of "Built-up area" to delineate the urban area for the year 2021 is illustrated on the left. On the right, the data of the whole golf course is used, increasing its area considerably. (Source of the background image: Environmental Systems Research Institute (2022))

### 4.1.3 Non-Rural Area

The delineation of the "Non-rural area" is identical to the "Area with strong anthropogenic influence", except that it excludes SIUAFs. The "Non-rural area" delineation is depicted in Figure 25. From afar, the urban areas derived from this definition appear quite similar to those based on the definition of "Area with strong anthropogenic influence". However, as the "Non-rural area" especially considers the handling of small outlying areas, the differences in delineating it and the urban area delineation based on the two other concepts are best visible in an image of a small area, like the one that can be seen in Figure 26.

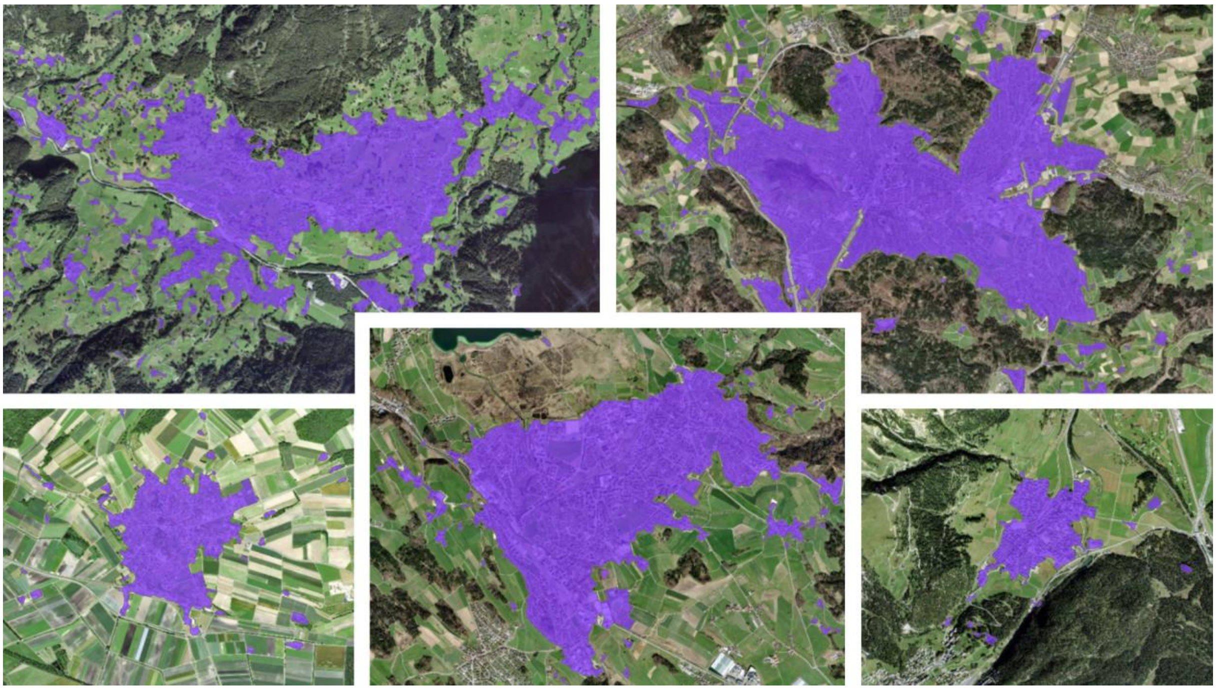


Figure 25: Delineations of the "Non-rural area": Shown are the urban areas (violet) of 2021 of the five analyzed municipalities– Grindelwald (upper left), Winterthur (upper right), Ins (lower left), Wetzikon (in the middle), and Celerina (lower right) – defining urban area as the "Non-rural area". (Source of the background image: Environmental Systems Research Institute (2022))

Locally, the disparity in the urban area between the delineations based on the concepts of "Area with strong anthropogenic influence" and "Non-rural area" can be substantial. How big the differences are, depends on the amount of SIUAFs. Such areas are common in the elevated parts of Grindelwald. This can be observed in Figure 26, where these fragmented urban areas occupy a substantial area. Thus, in the case of Grindelwald, the delineations based on the concepts of "Area with strong anthropogenic influence" and "Non-rural area" exhibit significant differences in terms of the area they encompass. In the other municipalities covered in this thesis, the difference between the delineations of the "Area with strong anthropogenic influence" and "Non-rural area" is not that striking, but nevertheless noticeable. This is especially the case when observing the delineation at a more local scale, rather than at the scale of the entire municipality.

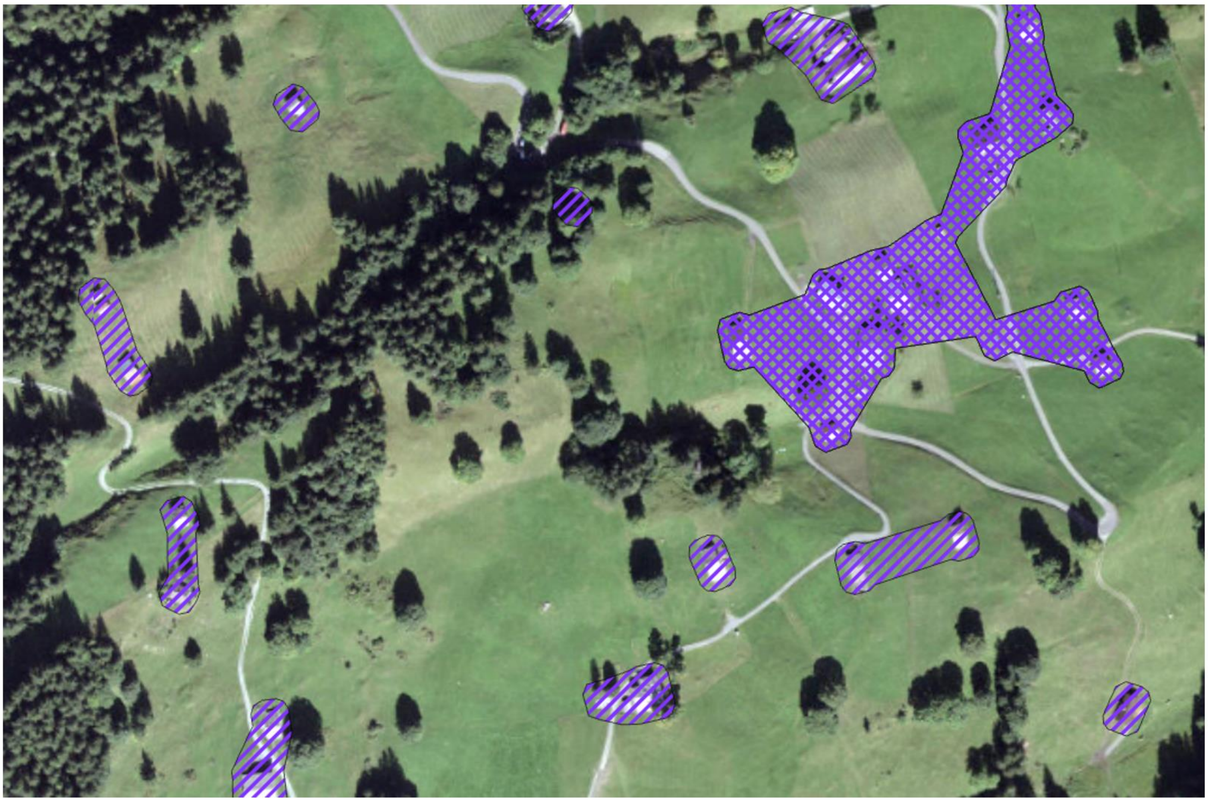


Figure 26: Differences in urban areas between the delineation based on the concept of "Area with strong anthropogenic influence" (diagonal lines and crosshatched areas) and "Non-rural area" (crosshatched areas only) in Grindelwald, using data from 2021. (Source of the background image: Environmental Systems Research Institute (2022))

## 4.2 Urban sprawl calculations

In the following, the results of the urban sprawl calculations will be presented. For every urban sprawl calculation method, the results of each year will be presented first, followed by the difference between the years. The absolute values of sprawl that are presented for the years 2015 and 2021 show the differences in urban sprawl values from a standpoint of urban sprawl as an absolute phenomenon that can be measured by considering only one point in time. The comparison of the urban sprawl values between 2015 and 2021 highlights the changes in the calculated urban sprawl values, viewing urban sprawl as a phenomenon that describes change over time. As already mentioned, both paradigms are common in urban sprawl literature, although their results can be fundamentally different: Even if a region has high absolute values for urban sprawl, it may still appear to have low sprawl when measured as a phenomenon between two points in time, if the change between those points is small. This difference in concepts becomes even more obvious when the sprawl decreases between two years.
In this thesis, the difference between the years is determined by calculating the value of the year 2021 minus the one of the year 2015. The measurements of density, WUP, and Moran's I will be employed to make statements on the state of urban sprawl of the municipalities, using the formulas, concepts and values previously discussed. The next sections focus on comparing the values of urban sprawl obtained from the various urban area delineations, in order to assess the impact of the use of different urban area delineations on the measurement of urban sprawl. In the following, the "values" of the "Built-up area", "Non-rural area", or "Area with strong anthropogenic influence" are mentioned. This terminology is used to denote the values resulting from the discussed sprawl measurement method on the basis of the "Built-up area", "Non-rural area", or "Area with strong anthropogenic influence", respectively.

### 4.2.1 Density

The results of the density calculation for the years 2015 and 2021 are illustrated in Figure 27. Higher density values are associated with lower sprawl in urban sprawl literature. In Wetzikon and Winterthur, the density measures of the "Built-up area" delineation were clearly higher than the ones of the other two delineation methods in both years. The value of the "built-up area" delineation was in both cases more than 1.15 times as large as the value of the other two delineation types, regardless of the analyzed year. The absolute difference in value between the types of delineations was not that pronounced in the other municipalities, due to the generally lower density values. However, here too, differences were noticeable, for example in the case of Grindelwald. In this municipality, a notable difference between the values of the versions of the "Built-up area" and the "Area with strong anthropogenic influence" was present. Especially in the case of Grindelwald and Celerina, the differences between the values of the delineation types of "Built-up area" and "Area with strong anthropogenic influence" were larger compared to the ones between the delineation types of "Built-up area" and "Non-rural area". Despite not being that noticeable in the other municipalities, this pattern was nevertheless present, regardless of the analyzed year.

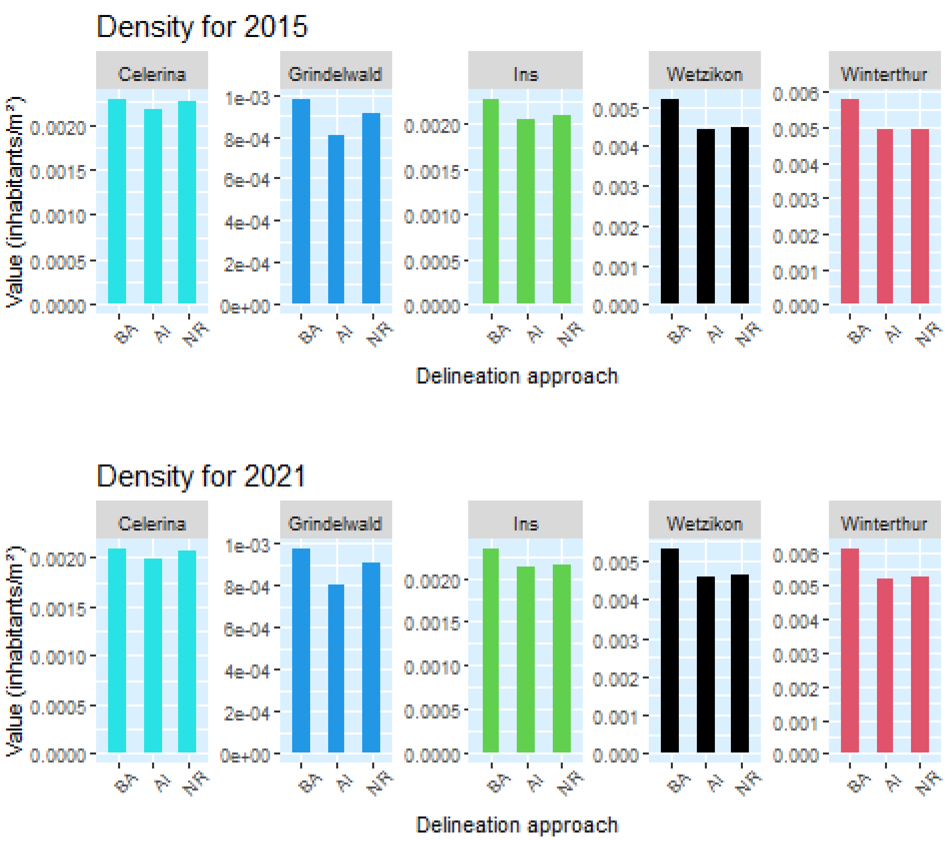



Figure 27: Density calculations for the years 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

The comparison between the change values of the years 2015 and 2021 is depicted in Figure 28. A positive change means that a densification took place between 2015 and 2021, whereas negative values signal a lower density in the year 2021 than in the year 2015. In the cases of Celerina, Grindelwald, Ins, and Wetzikon, the most positive absolute change was the one of the delineation of the "Area with strong anthropogenic influence". An exception to this rule was Winterthur, where the highest absolute change value was the one of the "Built-up area". The direction of the change was distinct in the different municipalities. In the case of Celerina and Grindelwald, all values were negative, whereas in the other three municipalities, they were positive. The change in Grindelwald was strikingly low, especially compared with the rather strong negative change in Celerina and the strong positive change in Winterthur.

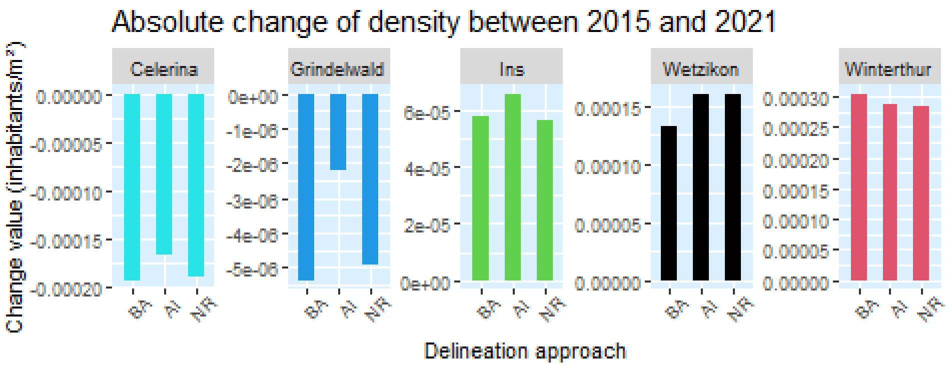



Figure 28: Absolute change in density values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

### 4.2.2 Moran's I

The sprawl values using Moran's I for the different municipalities, utilizing the three delineation methods, can be seen in Figure 29. Higher values of Moran's I indicate more spatial autocorrelation (i.e., more compactness) and thus a less sprawled settlement. The sprawl results of the Moran's I calculation resulted in many values near to 1. In certain cases, especially in Celerina, the differences between the distinct delineations were not striking. The differences were much more evident in the case of Grindelwald, where the "Built-up area" delineation clearly yielded the lowest result and the "Non-rural area" delineation the highest one. In this case, the value based on the "Non-rural area" was more than 1.20 times as large as the one using the "Built-up area" delineation, regardless of the year. However, the pattern of the lowest value for the "Built-up area" and the largest value for the "Non-rural area" was not exclusive to Grindelwald, but was present in every single one of the five municipalities, regardless of the analyzed year.

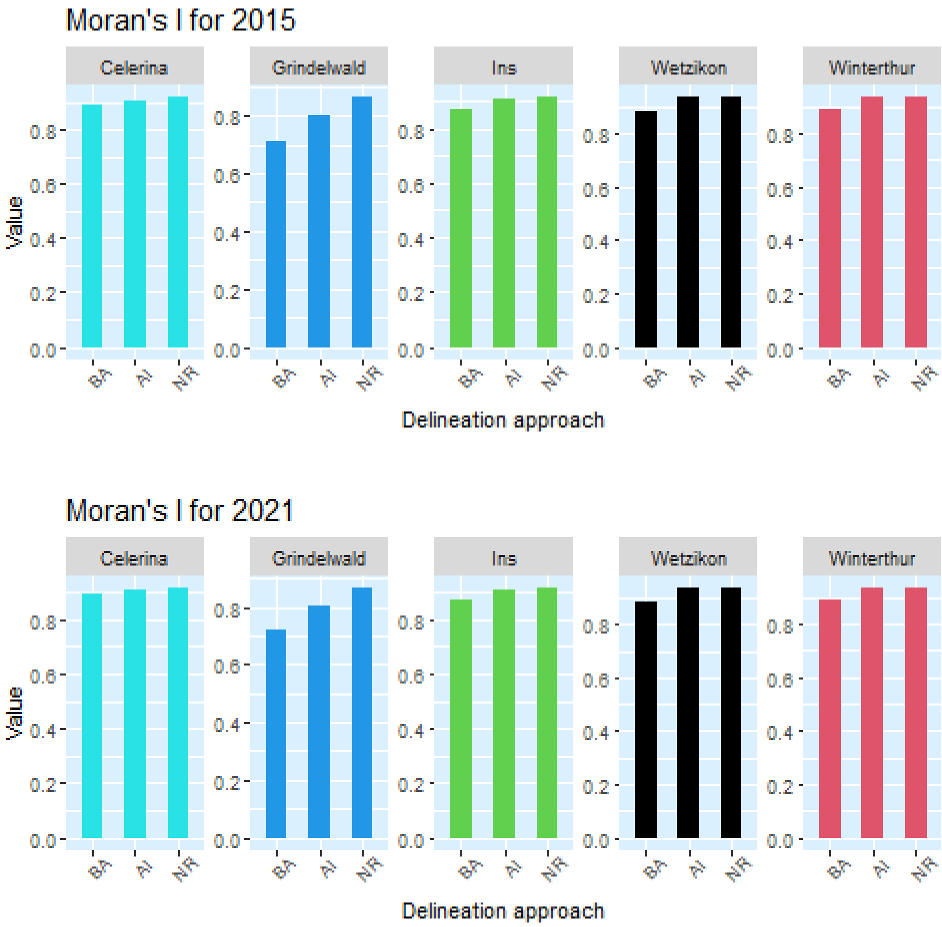



Figure 29: Moran's I values for the years 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

The absolute changes in Moran's I are depicted in Figure 30. A positive change indicates a growth of Moran's I between 2015 and 2021, meaning the urban area became more compact and thus less sprawled. The difference between the values of the distinct delineation methods in 2015 and the ones in 2021 was mostly quite pronounced. Wetzikon and Winterthur, the two municipalities with the largest population among the five, exhibit a stark contrast between the values of the "Built-up area" delineation and the other two urban area delineations. In the case of Wetzikon, the change of the sprawl value employing the "Built-up area" definition was more than 12 times as large as the one using "Non-rural area" to define urban areas. Whereas in the cases of Wetzikon and Winterthur the absolute changes using the delineations of the "Area with strong anthropogenic influence" and the ones of "Non-rural area" were quite similar, in the case of Grindelwald and Ins this change was more pronounced. Of the five municipalities, Celerina was the only one that contained a negative change value. This negative value was only present in the change value of the "Non-rural area", which was the lowest value in the other municipalities, too. This negative value in Celerina was quite pronounced, compared with the other two change values of this municipality.

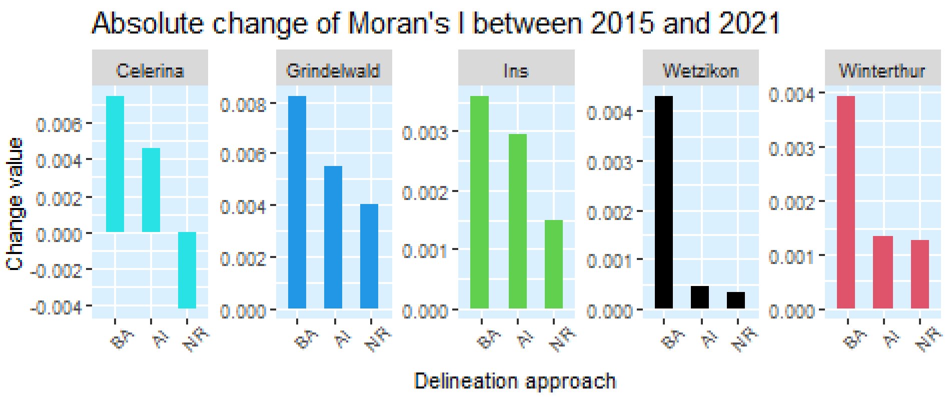



Figure 30: Absolute change in the Moran's I values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

### 4.2.3 WUP

In the following, the final Adjusted WUP result (i.e., WUP, adjusted to consider the uninhabitable area) will be presented. In addition, the results of the different components of WUP (DIS, UP, and UD) will be shown. These calculations rely solely on population and employment data pertaining to the delineated urban area. Apart from that, a version of the WUP will be presented that does not consider the uninhabitable area, here referred to as Unadjusted WUP.

#### 4.2.3.1 Adjusted WUP

Figure 31 displays the measured urban sprawl values using the Adjusted WUP calculation. More positive values indicate higher levels of urban sprawl. These values showed differences, depending on the employed delineation method. Regardless of the year, the "Area with strong anthropogenic influence" yielded the highest value when using Adjusted WUP. However, which delineation method resulted in the lowest sprawl value varied: In the cases of Celerina, Grindelwald, and Ins the "Non-rural area" led to the lowest sprawl values, whereas in the cases of Wetzikon and Winterthur, the "Built-up area" generated the lowest values. The contrast between the value obtained from the "Built-up area" and the "Area with strong anthropogenic influence" was most pronounced in Winterthur in the year 2021, with the former being roughly half the size of the latter. In contrast to this large difference in values, the values of Ins were more similar. In the case of Grindelwald, the difference between the sprawl values based on the "Area with strong anthropogenic influence" and the "Non-rural area" was clearly noticeable. This is the case because the sprawl value based on the "Area with strong anthropogenic influence" was more than 1.4 times the value of the "Non-rural area", regardless of the year.

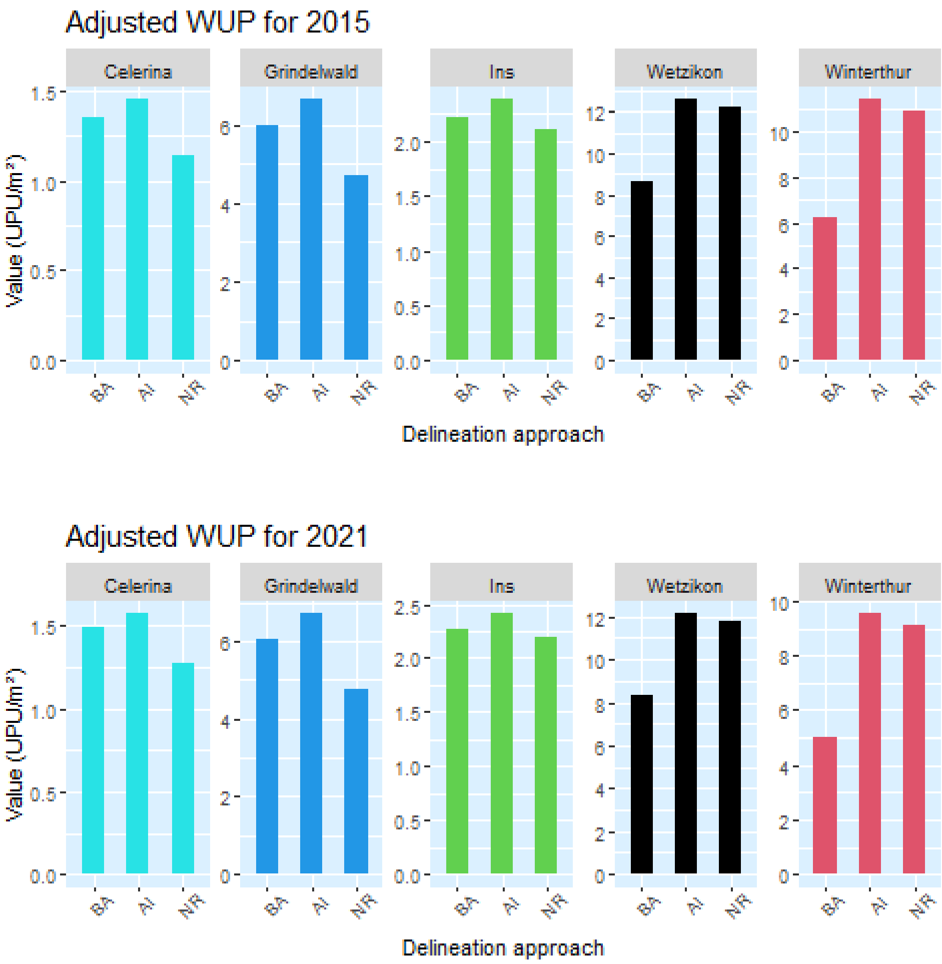



Figure 31: Adjusted WUP values for the years 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

The changes between 2015 and 2021 using Adjusted WUP can be seen in Figure 32. A positive value indicates an increase in sprawl while a negative value indicates a decrease. In three villages – Celerina, Grindelwald, and Ins – the WUP value went up during the six years between 2015 and 2021. On the contrary, the opposite was true for Wetzikon and Winterthur. The way the urban area delineation affects the values of the change of Adjusted WUP was highly variable, depending on the municipality. In the case of Celerina and Ins, the biggest change was the one based on the concept of "Non-rural area" and the smallest the one based on the definition of "Area with strong anthropogenic influence". In the case of Grindelwald, the "Area with strong anthropogenic influence" delineation led to the largest values, followed by the delineation of the "Built-up area". In Wetzikon and Winterthur, the largest change stemmed from the implementation of the concept of "Area with strong anthropogenic influence", followed by the one based on the "Non-rural area". The differences between the values stemming from the different delineation methods could be quite large: In Winterthur, the change value of the version using the "Area with strong anthropogenic influence" was almost 1.5 times as large as the one of the version based on "Built-up area". Moreover, in the case of Ins, the change value of the "Non-rural area" was more than 2.4 times as large as the one of "Area with strong anthropogenic influence".

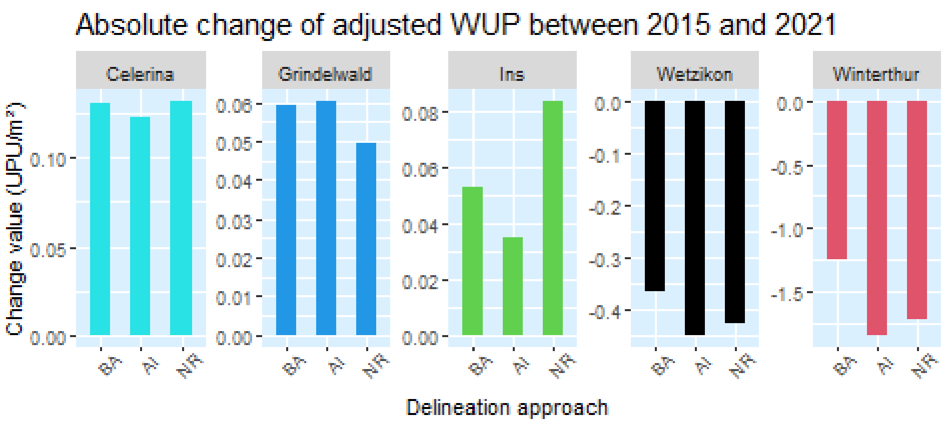



Figure 32: Absolute change in the Adjusted WUP sprawl values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

### 4.2.3.2 Unadjusted WUP

The Unadjusted WUP values of the year 2015 and 2021 led to similar patterns as the ones described for the Adjusted WUP values. However, the sprawl values were generally lower than in the adjusted version. The pattern of the Unadjusted WUP values can be seen in Figure 33. Like for Adjusted WUP, larger values mean more sprawl.

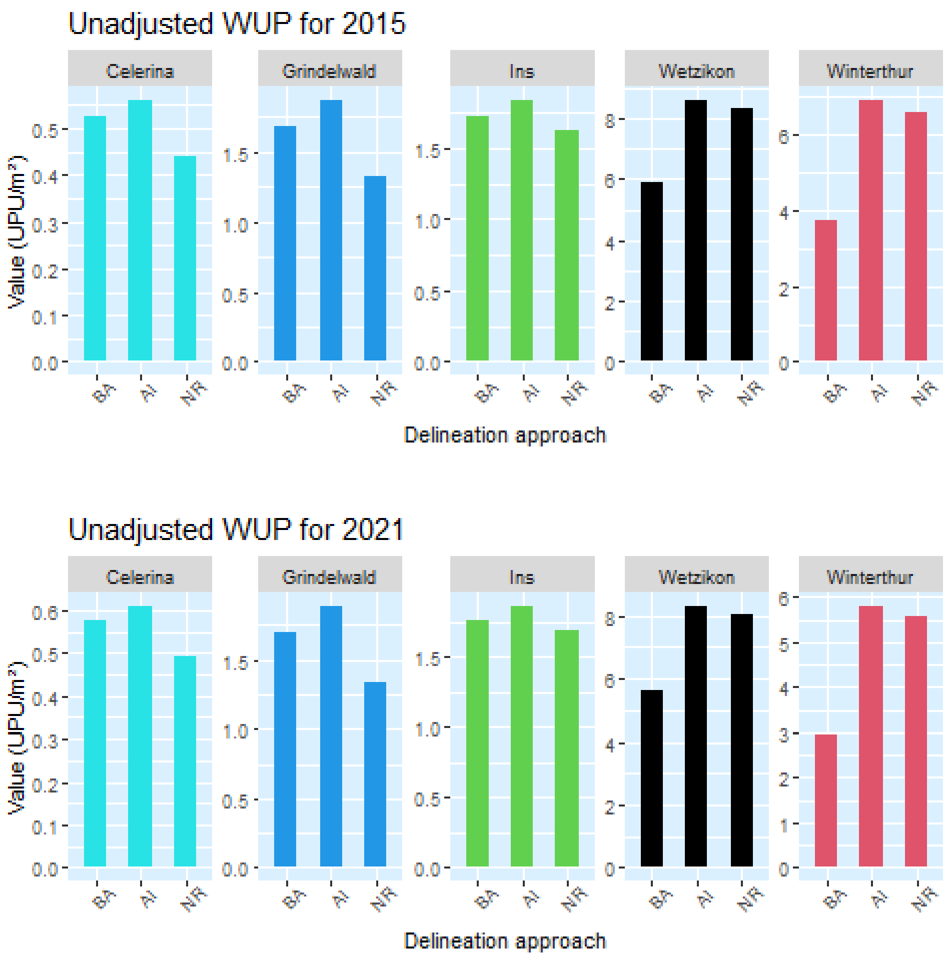



Figure 33: Unadjusted WUP values for the years 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

The same goes for the change in Unadjusted WUP, as can be seen in Figure 34. Positive values mean an increase in sprawl, negative ones a decrease.

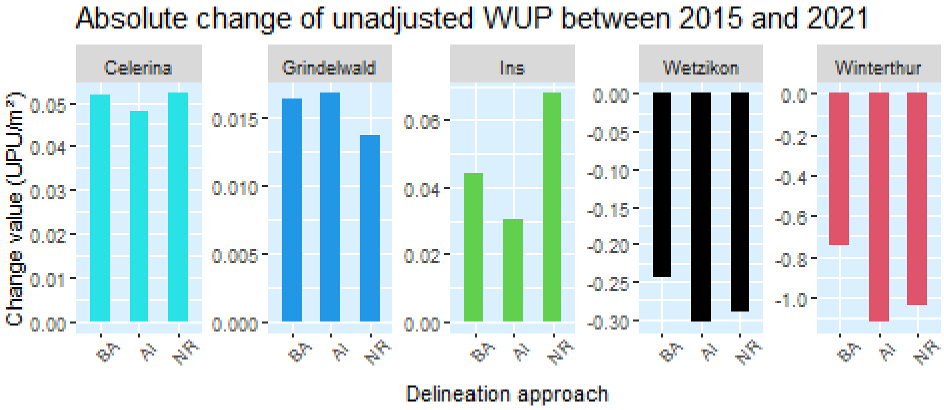



Figure 34: Absolute change in the Unadjusted WUP sprawl values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

### 4.2.3.3 UP

The UP value is a component of the WUP value, which itself is also used to assess sprawl (Nazarnia et al., 2016). The values of UP in the years 2015 and 2021 can be seen in Figure 35. Larger values mean higher sprawl. Here too, differences between the values could be observed. Regardless of the year, the delineations based on the concept of "Area with strong anthropogenic influence" led to the highest value. The urban sprawl delineation leading to the lowest value varied, depending on the municipality which was analyzed. The largest proportional difference between two delineation methods in a municipality was the one between the delineation based on the viewpoint "Area with strong anthropogenic influence" and the one based on the perspective of "Non-rural area" in Grindelwald for the year 2015. In this case, the value of the delineation based on the concept of "Area with strong anthropogenic influence" was equivalent to more than 1.30 times the one based on the definition of "Non-rural area".

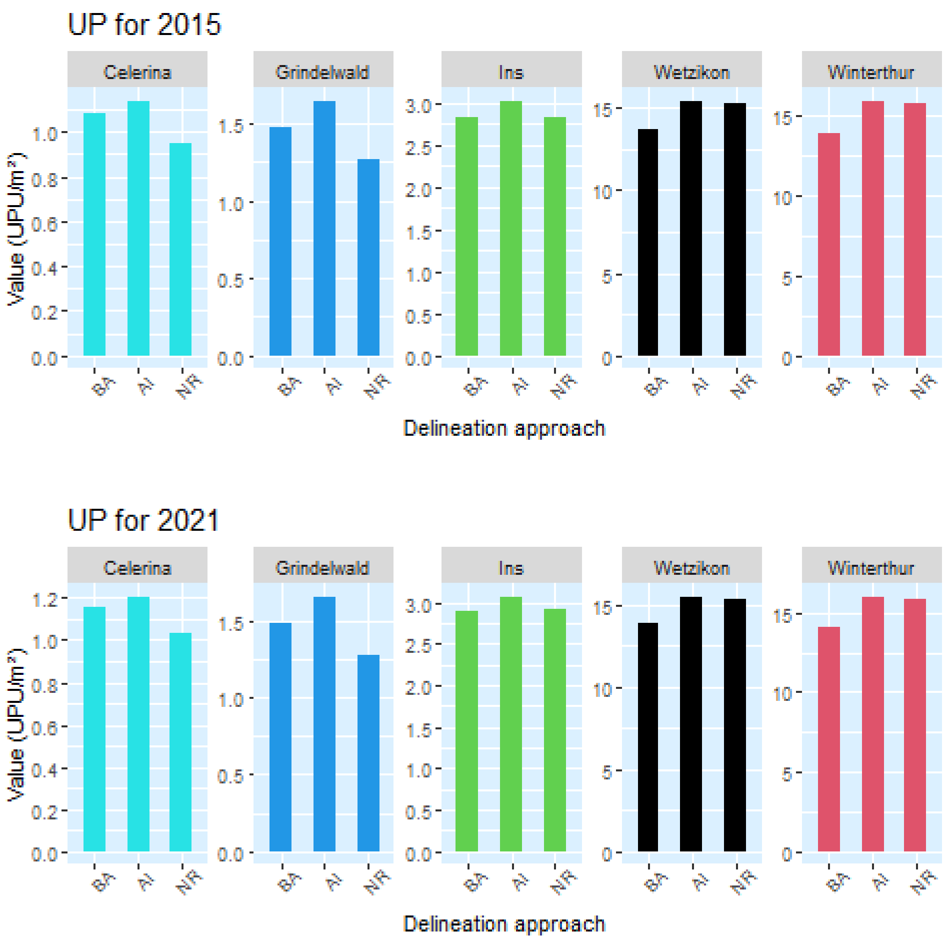



Figure 35: UP values for the years 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

The changes in UP between 2015 and 2021 exhibited substantial disparities depending on the delineation method utilized. This was especially the case in Wetzikon and Winterthur. In Wetzikon, the difference between the change value of the delineation reflecting the "Built-up area" was approximately 1.85 times as large as the one of the delineation of the "Area with strong anthropogenic influence". It was also approximately 1.75 times as large as the one outlining the "Non-rural area". In Winterthur, the change in value between 2015 and 2021 based on the "Built-up area" was equivalent to more than 2.50 times the value for the "Area with strong anthropogenic influence". In Ins, there was also a large discrepancy in the change value between the delineation approach of the "Area with strong anthropogenic influence" and of the "Non-rural area": The change value corresponding to "Non-rural area" nearly equaled 1.80 times the change value belonging to the "Area with strong anthropogenic influence". In Wetzikon and Winterthur, there was a pattern where the change values were the largest when using the "Built-up area" as a basis for the calculation, while the lowest change values were obtained when using the "Area with strong anthropogenic influence". In Celerina and Ins, the lowest change value was still produced based on the "Area with strong anthropogenic influence", while the largest change value was generated using the "Non-rural area". The change values for UP between the years of 2015 and 2021 are illustrated in Figure 36. The magnitude of the positive values indicates the increase in the extent of urban sprawl over the specified time period, based on the UP measurement.

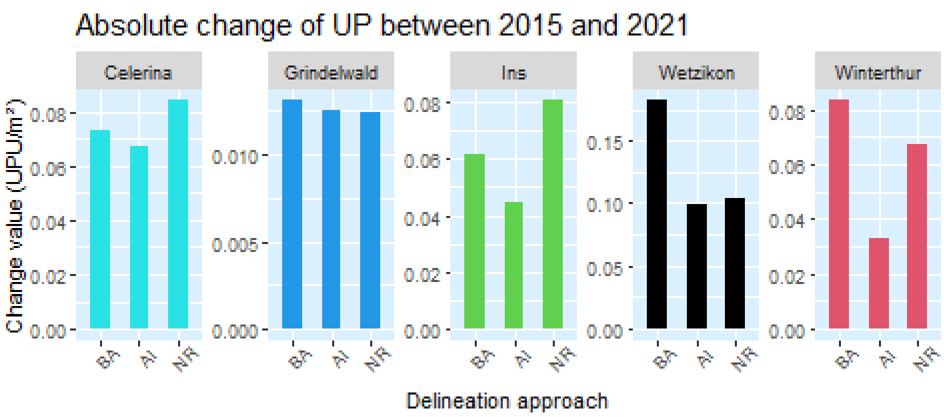



Figure 36: Absolute change in the unadjusted UP values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

### 4.2.3.4 UD

The values of UD, another component of WUP, are shown in Figure 37. Higher UD values led to more detected sprawl in the WUP calculation. In this analysis, only the population and employment values of those areas that were within the defined urban area were considered, as opposed to those of the entire municipality. Irrespective of the year or municipality studied, the approach of delimiting “Area with strong anthropogenic influence” consistently resulted in the highest UD-values. The differences between the values obtained through the “Area with strong anthropogenic influence” delineation method and the values obtained through the other two concepts of urban area delineation varied across municipalities. Also, which delineation method led to the lowest value depended on the municipality analyzed. The greatest proportional disparity in the UD values within a municipality could be observed between the “Area with strong anthropogenic influence” and the “Non-rural area” in Grindelwald in 2021: The value for the “Area with strong anthropogenic influence” was almost 1.25 times as large as the one of the “Non-rural area”.

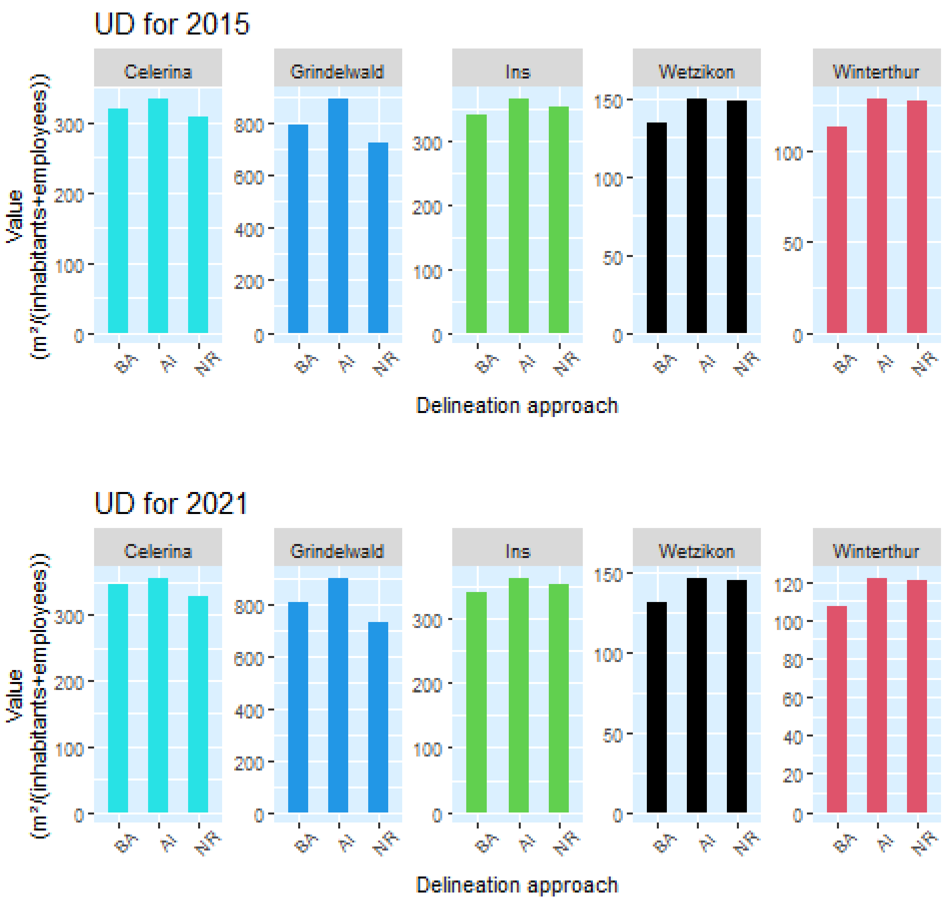



Figure 37: UD values for the years 2015 and 2021. The different delineation methods are abbreviated: “Built-up area” (BA), “Area with strong anthropogenic influence” (AI), and “Non-rural area” (NR).

The change in the values between 2015 and 2021 in the five municipalities can be observed in Figure 38. Positive values signify growth in the UD over the specified time frame, indicating an increase in the usage of space by people and employees. Conversely, negative values denote a decrease in the UD. In all municipalities except for Ins, the different delineation methods resulted in either all positive or all negative change values. Ins stood out from the other municipalities because the delineation methods of the "Built-up area" and "Non-rural area" resulted in positive change values, while the "Area with strong anthropogenic influence" delineation method produced a negative change value. Ins deviated from the general trend in yet another way, since the delineation of the "Built-up area" did not yield the highest positive change value. Among the four other municipalities, the greatest proportional discrepancy in change values could be observed in Grindelwald: The change value resulting from the delineation of the "Built-up area" was more than 1.60 times as large as the one from the delineation of the "Non-rural area".

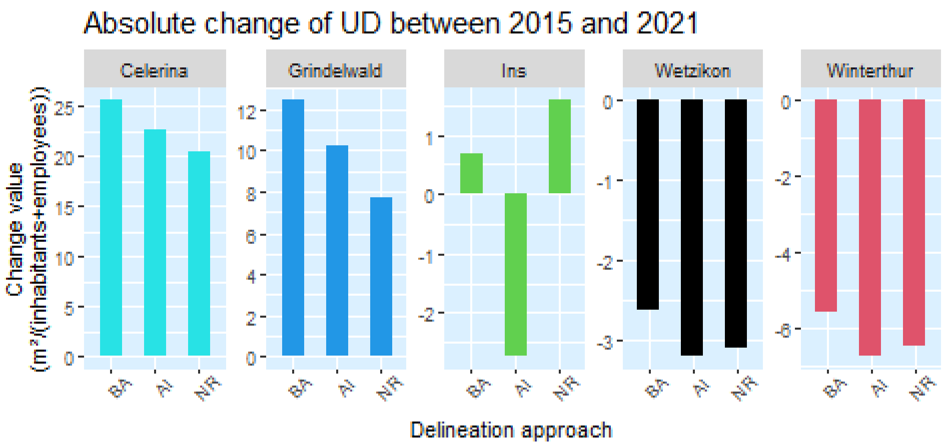



Figure 38: Absolute change in the UD values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

#### 4.2.3.5 DIS

The values of DIS, yet another component of WUP, are visualized in Figure 39. Higher values led to higher WUP values, i.e., more sprawl. Although differences between the values of the different delineation methods exist, they were not huge. In the cases of Wetzikon and Winterthur, the differences between the values obtained by the different delineation methods were particularly small.

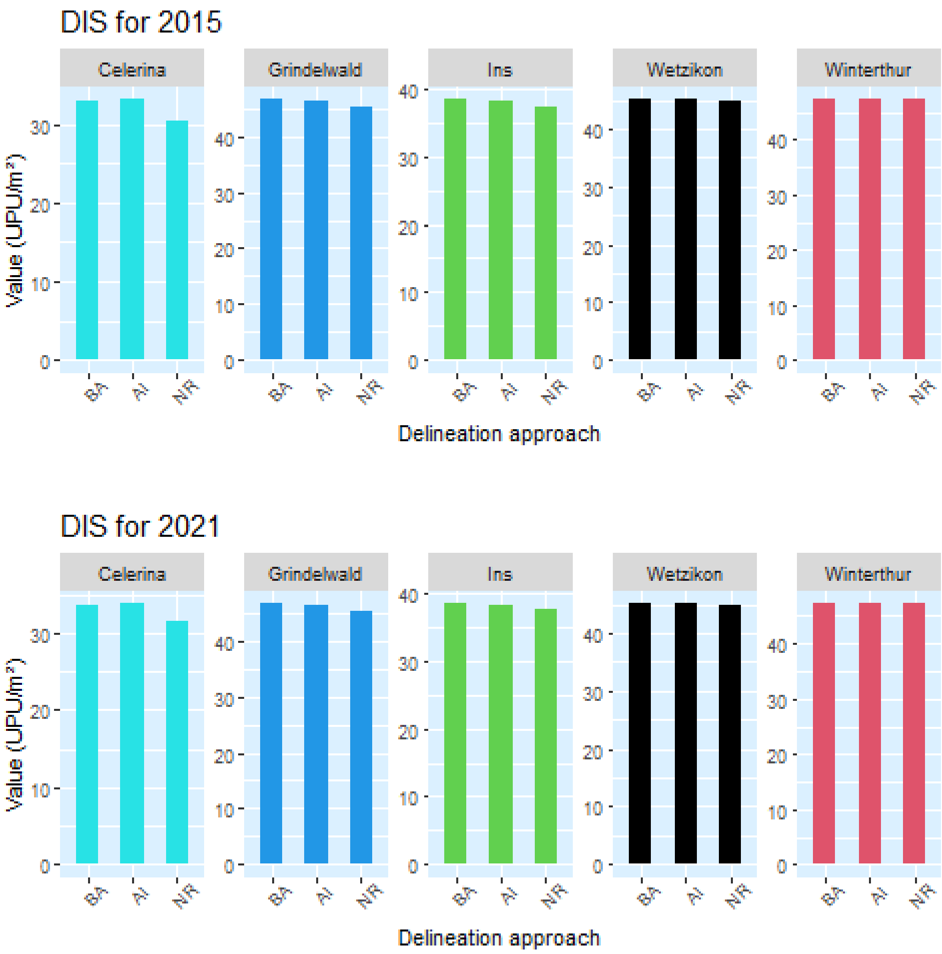



Figure 39: DIS values for the years 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

The changes in the DIS values between 2015 and 2021 are displayed in Figure 40. Higher values mean that DIS grew in the mentioned timeframe. The change values resulting from the different delineation approaches exhibited striking disparities. In all cases, except Winterthur, the different change values were consistently positive. Among these four cases, the largest proportional differences in change between the different delineation methods occurred in the municipality of Grindelwald, specifically between the delineation methods of "Area with strong anthropogenic influence" and "Non-rural area". Here, the change value of the "Area with strong anthropogenic influence" was more than 12 times as large as the one of the "Non-rural area". For the municipality of Ins, considerable differences between the delineation methods of "Area with strong anthropogenic influence" and "Non-rural area" existed, too. In this case, the change value of the "Non-rural area" was more than 5.25 times as large as the one of the "Area with strong anthropogenic influence". In Winterthur, the delineation method of the "Non-rural area" resulted in a positive change from 2015 to 2021, while those of the "Built-up area" and "Area with strong anthropogenic influence" resulted in a more pronounced negative change.

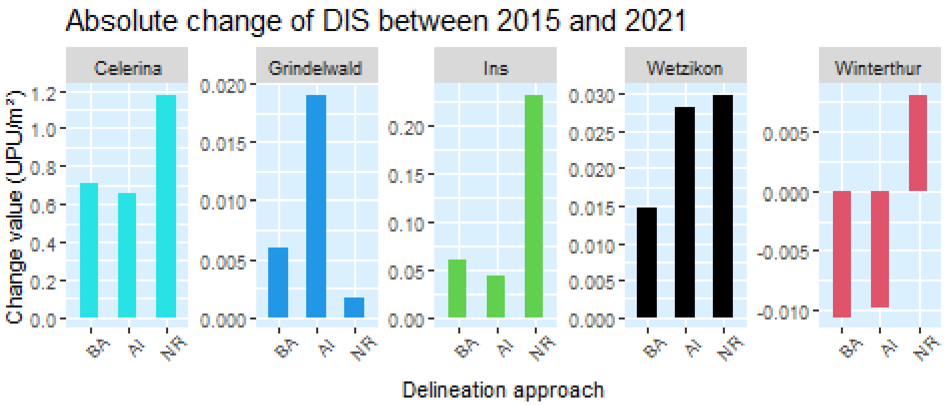


Figure 40: Absolute change in the DIS values between 2015 and 2021. The different delineation methods are abbreviated: "Built-up area" (BA), "Area with strong anthropogenic influence" (AI), and "Non-rural area" (NR).

# Chapter 5 | Discussion

## 5.1 Analysis of the Results

The results of the different sprawl measurement methods, i.e., the density calculation, Moran's I, and WUP, including its components, shows disparities in the values depending on the employed method of urban area delineation. In the following, the results of the density calculations and the Moran's I calculations will be critically discussed, and reasons for the observed differences in their values will be presented. This topic will not cover WUP and its components. This is partly due to their complex calculation, rendering it difficult to associate their values with observable phenomena. In addition, the values of the WUP calculation are, at least to a large part, subject to the same phenomena as the density and the Moran's I calculation. This means that the reasons for the changes in WUP can also be deduced – at least in part – from the information provided concerning the calculations of density and of Moran's I.

### 5.1.1 Absolute Density Values

The differences between the density values obtained using the three delineation methods in both 2015 and 2021 show variations across the five municipalities. The similarity between the values of the "Area with strong anthropogenic influence" and "Non-rural area" compared with the larger value of the "Built-up area" in Wetzikon and Winterthur may be attributed to two factors: 1) Both municipalities contain holes in the "Built-up area" that lead to a larger urban area in the other two delineations; and 2) both municipalities lack a significant presence of SIUAFs. As a result of this second factor, the value of the "Non-rural area" remains comparable to the value of the "Area with strong anthropogenic influence", unlike in Grindelwald, where the inclusion of urban green spaces is offset by the exclusion of the substantial area of SIUAFs.

### 5.1.2 Change in Density Values Between 2015 and 2021

The cause of the differing levels of change between the various delineation methods remains uncertain, but it could be the result of the observed construction and destruction of individual isolated structures and the development of housing within urban green spaces. Moreover, the shift in density between 2015 and 2021 can also be attributed to the emergence of new holes within the urban areas, resulting from a change in building delineation quality, as discussed in Section 5.5.5.

Interestingly, the change values reveal a lowering of the density between 2015 and 2021 in Celerina and Grindelwald, whereas the other municipalities underwent densification. This probably is to a large part due to the fact that the population of Celerina and Grindelwald shrunk in the years between 2015 and 2021 (Bundesamt für Statistik, 2022d).

### 5.1.3 Absolute Moran's I Values

In the case of the Moran's I, all values indicate a positive autocorrelation. This makes sense, given the small pixel size of 15x15 m. The trend towards increased spatial autocorrelation from the "Built-up area" delineation method to the "Area with strong anthropogenic influence" and ultimately to the "Non-rural area" can be easily explained: The delineation of the "Built-up area", which contains small outlying urban areas and holes comprises more urban pixels surrounded by non-urban area and vice versa. In the case of the "Area with strong anthropogenic influence", this effect is dampened, as the holes are classified as urban. Finally, the "Non-rural area", which does not contain the small outlying urban areas, has the least number of urban pixels that border non-urban area pixels and vice versa. The pronounced effect of the delineation method on the Moran's I values in the municipality of Grindelwald is comprehensible, given its abundance of SIUAFs and a main settlement characterized by numerous large holes in the delineated "Built-up area".

### 5.1.4 Change in Moran's I Values Between 2015 and 2021

The fact that the change values for Moran's I for the "Built-up area" are more positive than the ones of the "Area with strong anthropogenic influence" and the "Non-rural area" can be explained by the fact that the "Built-up area" presents a larger perimeter on which change can take place.

The positive change values indicate a net expansion of the urban area, which, due to the mentioned principle, is most effectively captured by the "Built-up area" delineation, followed by the "Area with strong anthropogenic influence", which exhibits the second largest perimeter. The negative change value in Celerina for the "Non-rural area" is intriguing, as no easy explanation could illuminate why such a trend is possible, especially as the other two delineation methods lead to positive change values. A possible explanation could be data quality issues, as discussed in Section 5.5.5.

## 5.2 Significance of Moran's I

The results of the Moran's I calculation are highly significant in all cases. The exact p-value remains unknown, but must be very low, as the displayed value was 0.000000. The z-value varied from positive lower to higher three-digit figures.

## 5.3 Comparison Between the Measurements

The measurements employed in this thesis, density, Moran's I, and WUP, are all used in literature to assess urban sprawl, either by solely relying on one of these measures or in combination with each other or with other measurements. However, if each of these measures is used as a standalone indicator of sprawl, it becomes evident that the results obtained are different. Despite the differences in urban area delineation, two municipalities – Celerina and Grindelwald – experienced a decrease in density between 2015 and 2021, while the remaining three municipalities exhibited an increase in density. In contrast to the change in density, the change values of the Moran's I calculation show an increase in spatial autocorrelation in all municipalities over the same time period, with the exception of one delineation method in Celerina. Thus, the results of the Moran's I calculation obviously contradict the results of the density measurement. The analysis using WUP indicates that three out of the five municipalities exhibit an increase in urban sprawl, while the remaining two show a decrease between 2015 and 2021. This is inconsistent with the results obtained through the calculation of density and Moran's I. The UP, which is not only a component of WUP, but can also be employed as a sprawl measurement on its own (Nazarnia et al., 2016), exhibits a positive change in every municipality, indicating an increase in sprawl in every case. This again contradicts the results of the other urban sprawl measurements. Therefore, it becomes evident that the calculation methods can lead to contradicting results. Hence, it is advisable to consider using multiple measurements to provide a comprehensive understanding of urban sprawl in a region, as suggested by Zhou et al. (2019). The discrepancy in the results produced by the UP and the WUP measurements, despite both satisfying all or almost all the suitability criteria of Jaeger, Bertiller, Schwick, & Kienast (2010), suggests that in spite of all the efforts to find the best way to calculate urban sprawl, different methods that meet the criteria for a "good" measurement can still produce different results.

The varying impact of the different urban sprawl calculation methods on the results is also apparent when comparing which municipalities had the highest or lowest levels of sprawl in 2015 or 2021. The lowest density value, irrespective of the year, can be found in Grindelwald, the highest ones in Wetzikon and Winterthur. The same pattern is present in the results of the Moran's I calculation. Thus, based on the density values and the Moran's I values, it can be inferred that Wetzikon and Winterthur exhibit low levels of sprawl, while Grindelwald has high levels of sprawl. Similarly, high values for Wetzikon and Winterthur are also reflected in the WUP metric. However, using this metric, higher values indicate greater levels of sprawl, suggesting that Wetzikon and Winterthur are actually the most sprawled municipalities. This is likely because WUP, in contrast to the density measurement and the Moran's I measurement, penalizes municipalities with a large amount of urban area.

## 5.4 Comparison With Existing Urban Sprawl Delineations

To assess the similarity of the urban area delineation of this thesis with already existing ones, a comparison was conducted between the urban areas determined in this study and already existing delineations of urban area in Switzerland. The urban areas used for comparison in this thesis are the delineations of the Swiss Arealstatistik (Bundesamt für Statistik, 2021) and of the OpenStreetMap (*OSM*) (Geofabrik & OpenStreetMap Contributors, 2022). The data of the Swiss Arealstatistik were collected during the survey period of 2013 to 2018, whereas the OSM data are from the 12th November 2022, 21:21 o'clock (Bundesamt für Statistik, 2021; Geofabrik & OpenStreetMap Contributors, 2022). The data of the Arealstatistik are a product of the Swiss federal government and are made up by data points in a 100x100 m grid, which are assigned to a class. There are different versions of the Arealstatistik: one about the land cover of Switzerland, one about the land use of Switzerland and a combined map that encompasses both (Bundesamt für Statistik, 2018). This last version has the greatest number of classes (Bundesamt für Statistik, 2018) and is discussed here. The urban areas delineated in this thesis were compared with the classes of the Arealstatistik that lie in the category of "Settlement and urban areas". The areas classified under said category, as well as the overall classification scheme of the Arealstatistik, can be viewed in Figure 41.

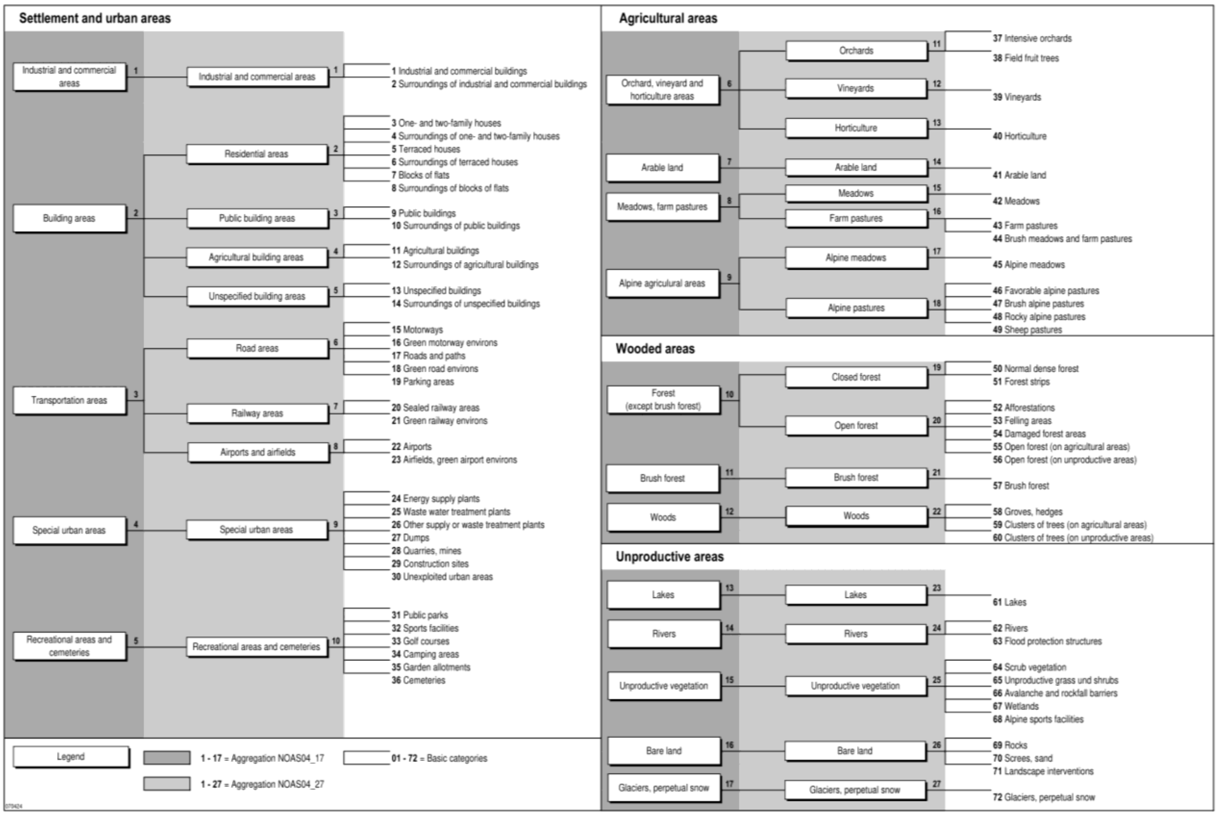



Figure 41: Classes of the combined map of land cover and land use of the Arealstatistik. The 72 basic categories (in the white areas) can be aggregated to higher-level categories (in light and dark grey or in bold) (Bundesamt für Statistik, 2006). One of them is the category "Settlement and urban areas", which is of particular interest in this thesis. (Image source: Bundesamt für Statistik BFS (2006))

Due to the fact that the Arealstatistik data are delivered in the form of points, these points were assigned 100x100 m squares centered on the points. This method of converting the point data of the Arealstatistik into polygon data was inspired by the federal Arealstatistik map on map.geo.admin.ch (Swiss Confederation, 2023), which also visualizes the data as squares. The Swiss Arealstatistik offers detailed data compared with the CORINE land cover dataset, which was considered as another option for comparison with the urban areas outlined in this thesis (Steinmeier, 2013).

The OSM data represents data created by a community of volunteer contributors. This dataset was selected as a comparison to the urban areas outlined in this thesis, as it provides insights into how individuals perceive and categorize the landscape. This relates to the viewpoint of Jaeger, Bertiller, Schwick, Cavens, et al. (2010), which argue that urban sprawl is a visually perceptible phenomenon, and also to the approach used by Schwick et al. (2018) to delineate urban areas – a delineation based on human subjective perception. The OSM dataset used for comparison consists of areas classified under the "Developed land" category in the OSM classification system (OpenStreetMap Wiki contributors, 2022). This class encompasses areas characterized by the presence of commercial activity, construction, educational areas, fairgrounds, industry, residences, retail, or institutions (OpenStreetMap Wiki contributors, 2022). It is important to note that both the Arealstatistik data and the OSM data were not specifically designed for urban area delineation purposes. Still, they served as a reference to evaluate the similarity of the urban area delineation presented in this thesis to already existing urban area delineations.

To calculate the similarities between the urban area delineations created for this thesis with already existing urban areas, the Jaccard index was employed, which is calculated as:

$$Jaccard\ index = \frac{\mathrm{a}(\mathrm{A} \cap \mathrm{B})}{\mathrm{a}(\mathrm{A} \cup \mathrm{B})}$$

where A is the urban area according to one delineation, B is the urban area according to another delineation, and a() denotes the area size. The larger the overlap between the two areas, the higher the value of the Jaccard index will be. A value of 0 indicates that there is no overlap at all between the two urban areas, while a value of 1 signifies perfect overlap. The results of the mentioned calculation comparing the urban area delineated for this thesis with the OSM dataset and the Arealstatistik data can be viewed in the Tables 1 and 2. Ideally, the data of the compared urban areas should have been collected for the same year. However, this is not the case in the current comparison, which probably slightly affects the disparities in the delineation of urban areas. Nonetheless, the impact of these differences is expected to be minimal and should still permit a meaningful comparison between the delineations.

Table 1: Jaccard indices of the urban area delineations of this thesis of the year 2015 and the OSM and Arealstatistik delineations.

Celerina 2015

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.735 | 0.430 |
| AI | 0.756 | 0.446 |
| NR | 0.789 | 0.447 |

Grindelwald 2015

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.386 | 0.288 |
| AI | 0.407 | 0.298 |
| NR | 0.451 | 0.320 |

Ins 2015

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.627 | 0.393 |
| AI | 0.681 | 0.416 |
| NR | 0.693 | 0.417 |

Wetzikon 2015

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.791 | 0.692 |
| AI | 0.863 | 0.732 |
| NR | 0.863 | 0.733 |

Winterthur 2015

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.755 | 0.681 |
| AI | 0.762 | 0.716 |
| NR | 0.764 | 0.717 |

Table 2: Jaccard indices of the urban area delineations of this thesis of the year 2021 and the OSM and Arealstatistik delineations.

**Celerina 2021**

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.744 | 0.452 |
| AI | 0.758 | 0.463 |
| NR | 0.787 | 0.464 |

**Grindelwald 2021**

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.391 | 0.292 |
| AI | 0.409 | 0.298 |
| NR | 0.453 | 0.318 |

**Ins 2021**

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.636 | 0.395 |
| AI | 0.687 | 0.414 |
| NR | 0.695 | 0.416 |

**Wetzikon 2021**

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.800 | 0.700 |
| AI | 0.864 | 0.731 |
| NR | 0.865 | 0.732 |

**Winterthur 2021**

| | OSM | Arealstatistik |
|---|---|---|
| BA | 0.769 | 0.685 |
| AI | 0.767 | 0.719 |
| NR | 0.769 | 0.720 |

In general, the delineation of urban areas used in this thesis is more similar to the delineation by the OpenStreetMap community than to the delineation of the Arealstatistik. This could be attributed to the fact that the delineation of urban areas in the Arealstatistik is carried out using a 100x100 m raster, which causes inaccuracies. The OSM dataset is free from such inaccuracies. The values of the Jaccard index vary very much depending on the municipality. A rather high overlap is present in Wetzikon: The comparison between the delineation of the "Non-rural area" for the year 2021 and the OSM dataset using the Jaccard index revealed a score of 0.865. The lowest Jaccard index value occurs when comparing the "Built-up area" delineation of Grindelwald in 2015 with the delineation of the Arealstatistik. In this particular scenario, the Jaccard index yields a value of 0.288. The exact reasons for the varying Jaccard index values across the municipalities are not entirely clear, but several factors could contribute to the differences. A potential cause of the discrepancies between some of the urban area delineations used in this thesis and the one created by the OpenStreetMap community could be the complex nature of certain urban areas. This explanation may particularly shed light on why the Jaccard index in Grindelwald exhibits such low values. Apart from varying definitions what constitutes "urban areas", the OpenStreetMap contributors may also not have a strong motivation to precisely delineate the urban area, and instead opt for a more generalized approach. It is important to note that in contrast to the Arealstatistik, OSM allows for areas to be unclassified, potentially resulting in missing data. The urban area in Grindelwald according to OSM, which is depicted in Figure 42, is delineated in a much more generalized way than the delineations outlined in this thesis. Additionally, the same figure shows the delineation of the urban area as determined by the Arealstatistik.

Figure 42: The class "Developed land" of the OSM dataset (left) and the "Settlement and urban areas" of the Arealstatistik (right) in Grindelwald. (Source of the background image: Environmental Systems Research Institute (2022))

Municipalities like Wetzikon experience less problems with imprecise delineation, as the border between urban and non-urban area is not as complex anyway. The level of detail in the OSM delineation of urban areas could also be influenced by the number of inhabitants of a municipality, as a larger number of residents may be interested in accurately mapping their surroundings. However, the underlying causes behind the low Jaccard index values obtained from comparing the urban area delineations of this thesis for Grindelwald with the Arealstatistik data cannot be explained in a similar fashion as those stemming from the comparison with OSM data. It is possible that the large raster size used here results in a significantly different delineation compared with the one employed in this thesis. Additionally, as already mentioned, the Arealstatistik dataset initially only refers to a single point. Thus, particularly in the case of Grindelwald, the possibility exists that many outlying buildings or thin strips of development might not have been captured in the data of the Arealstatistik. Simultaneously, certain small buildings will result in 100x100 m pixel cells being designated as part of the urban area, despite the building and its garden not taking up nearly this amount of area.

It is commonly observed that the lowest Jaccard index values are produced by the "Built-up area" delineations, whereas the highest Jaccard indices are typically achieved for the "Non-rural area". Thus, the assumption could be made that the OpenStreetMap community and the Arealstatistik tend to delineate urban areas without outlying buildings and including green urban areas. This observation appears to hold true, at least to some extent, for the Arealstatistik: It explicitly includes public parks into the urban area (Bundesamt für Statistik, 2006). The definition of "public parks" in the Arealstatistik classification encompasses not just maintained lawns or shrubs in urban settings, but also open forests within the urban area, without the requirement for additional evidence of human impact (BFS, 2018). This contradicts the viewpoint of Schwick et al. (2018), which asserts that the extent of urban sprawl is determined by the distribution of built-up area, including outlying buildings and excluding

large parks, and that the boundaries of urban areas can be identified through human observation. The OSM delineation appears to also challenge this idea: It appears that humans, or at least those responsible for delineating urban areas on OpenStreetMap, have opted not to rely solely on the existence of built-up areas as a defining criterion for urban areas.
The delineation of the OpenStreetMap contributors and the Arealstatistik delineation were also compared with each other using the Jaccard index. The result can be seen in Table 3 and shows that, here too, the differences between the urban area delineations can be quite substantial. The values of the Jaccard index between the OSM data and the Arealstatistik data tend to be similar to the lowest values obtained when comparing the delineation methods used in this thesis with the OSM and Arealstatistik data. An exception is Winterthur, where the Jaccard index between the OSM data and the Arealstatistik data is quite high. From the overall pattern, it can be inferred that the delineation of urban areas in this thesis is comparatively similar to the OSM delineation, while the dataset of the Arealstatistik differs from both other delineations to a greater degree.

Table 3: Jaccard indices between the OpenStreetMap delineations and the ones of the Arealstatistik.

**Celerina**

| | Arealstatistik |
|---|---|
| OSM | 0.461 |

**Grindelwald**

| | Arealstatistik |
|---|---|
| OSM | 0.290 |

**Ins**

| | Arealstatistik |
|---|---|
| OSM | 0.381 |

**Wetzikon**

| | Arealstatistik |
|---|---|
| OSM | 0.725 |

**Winterthur**

| | Arealstatistik |
|---|---|
| OSM | 0.772 |

## 5.5 Limitations

The outcomes of this thesis are impacted by the choices made during the development of the final urban area delineations. These choices and their implications will be presented in the following.

### 5.5.1 Definitions of Urban Areas

The three ways to delineate urban areas, namely “Built-up area”, “Area with strong anthropogenic influence”, and “Non-rural area”, despite being based on urban sprawl literature, do not constitute widely recognized classes in urban sprawl literature. Among them, only the concept of “Built-up area” is explicitly mentioned in urban sprawl literature. The other two ways to delineate urban areas – “Area with strong anthropogenic influence” and “Non-rural area” – are newly proposed classes. This classification originated from an effort to consolidate the overarching perspectives in current urban sprawl research. Thus, the study of the effects of the three archetypes of urban area delineation does not reflect an examination of three widely recognized methods of defining urban areas, but rather an evaluation of three newly introduced classes.

### 5.5.2 Unusual Delineation Methods

The delineation in this thesis revolves around three archetypical delineation methods, namely “Built-up area”, “Area with strong anthropogenic influence”, and “Non-rural area”. In reality, most delineations used to study urban sprawl are not strictly based on any single method, but rather combine elements from different methods to varying degrees. However, in this thesis, the three mentioned “extreme cases” were compared with each other, in order to show the differences between these delineation methods. Thus, it can be concluded that the majority of delineation approaches in urban sprawl literature likely do not show such pronounced disparities between each other as those presented in this thesis purely due to differences in urban area delineation methods.

### 5.5.3 General Concept

The delineation of urban area in this thesis is based on the premise that a fixed distance can be utilized to connect areas of buildings and other urban features to form an urban area. However, this notion is not necessarily correct. Instead, it is possible that the distance for connecting these areas should not be constant but should vary between municipalities or even parts thereof. It can be argued that different municipalities and distinct areas within a municipality exhibit varying spatial distributions of components such as houses, due to differences in factors such as the width of roads. The idea that certain areas like protected wetlands or waterbodies do not necessarily interrupt urban areas, which was knowingly employed in this thesis, is also not beyond dispute. Overall, it is worth noting that while the use of the $\alpha$-shape algorithm in this thesis aligns with the definitions of urban areas found in literature to a large degree, the chosen approach may still be subject to criticism as it may clash with certain perspectives held by individual urban sprawl researchers.

### 5.5.4 Study Areas

The study areas analyzed in this thesis are five Swiss municipalities. Three of them are villages, one is a town and one is a big city. All of these municipalities are German-speaking, albeit the municipality of Celerina is bilingual. Apart from the fact that the choice of these municipalities excludes areas outside the country of Switzerland, it also excludes the municipalities in French, Italian, and (monolingual) Romansh-speaking regions within the country. In addition, all of the five municipalities are traditionally protestant and four of them – Wetzikon, Winterthur, Ins and Grindelwald – are located in the two most populous cantons in Switzerland – Zurich and Berne. It is very possible that the inclusion of municipalities with other backgrounds would uncover additional impacts of urban area delineations on urban sprawl calculations, which are not observable in this thesis.

### 5.5.5 Data Quality

Apart from effects like the increase of built-up area, the differences between the values of 2015 and 2021 are likely influenced by the changing data quality, too. An example of inconsistently delineated buildings in the utilized data from TLM3D can be seen in Figure 43. If inconsistencies are present, the delineation of 2021 generally seems to more accurately depict real-life conditions. This also results in a shrinking of the footprint of certain buildings. This has two effects for the further delineation of the urban areas. *First*, areas that were previously connected may now become disconnected as the distance between the buildings could have increased. *Second*, some areas that were sufficiently large in the 2015 delineation to not be excluded in the delineation of the "Non-rural area", may be excluded in the delineation of 2021. This is because the area attributed to the building and its garden is smaller, leading to the possibility that the size of the area falls below the threshold value. The disconnection of areas that were connected in the delineation of the year 2015 could intensify this effect. This effect is visible in the case of Celerina and may account for the substantial decrease in its Moran's I between 2015 and 2021, using the "Non-rural area" delineation. The large drop in Moran's I in Celerina can be attributed to the small size of the urban area of this village. Even a single occurrence of the mentioned effect could result in a considerable reduction of the urban area. In instances where the building footprints have increased in size between 2015 and 2021, the same two mentioned effects can also occur but with opposite consequences.

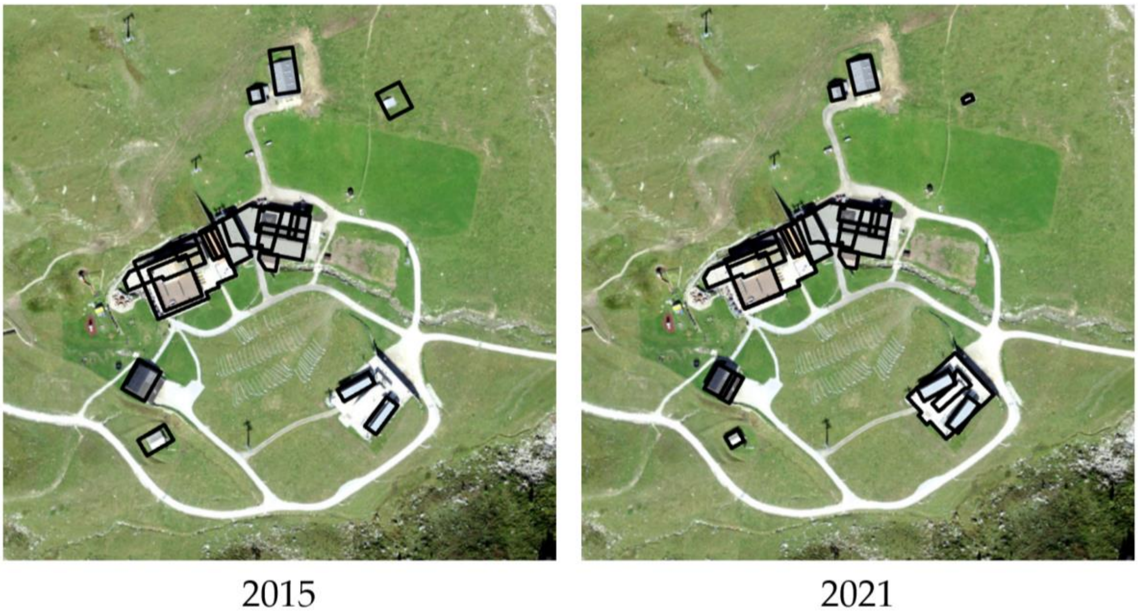



Figure 43: Change in delineation of buildings between 2015 and 2021 in the utilized data from TLM3D. Analyzing past Google Earth imagery (Google Earth, 2023), no changes in the buildings can be observed that would account for the differences in the delineation between 2015 and 2021. (Source of the background image: Environmental Systems Research Institute (2022))

### 5.5.6 Uninhabitable Areas

For the approaches of Moran's I and WUP, uninhabitable areas like forests and waterbodies were considered inhabitable when they overlapped with the delineated urban area. This made sense in most cases, as forests within the urban area were technically uninhabitable but were still used in an urban context, at least according to the delineations of the "Area with strong anthropogenic influence" and "Non-rural area". However, the overlap between the designated urban areas and the utilized data on uninhabitable areas raised concerns about the accuracy of the latter's delineation in some instances: Some areas classified as uninhabitable due to steep terrain occasionally contained buildings, for example. Thus, it has to be assumed that the delineation of the uninhabitable areas used in this thesis may not be completely accurate.

### 5.5.7 Included Areas

The areas included to form the urban areas are based on the delineations of the three discussed urban area definitions, as described in Chapter 3 of this thesis. However, it can be debated whether the included areas are truly adequate in all cases. First, the data used to determine built-up areas may not always reflect actual built-up areas. For instance, uncertainty arises in the case of sports grounds, which can be areas such as basketball courts, which exhibit hard surfaces, or areas covered by grass, like soccer fields. Second, for simplicity's sake, green areas and non-green areas were not always distinguished in this thesis, although this could have been possible using the available data. Specifically, in this thesis, all airstrips were treated as urban areas in every delineation method, regardless of their landcover. These two factors could have resulted in even greater disparities in urban sprawl calculations than the ones presented in this thesis. Additionally, the buffer distances used in some calculations, like the buffer distance to delineate the gardens, are an approximation that does not reflect reality in all cases. Larger gardens undoubtedly exist and there are also buildings where it is questionable if something like a "garden" is even present, as can be seen in Figure 44.

Figure 44: Buildings in the municipality of Grindelwald. Although for this thesis, gardens were assigned to these buildings, there is no noticeable area used as garden. (Image source: Environmental Systems Research Institute (2022))

## 5.5.8 Urban Area Delineation

The use of the $\alpha$-shape algorithm to define urban areas is not a widely adopted method in the literature on urban sprawl. In fact, no study researching urban sprawl has been found to rely on the $\alpha$-shape algorithm to delineate the urban area. The use of the $\alpha$-shape algorithm for delineation, despite its advantages, has also several shortcomings and has produced some results that are open to criticism, which will be presented more in detail in the following sections.

### 5.5.8.1 Spikes

The use of the $\alpha$-shape algorithm to outline urban areas resulted in certain shapes that may not be present in a manual delineation based on visual inspection. The most notable among these cases is the occurrence of “spikes” within the delimited urban area. One such spike can be seen in Figure 45. These spikes can arise from a point in the $\alpha$-shape algorithm having four connections. The presence of these spikes can not only affect the urban sprawl calculation directly due to their occupied area, but in some cases, they can also isolate green areas within the urban area from those outside it, as demonstrated in Figure 45. This results in a larger urban area in the delineations of the “Area with strong anthropogenic influence” and the “Non-rural area”, where the holes within the urban area are filled. The presence of spikes could have been mitigated by using a negative buffer followed by a positive buffer, but this approach was not pursued after trials revealed unintended consequences on other parts of the urban area. One could also argue that the spikes should not be removed since they may be considered a normal feature of the urban area delineated in this thesis, consistent with the approach and parameters utilized for the delineation of the other regions.

Figure 45: “Spike” of urban area. In this case, the spike forms a hole, which would be filled in the delineations of “Area with strong anthropogenic influence” and “Non-rural area”. (Source of the background image: Environmental Systems Research Institute (2022))

#### 5.5.8.2 Random Inclusion of Areas

As discussed in Chapter 3 of this thesis, there were instances where it was necessary to remove certain connections created by the $\alpha$-shape algorithm. During this deletion process, certain areas may or may not be included in the urban area by chance. This can be seen in Figure 18 in Section 3.5 of this thesis. In such cases, there can be inconsistencies in the urban area delineation, which is unfavorable. Fortunately, such occurrences were rare in the urban area delineation for this thesis, partly due to the pre-processing of the points prior to running the $\alpha$-shape algorithm.

#### 5.5.8.3 Manual Correction

In one case, namely the delineation of Celerina for the year 2021, a manual correction was required, as the code was unable to handle the intricate shape that was formed between two buildings in the village. The problem was caused by a shape similar to scenario "c)" in Figure 18 in Section 3.5 of this thesis.

#### 5.5.8.4 Paved Areas Inside Settlements

Certain paved areas fully surrounded by urban area were considered to be non-urban and therefore treated like greenery. This error occurred because certain asphalted areas inside the settlements were not delineated in any of the classes of the used data. Thus, certain holes inside the delineation method of "Built-up area" are not green areas, but paved ones.

### 5.5.9 Railway Tracks Between and Alongside Urban Areas

Although areas characterized by parallel railway tracks were classified as urban areas, as long as they lied in holes of urban areas, no easy way was found to incorporate such areas that were not inside the holes of urban areas. This leads to the effect that in Winterthur, such an area lying between buildings was not included into the urban area. Therefore, the delineation fails to include certain areas characterized by parallel railway tracks in an urban context, as can be seen in Figure 46. Thus, the urban sprawl value is influenced by the absence of this portion of the urban area. However, as this issue is ubiquitous in all urban area delineations in this thesis, this does not affect the differences in urban sprawl delineations. In theory, this issue could be solved by applying a positive buffer and a subsequent negative buffer. However, this idea was not implemented in the final delineation, as trials showed unwanted effects on different parts of the urban area and its holes.

Figure 46: Parallel railway tracks in an urban context: At the center, an area in Winterthur is visible that is covered by several railway tracks. Although this area is located between buildings, this area could not be included into the urban area. (Source of the background image: Environmental Systems Research Institute (2022))

### 5.5.10 Roads Between and Alongside Urban Areas

In the final delineation of the urban areas, no correction was made to include roads into the urban areas. This had several reasons, but also led to some suboptimal results. For example, some urban areas, like the one in Figure 47, were divided by the presence of roads. A reason to ignore the road system was the fact, that the $\alpha$-shape distance was in principle already designed to include roads. However, the chosen radius is not sufficient for the inclusion of wide roads like highways. Additionally, roads at the edge of the delineated urban area were not always within its boundaries. This is not that much of a problem due to the lack of strong opinions regarding the inclusion of such areas in urban sprawl literature. Aside from theoretical considerations, practically speaking, no straightforward method was found to include roads into the urban area in a sensible way. Attempts to incorporate road data resulted in issues like the inclusion of excessive road area, the alteration of the urban area's overall shape, or the formation of narrow, non-urban strips due to vegetation such as grass verges along the roads.

Figure 47: Excerpt of the urban area delineation of Winterthur. One can observe that the wide road area, despite being situated between buildings on either side, was not considered a part of the urban area. (Source of the background image: Environmental Systems Research Institute (2022))

### 5.5.11 Hole Filling

In the delineation methods of "Area with strong anthropogenic influence" and "Non-rural area" all holes are filled, including large ones. This would mean that when a group of buildings form a large indentation based on the data of the year 2015 and using data for the year 2021 this indentation becomes fully encircled by buildings, the urban area substantially increases. This can be seen in Figure 48. This rapid rise in urban area does not seem sensible, as a small number of buildings or even one single building can finalize this process. This rapid rise in urban area clashes with the point "monotonous reaction to increases in urban area" of the 13 indicators proposed by Jaeger, Bertiller, Schwick, Cavens, et al. (2010), as it may lead to an abrupt rise in the utilized urban sprawl indicator. Additionally, it is questionable if it is compatible with the proposed criteria to select data for environmental indicators by Niemeijer and de Groot (2008), as it can be debated if this sudden rise in urban area meets the requirement of being predictable. In urban sprawl literature, this issue is sometimes addressed by excluding large green areas from the urban area even if they are completely encircled by it (whereas the definition of what constitutes "large" varies). However, this issue persists on a smaller scale, and most notably, the potential urban character of green areas within urban areas is disregarded. This may make sense for a delineation that focuses solely on the built-up areas but can be problematic otherwise. A potential solution to mitigate such rapid changes in urban area would be to already include the areas in indentations into the urban area. An approach leading to such a delineation could be a similar one to the convex-hull approach. However, this would lead to the inclusion of potentially large swaths of green area on the urban fringe, which is not considered "urban" in the general understanding of urban areas in the context of urban sprawl. The best approach could be to utilize data on green areas that are being utilized by humans in an urban context. One the one hand, this would effectively tackle the addressed issue. One the other hand, when human use of green areas increases (or decreases) gradually, a binary classification of urban versus non-urban may again result in a sudden increase (or decrease) in urban area.

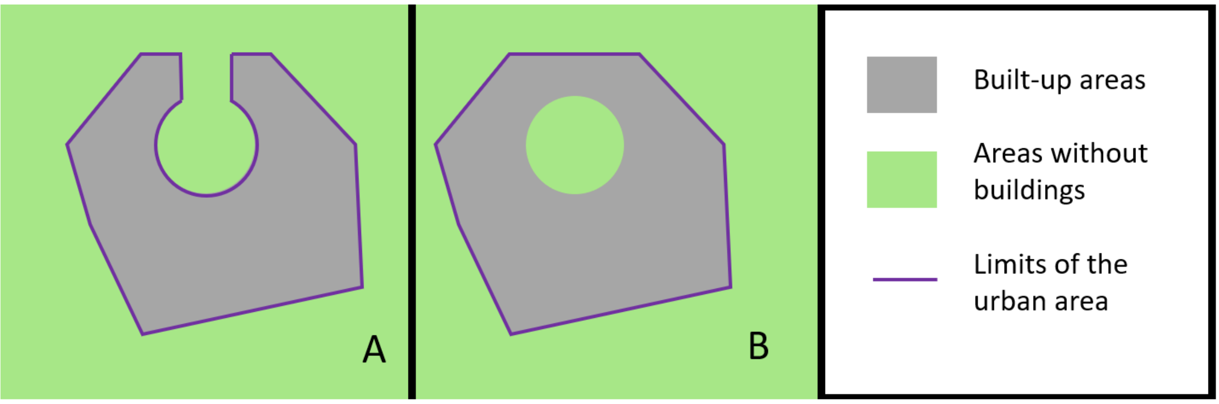



Figure 48: Change in urban area between two years. "A" shows an incompletely surrounded hole (i.e., an indentation), whereas in "B" the hole is completely surrounded by urban area. In the delineations of "Area with strong anthropogenic influence" and "Non-rural area" this would lead to an urban area delineation as shown in the figure.

## 5.5.12 Rasterization and Variations Therein

Differences in the delineations of Moran's I and WUP are probably also influenced by the different choices of the rasterization method. Contrary to the rasterization applied to the Moran's I delineation that categorized cells based on the largest area within the cell, being either urban or non-urban area, the raster used for the calculation of WUP was constructed differently. In the latter case, the software FME (Safe Software, 2023b) was used by Yves Maurer to rasterize the urban area (Y. Maurer, personal communication, January 24, 2023). For this, the "NumericRasterizer" function was utilized, as according to Yves Maurer's experience, this approach creates a fast and accurate rasterized image of urban areas. The tolerance value of 0.4 was chosen, as according to Yves Maurer's expertise this value leads to a particularly precise depiction of the urban form (Y. Maurer, personal communication, January 24, 2023). This tolerance parameter means that cells were classified as part of the urban area if they contained urban areas in vector format within a 40% distance from the center point of the pixel to the border of the pixel (Safe Software, 2023a).

It is important to note that the process of rasterization itself can result in changes in the urban area delineation. This not only relates to the urban area size but also to the connectivity of the form, for example.

Despite its drawbacks, a rasterization is necessary for the calculation of WUP and Moran's I. Distinct methods of rasterization finally also have effects on the sprawl result. For instance, the rasterization for WUP was found to contain more urban pixels than the one used for Moran's I. This leads to a larger area and affects the urban sprawl values this way. At the same time, the rasterization utilized for WUP helps to preserve connectivity at certain locations. However, in theory, the rasterization process for WUP could also create connections where none existed in the original vector data. Conveniently, due to the large distance between disconnected urban areas in comparison to the pixel size, this potential effect is likely not observed in the delineations in this thesis.

However, the degree to which the retention of connectivity is a positive aspect is up to debate: On the one hand, it preserved the original urban form, but on the other hand, this also applies to cases like spikes which may or may not truly be considered parts of the urban area. The preservation of connectivity or lack thereof again can have a significant impact on urban sprawl metrics, for example on measurements that quantify phenomena such as clustering.

Ultimately, it is difficult to state which rasterization method is most suitable for the task. As within the measurement method all variants of urban area delineation were rasterized in the same manner, the choice of rasterization method does not influence the comparison of sprawl values between different delineation methods. However, regardless of the choice of the rasterization method, changes in the urban area will occur and potentially amplify or diminish the differences present in the different urban area delineations.

### 5.5.13 Effects of Raster Size on Spatial Autocorrelation

As already discussed in Chapter 3, using small cell sizes for the rasterization has several advantages. A small cell size, however, does not only increase computation time for the calculation of the urban sprawl, which in this case was still very acceptable, but also affects the Moran's I value, irrespective of the mentioned concern of quality loss. The reason for this is the following: Moran's I, as already presented, measures spatial autocorrelation. For this, the values of nearby cells are compared to the cell currently analyzed. When employing smaller cell sizes, the immediate neighbors of a cell are more likely to be of the same value. This effect can be seen in Figure 49. This relationship is also visible using the data on the urban areas of the five municipalities, which show high variations in their Moran's I values, employing different cell sizes.

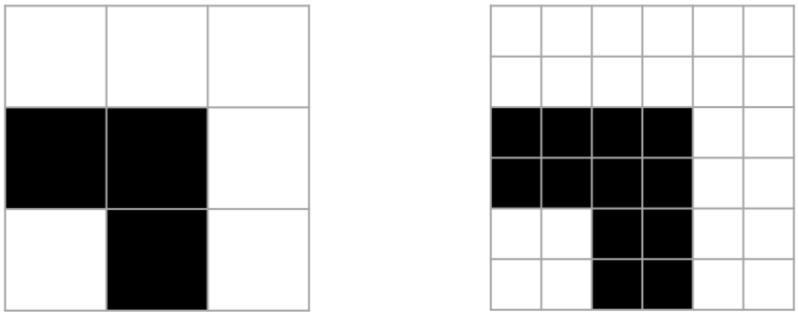


Figure 49: Increase of spatial autocorrelation due to different cell sizes. Black and white cells stand for different values, for example 0 and 1. On the right, where the cell size is smaller, certain cells, like for example the whole first row, do not touch any cell of the other value. This is not the case on the left. The autocorrelation will be larger for the right image, as cells with the same value touch each other more frequently.

## 5.6 Comparison With Existing Research

Given the limited research with a similar scope as this thesis, it is difficult to state how the results of this study fit into the existing literature. In addition to general comments on the impact of delineation methods on urban sprawl without further clarification (Schwick et al., 2018), a study by García-Álvarez and Camacho Olmedo (2017), discussed in Chapter 2 of this thesis, provides a basis for comparing the effects of different urban area delineation methods on the calculated urban sprawl. Despite its discussed limitations, the findings of García-Álvarez and Camacho Olmedo (2017) align with the results of this thesis: The way an urban area is delineated matters. Both the research by García-Álvarez and Camacho Olmedo (2017) and the findings presented in this thesis demonstrate that the use of different urban area delineations can result in varying measurements of sprawl. The ongoing debate in the field of urban sprawl literature regarding the validity of various calculation methods is further reinforced by the results of this thesis: The use of different methods indeed results in varying urban sprawl values. This thesis also highlights specific discrepancies between different urban sprawl calculation methods, for example highlighting the inconsistencies between the results of the density calculation and those of Moran's I, which were also observed by Zhou et al. (2019).

## 5.7 General Assessment of the Methodology

Despite the limitations and considerations raised regarding the methods employed in this thesis, the results of the delineation and sprawl analysis can be considered useful in addressing the research question. The different employed delineations map the urban area based on different paradigms of urban areas in urban sprawl research. As the delineation is very detailed, filigree phenomena, like ribbon-development, are less likely to be overlooked, as could happen when delimiting the urban area at smaller scales. This is particularly crucial, as such developments are one of the main characteristics of urban sprawl (Altshuler et al., 1993). The utilized sprawl indicators, namely, density, Moran's I, and WUP, represent state of the art indicators – or at least dimensions of it – which are especially apt to be used in urban sprawl measurement, as described in Chapter 2 of this thesis. Density, a very important, if not the most important sprawl indicator (Lopez & Hynes, 2003; Torrens, 2008), Moran's I, a widely employed measure that quantifies clustering and the second most important sprawl measure (Zhou et al., 2019), and WUP, which takes into account various characteristics of sprawl and is recognized for its conformity to a high number of suitability criteria (Behnisch et al., 2022; Jaeger, Bertiller, Schwick, & Kienast, 2010) and also otherwise has been demonstrated to measure sprawl in an especially pleasing manner (Behnisch et al., 2022), were used in this thesis to calculate sprawl. Moreover, the components of WUP were analyzed for differences arising from the delineation method in order to reveal that multiple components, and not just one, are affected by changes in delineation and contribute to the varying results of WUP. In addition, it can be determined which of these components are particularly impacted by changes in the delineation method. The selected municipalities contain settlement areas of different sizes and with different urban forms, aiming to provide results on a diverse assortment of urban forms. As the three delineation methods used the same parameters and followed the same rules in all municipalities, issues arising from urban areas that were delineated at varying scales were avoided.

## 5.8 Discussion of the Findings

The differences between the three urban area delineation methods in this thesis were largely dependent on the method used to quantify urban sprawl. This is not only the case for the sprawl values of the years 2015 and 2021 but also for the difference between these two years. The importance of the observed variations between the different urban area delineation methods for research on urban sprawl is uncertain as there is no agreement on the meaning of specific sprawl values. Nevertheless, the fact that there are cases in which the sprawl values based on one delineation method are a multiple of the one of another delineation method props up the assertion, that striking differences were found. Apart from this, the delineation method also could lead to differences in the direction the detected sprawl was heading between the years. Furthermore, even minor differences found in the sprawl measurements arising from different delineations could translate to extreme differences in urban sprawl. This is due to the fact that some urban sprawl values that could theoretically be obtained through the urban sprawl measurements are not achievable in real-world conditions. For instance, in the realm of urban sprawl, urban development ideally exhibits an infinite density, while in the worst-case scenario, it possesses a density close to zero. Similarly, a development with an optimal Moran's I is perfectly clustered (i.e., is an area with only urban or only non-urban area) and in the worst-case scenario exhibits a checkerboard pattern. However, urban developments will in reality always fall between these two extremes. Especially a municipality featuring a perfect checkerboard pattern of 15x15 m cells of urban and non-urban area is far removed from the patterns observable in real-life settlements. Therefore, as in the real world the most sprawled municipality will still have lower sprawl values than the largest theoretical sprawl value, and the same, in reverse, holds true for the most compact municipality, the differences between the three distinct delineation methods in this thesis are underestimated. Thus, the actual sprawl value of the most sprawling municipality will always be lower than the theoretical maximum, while the sprawl value of the most compact municipality will be higher than the theoretical minimum. Therefore, the disparities between the sprawl values stemming from the three distinct urban area delineation methods in this thesis may be way more substantial as they seem based solely on the difference of their measured values.

Interestingly, in certain municipalities, considerable differences in sprawl values could be observed between the delineations of "Areas with strong anthropogenic influence" and "Non-rural area" using the WUP calculation method. This is the case, although this calculation method is reported to be largely insensitive to very small patches of urban area, like single buildings (Jaeger, Bertiller, Schwick, & Kienast, 2010). This demonstrates that even measurements that are reported to handle differences in delineation well have their limits.

Disparities between the different delineation methods were noted in every municipality. Depending on the calculation method, some municipalities presented larger differences in values due to the variation in the applied urban area definition. However, large differences were not confined to a certain urban form. For example, using some calculations, Grindelwald, a touristic village in the Alps with a multitude of scattered buildings shows great variations in sprawl values stemming from the different delineation methods whereas in others, Wetzikon or Winterthur, which are a town and a big city, respectively, are the ones exhibiting the large variations in sprawl values. This is to show that the found differences in calculated sprawl stemming from the employed urban area definitions is not confined to a certain settlement type, but is rather an overarching issue affecting a whole range of contemporary urban forms.

## 5.9 Implications for Future Research

This thesis highlights the considerable differences that exist in urban sprawl measurements when different concepts are employed to delineate urban areas. This applies for urban sprawl as an absolute phenomenon or as a change over time. However, it is not possible to definitively conclude that the choice of delineation method leads to a meaningful impact on the assessment of sprawl, as there are no established benchmarks for when sprawl begins or ends. If the impact of the urban area delineation on urban sprawl measurements can be viewed as meaningful, one can hope than in future, more awareness emerges regarding the appropriate way to delineate urban areas for accurate urban sprawl measurements. This includes the critical discussion of questions about what constitutes "urban", the type of data that should be used to delineate urban areas and the standards that should be followed by existing maps if they are to be used as the foundation of urban sprawl research. By addressing these issues, it is possible to not only enhance the accuracy of urban sprawl measurements, but also to adopt consistent standards for urban area delineation which could promote fair regulations and foster a more uniform understanding of urban sprawl among researchers in the field. In literature, the vast array of proposed calculation methods for measuring urban sprawl seems immeasurable, whereas the principles for defining the urban area are often overlooked. It would be unfortunate if the progress in finding the most accurate method for measuring urban sprawl were hindered due to the lack of standardization in the urban area delineations used for the calculations.

Given the potential significance of the findings in this thesis, it is recommended that they are validated through further research. This also encompasses addressing the various limitations of the adopted methodology. Additionally, other regions could be analyzed using different or improved algorithms to delineate urban areas and calculation methods to quantify urban sprawl to revise the results of this thesis.

The approach utilized in this thesis for urban area delineation, which is based on the $\alpha$-shape algorithm, in spite of its downsides, proved to be a powerful algorithm to delineate urban areas. This algorithm can mitigate or significantly reduce the problems associated with subjective delineation and replace the time-consuming process of visual inspections if the required data to run it are available. The data used for the purpose of delimiting urban areas in urban sprawl research do not have to be limited to vector data from TLM3D, but can also include other vector data on buildings and areas. Such data could be for example extracted from OpenStreetMap (Geofabrik & OpenStreetMap Contributors, 2023) or the Global ML Building Footprints-dataset (Microsoft, 2022), which provides AI-generated building footprints. In this way, urban area delineations for urban sprawl measurement could be carried out for other locations, and similar analyses to the one in this thesis can be performed.

# Chapter 6 | Conclusion

This study analyzes if the delineation method used for urban areas has an impact on the measured urban sprawl values – a largely overlooked aspect in previous urban sprawl research. To do so, the existing heterogeneous approaches used in the literature for defining the urban area had to be studied and clustered, in order to identify the typical methods for urban area delineation. The three identified archetypical urban area definitions were then compared to each other calculating the urban sprawl resulting from their delineation. To achieve this, a delineation based on the $\alpha$-shape algorithm was implemented, aiming for an automatic delimitation of urban areas, intended to reduce error-prone human judgement when delineating the urban area to a high degree. This new method to delineate urban areas for urban sprawl calculation produces maps with a high spatial resolution on five distinct Swiss municipalities, ranging from alpine villages to a big city in the Swiss plateau. The differences between the constructed urban areas were assessed using three calculation methods for urban sprawl: density, Moran's I, and WUP. This study highlights the significance of carefully defining the urban area used for calculating urban sprawl, as it shows that variations in the calculated urban sprawl measurements result from the use of distinct urban area delineations. These variations can be massive, leading to great inconsistencies in the way urban sprawl is ultimately measured. Having said this, the results of this thesis urge for a more thoughtful consideration of how urban areas should be delineated for urban sprawl research. Doing so could improve the comparability of urban sprawl calculations and help to achieve more consistent and fairer urban sprawl legislation.

# Literature


Altshuler, A. A., Gómez-Ibáñez, J. A., & Howitt, A. M. (1993). *Regulation for revenue: The political economy of land use exactions*. The Brookings Institution; The Lincoln Institute of Land Policy. http://www.worldcat.org/oclc/604752082

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensor data* (Geological Survey Professional Paper 964). USGS Publications Warehouse. https://doi.org/10.3133/pp964

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. (2018). *Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens, GeoInfoDok* (ATKIS-Objektartenkatalog Basis-DLM Version 7.1 rc.1; ATKIS-Katalogwerke, pp. 1–408). https://www.adv-online.de/icc/extdeu/nav/a63/binarywriterservlet?imgUid=9201016e-7efa-8461-e336-b6951fa2e0c9&uBasVariant=11111111-1111-1111-1111-111111111111

Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. (2021). *Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens, GeoInfoDok* (Anwendungsschema ATKIS Version 6.0.1; Erläuterungen zum ATKIS® Basis-DLM, pp. 1–174). http://www.adv-online.de/GeoInfoDok/GeoInfoDok-6.0/binarywriterservlet?imgUid=34f23fbd-a20a-771d-478d-9d43b36c4c2e&uBasVariant=11111111-1111-1111-1111-111111111111

Arribas-Bel, D., Garcia-López, M.-À., & Viladecans-Marsal, E. (2021). Building(s and) cities: Delineating urban areas with a machine learning algorithm. *Journal of Urban Economics*, *125*, Article 103217. https://doi.org/10.1016/j.jue.2019.103217

Arribas-Bel, D., & Wolf, L. J. (2022). *Alpha_shapes.py* [Python]. https://github.com/pysal/libpysal/blob/969039d32069a15e1676524e8d2dd76780a9ff56/libpysal/cg/alpha_shapes.py

Barber, C. B., The Geometry Center, & Laurent, S. (2022). *Package 'cxhull'* (0.7.0) [R]. https://github.com/stla/cxhull

Baur, I., Liechti, K., & Binder, C. R. (2014). Why do individuals behave differently in commons dilemmas? The case of alpine farmers using common property pastures in Grindelwald, Switzerland. *International Journal of the Commons*, *8*(2), 657–685. https://doi.org/10.18352/ijc.469

Behnisch, M., Krüger, T., & Jaeger, J. A. G. (2022). Rapid rise in urban sprawl: Global hotspots and trends since 1990. *PLOS Sustainability and Transformation*, *1*(11), Article e0000034. https://doi.org/10.1371/journal.pstr.0000034

Bellock, K. E. (2021, April 16). *Alphashape 1.3.1*. Alpha Shape Toolbox. https://pypi.org/project/alphashape/

Böck, S., Immitzer, M., & Atzberger, C. (2017). On the objectivity of the objective function—Problems with unsupervised segmentation evaluation based on global score and a possible remedy. *Remote Sensing*, *9*(8), Article 769. https://doi.org/10.3390/rs9080769

Bundesamt für Landestopografie. (2022, November 2). *SwissTLM3D*. Schweizerische Eidgenossenschaft. https://www.swisstopo.admin.ch/de/geodata/landscape/tlm3d.html

Bundesamt für Landestopografie. (2023, January 12). *SwissTLM3D Strassen und Wege*. EEA geospatial data catalogue. https://sdi.eea.europa.eu/catalogue/fise/api/records/be285f33-0373-4882-b911-de402bcdfeb4

Bundesamt für Raumentwicklung, & Fahrländer Partner AG. (2008). *Bauzonen Schweiz: Wie viele Bauzonen braucht die Schweiz?* [Schlussbericht Projekt 30008]. Schweizerische Eidgenossenschaft. https://www.are.admin.ch/dam/are/de/dokumente/grundlagen/publikationen/wie-viele-bauzonen-braucht-die-schweiz.pdf

Bundesamt für Statistik. (2006). *Swiss land use statistics – Standard nomenclature NOAS04: Basic categories and aggregations* (p. 1). Schweizerische Eidgenossenschaft, Eidgenössisches Departement des Innern. https://dam-api.bfs.admin.ch/hub/api/dam/assets/171190/master

Bundesamt für Statistik. (2018). *Arealstatistik Nomenklatur, Standardnomenklatur NOAS04* (Standardkategorien, pp. 1–154). Schweizerische Eidgenossenschaft, Eidgenössisches Departement des Innern EDI. https://dam-api.bfs.admin.ch/hub/api/dam/assets/6948898/master

Bundesamt für Statistik. (2022a). *Stadtent* [Data set]. [Unpublished data]

Bundesamt für Statistik. (2022b). *Stadtpop* [Data set]. [Unpublished data]

Bundesamt für Statistik. (2021, November 25). *Arealstatistik nach Nomenklatur 2004, Erhebungen 1979-1985, 1992-1997, 2004-2009, 2013-2018*. Schweizerische Eidgenossenschaft. https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/boden-nutzung-bedeckung-eignung/arealstatistik-schweiz.assetdetail.20104753.html

Bundesamt für Statistik. (2022c, October 6). *Durchschnittliche Wohnfläche pro Bewohner nach Gebäudekategorie und Bauperiode*. Schweizerische Eidgenossenschaft. https://www.bfs.admin.ch/bfs/de/home/statistiken/bau-wohnungswesen/wohnungen/wohnverhaeltnisse/flaechenverbrauch.assetdetail.23524462.html

Bundesamt für Statistik. (2022d, October 6). *Gebäude nach Gebäudekategorie sowie Bauperiode und Geschosszahl*. Schweizerische Eidgenossenschaft. https://www.bfs.admin.ch/bfs/de/home/statistiken/bau-wohnungswesen/gebaeude/kategorie.assetdetail.23524379.html

Bundesamt für Statistik. (2022e, November 2). *Belegungsdichte*. Schweizerische Eidgenossenschaft. https://www.bfs.admin.ch/bfs/de/home/statistiken/bau-wohnungswesen/wohnungen/wohnverhaeltnisse/belegungsdichte.html

Bundesamt für Statistik. (2022f, November 10). *Demografische Bilanz nach institutionellen Gliederungen* [STAT-TAB – interaktive Tabellen (BFS)]. Schweizerische Eidgenossenschaft. http://www.pxweb.bfs.admin.ch/pxweb/de/px-x-0102020000_201/px-x-0102020000_201/px-x-0102020000_201.px/

Bundesamt für Statistik, Sektion GEO. (2007). *Siedlungsgebiet der Schweiz 2006, Polygondatensatz aus Daten des Gebäude- und Wohnungsregisters (GWR) sowie der Arealstatistik* (Version 1.0; Siedlungsgebiet Schweiz, Verfahren zur Erstellung des GIS-Datensatzes, pp. 1–19). Schweizerische Eidgenossenschaft, Eidgenössisches Departement des Innern EDI. [Unpublished manuscript]

Bundesamt für Verkehr. (2020). Ausführungsbestimmungen zur Eisenbahnverordnung. *AB zur EBV (SR 742.141.11)*. https://www.bav.admin.ch/bav/de/home/rechtliches/rechtsgrundlagen-vorschriften/ab-ebv.html

Copernicus. (2020). *Mapping guide v6.2 for a European urban atlas*. European Commission. https://land.copernicus.eu/user-corner/technical-library/urban_atlas_2012_2018_mapping_guide

Copernicus. (2021). *CORINE land cover product user manual* (Version 1.0; Copernicus Land Monitoring Service). European Environment Agency. https://land.copernicus.eu/user-corner/technical-library/clc-product-user-manual

Direção-Geral do Território. (2019). *Especificações técnicas da carta de uso e ocupação do solo (COS) de Portugal continental para 2018: Relatório técnico*. Direção-Geral do Território. https://www.dgterritorio.gov.pt/sites/default/files/documentos-publicos/2019-12-26-11-47-32-0__ET-COS-2018_v1.pdf

Environment Agency Austria. (2019). *Updated CLC illustrated nomenclature guidelines* (p. 126). https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/docs/pdf/CLC2018_Nomenclature_illustrated_guide_20190510.pdf

Environmental Systems Research Institute. (2022). *World imagery (for export)* [Map]. https://www.arcgis.com/home/item.html?id=226d23f076da478bba4589e7eae95952

Equipo Técnico Nacional SIOSE. (2022). *Manual de fotointerpretación: Versión 3.9* (pp. 1–65) [Sistema de Información de Ocupación del Suelo en España de Alta Resolución, Cartografía Digital]. Instituto Geográfico Nacional. https://www.siose.es/SIOSEtheme-theme/documentos/pdf/Manual_Fotointerpretacion_SIOSE_AR_v3.9.pdf

European Environment Agency. (2011). *Updated UMZs and corresponding methodological documentation: Method based on CLC v16* (pp. 1–48) [Task: 261-1#2.3_1: Further development of IUME and urban assessment]. http://cmshare.eea.europa.eu/s/Qti5GtdEpL2FRyi/download

European Environment Agency & Swiss Federal Office for the Environment. (2016). *Urban sprawl in Europe: Joint EEA-FOEN report* (No. 11). Publications Office. https://data.europa.eu/doi/10.2800/143470

Ewing, R., & Hamidi, S. (2014). *Measuring urban sprawl and validating sprawl measures* (p. 158). Metropolitan Research Center. https://gis.cancer.gov/tools/urban-sprawl/sprawl-report-short.pdf

Ewing, R., Pendall, R., & Chen, D. (2003). Measuring sprawl and its transportation impacts. *Transportation Research Record: Journal of the Transportation Research Board*, *1831*(1), 175–183. https://doi.org/10.3141/1831-20

Fabian Gattlen. (2023). *Celerina/Schlarigna* [Photograph]. https://vcdn.bergfex.at/images/resized/39/6159696800e91439_6f418d40c0ee2510@2x.jpg

Fulton, W., Pendall, R., Nguyen, M., & Harrison, A. (2001). *Who sprawls most? How growth patterns differ across the U.S.* (The Brookings Institution Survey Series, p. 23). The Brookings Institution. https://www.brookings.edu/research/who-sprawls-most-how-growth-patterns-differ-across-the-u-s/

Galster, G., Hanson, R., Ratcliffe, M. R., Wolman, H., Coleman, S., & Freihage, J. (2001). Wrestling sprawl to the ground: Defining and measuring an elusive concept. *Housing Policy Debate*, *12*(4), 681–717. https://doi.org/10.1080/10511482.2001.9521426

García-Álvarez, D., & Camacho Olmedo, M. T. (2017). Changes in the methodology used in the production of the Spanish CORINE: Uncertainty analysis of the new maps. *International Journal of Applied Earth Observation and Geoinformation*, *63*, 55–67. https://doi.org/10.1016/j.jag.2017.07.001

Gemeinde Ins. (2023). *Ins* [Photograph]. https://www.ins.ch/wAssets/img/ins/portrait/086489.jpg

Gemeindeverwaltung Celerina/Schlarigna. (2023a, January 13). *Celerina*. Leben in Celerina. https://www.gemeinde-celerina.ch/celerina#

Gemeindeverwaltung Celerina/Schlarigna. (2023b, January 13). *Kultur & Sprache*. Gemeinde Celerina. https://www.gemeinde-celerina.ch/leben-in-celerina/kultur-sprache#

Gemeindeverwaltung Grindelwald. (2023, January 13). *Zahlen und Fakten*. Einwohnergemeinde Grindelwald. https://www.gemeinde-grindelwald.ch/zahlen-fakten/

Gemeindeverwaltung Ins. (2023, January 13). *Zahlen und Fakten*. Gemeinde Ins. https://www.ins.ch/de/portrait/zahlen-und-fakten/

Geofabrik, & OpenStreetMap Contributors. (2022). *Switzerland-latest-free.shp.zip* [Map]. https://download.geofabrik.de/europe/switzerland.html

Geofabrik, & OpenStreetMap Contributors. (2023). *Geofabrik download server*. https://download.geofabrik.de/index.html

Geoportal of the Confederation. (2023, January 14). *Map.geo.admin.ch*. Swiss Confederation. https://map.geo.admin.ch

Gilbert, T., & Stys, B. (2004). *Descriptions of vegetation and land cover types mapped using Landsat imagery* (p. 16). Florida Fish and Wildlife Conservation Commission. https://myfwc.com/media/12977/veglulc2003desc.pdf

Glaeser, E. L., Kahn, M., & Chu, C. (2001). *Job sprawl: Employment location in U.S. metropolitan areas* (The Brookings Institution Survey Series, pp. 1–8). Center on Urban & Metropolitan Policy. https://www.brookings.edu/research/job-sprawl-employment-location-in-u-s-metropolitan-areas/

Glaeser, E. L., Kahn, M. E., Arnott, R., & Mayer, C. (2001). Decentralized employment and the transformation of the American city [with Comments]. *Brookings-Wharton Papers on Urban Affairs*, 1–63.

Gombin, J., Vaidyanathan, R., Agafonkin, V., & Mapbox. (2022). *Package 'concaveman'* (1.1.0) [R]. https://joelgombin.github.io/concaveman/

Google Earth. (2021). *Winterthur* [Map].

Google Earth. (2023). *Marguns* [Map].

Hahsler, M., Piekenbrock, M., & Doran, D. (2019). dbscan: Fast density-based clustering with R. *Journal of Statistical Software*, *91*(1), 1–30. https://doi.org/10.18637/jss.v091.i01

Jaeger, J. A. G., Bertiller, R., Schwick, C., Cavens, D., & Kienast, F. (2010). Urban permeation of landscapes and sprawl per capita: New measures of urban sprawl. *Ecological Indicators*, *10*(2), 427–441. https://doi.org/10.1016/j.ecolind.2009.07.010

Jaeger, J. A. G., Bertiller, R., Schwick, C., & Kienast, F. (2010). Suitability criteria for measures of urban sprawl. *Ecological Indicators*, *10*(2), 397–406. https://doi.org/10.1016/j.ecolind.2009.07.007

Jaeger, J. A. G., & Schwick, C. (2014). Improving the measurement of urban sprawl: Weighted Urban Proliferation (WUP) and its application to Switzerland. *Ecological Indicators*, *38*, 294–308. https://doi.org/10.1016/j.ecolind.2013.11.022

Jiang, S., & Liu, D. (2012). Box-counting dimension of fractal urban form: Stability issues and measurement design. *International Journal of Artificial Life Research (IJALR)*, *3*(3), 41–63. https://doi.org/10.4018/jalr.2012070104

Kalinina, D., Ingilevich, V., Lantseva, A., & Ivanov, S. (2018). Computing concave hull with closed curve smoothing: Performance, concaveness measure and applications. *Procedia Computer Science*, *136*, 479–488. https://doi.org/10.1016/j.procs.2018.08.258

Kawula, R., & Redner, J. (2018). *Florida land cover classification system* (p. 72). Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute. https://myfwc.com/research/gis/applications/articles/fl-land-cover-classification/

Knight, G. R., Knight, A., Hipes, D., NeSmith, K., Gulledge, K., Jenkins, A., Elam, C., Diamond, P., Oetting, J., & Newberry, A. (2010). *Development of a cooperative land cover map: Final report* (p. 108). Florida Natural Areas Inventory and Florida Fish and Wildlife Conservation Commission. https://www.nrc.gov/docs/ML1433/ML14336A379.pdf

Lagarias, A., & Sayas, J. (2018). Urban sprawl in the Mediterranean: Evidence from coastal medium-sized cities. *Regional Science Inquiry*, *10*(3), 15–32.

Lopez, R., & Hynes, H. P. (2003). Sprawl in the 1990s: Measurement, distribution, and trends. *Urban Affairs Review*, *38*(3), 325–355. https://doi.org/10.1177/1078087402238805

Manesha, E. P. P., Jayasinghe, A., & Kalpana, H. N. (2021). Measuring urban sprawl of small and medium towns using GIS and remote sensing techniques: A case study of Sri Lanka. *The Egyptian Journal of Remote Sensing and Space Sciences*, *24*(3), 1051–1060. https://doi.org/10.1016/j.ejrs.2021.11.001

Mayhew, S. (2015). Urban sprawl. In *A Dictionary of Geography* (5th ed.). Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780199680856.001.0001/acref-9780199680856-e-3236

Microsoft. (2022). *GlobalMLBuildingFootprints*. https://github.com/microsoft/GlobalMLBuildingFootprints

Mosammam, H. M., Nia, J. T., Khani, H., Teymouri, A., & Kazemi, M. (2017). Monitoring land use change and measuring urban sprawl based on its spatial forms. *The Egyptian Journal of Remote Sensing and Space Sciences*, *20*(1), 103–116. https://doi.org/10.1016/j.ejrs.2016.08.002

Musakwa, W., & Van Niekerk, A. (2013). Implications of land use change for the sustainability of urban areas: A case study of Stellenbosch, South Africa. *Cities*, *32*, 143–156. https://doi.org/10.1016/j.cities.2013.01.004

Musakwa, W., & Van Niekerk, A. (2014). Monitoring urban sprawl and sustainable urban development using the Moran index: A case study of Stellenbosch, South Africa. *International Journal of Applied Geospatial Research*, *5*(3), 1–20. https://doi.org/10.4018/ijagr.2014070101

Nazarnia, N., Harding, C., & Jaeger, J. A. G. (2019). How suitable is entropy as a measure of urban sprawl? *Landscape and Urban Planning*, *184*, 32–43. https://doi.org/10.1016/j.landurbplan.2018.09.025

Nazarnia, N., Schwick, C., & Jaeger, J. A. G. (2016). Accelerated urban sprawl in Montreal, Quebec City, and Zurich: Investigating the differences using time series 1951–2011. *Ecological Indicators*, *60*, 1229–1251. https://doi.org/10.1016/j.ecolind.2015.09.020

Nelson, A. C. (1999). Comparing states with and without growth management: Analysis based on indicators with policy implications. *Land Use Policy*, *16*(2), 121–127. https://doi.org/10.1016/S0264-8377(99)00009-5

Niemeijer, D., & de Groot, R. S. (2008). A conceptual framework for selecting environmental indicator sets. *Ecological Indicators*, *8*(1), 14–25. https://doi.org/10.1016/j.ecolind.2006.11.012

Olazabal, E., & Bellet, C. (2017). Análisis de las nuevas dinámicas de urbanización en España: Su estudio a través del uso de Corine land cover y SIOSE. *Naturaleza, territorio y ciudad en un mundo global*, 2639–2648. https://www.researchgate.net/publication/321033603

OpenStreetMap Wiki contributors. (2022, October 22). *Key:landuse*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Key:landuse

Pateiro-López, B., & Rodríguez-Casal, A. (2010). Generalizing the convex hull of a sample: The R package alphahull. *Journal of Statistical Software*, *34*(5), 1–28. https://doi.org/10.18637/jss.v034.i05

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Petrescu, F. (2019). Urban sprawl from Urban Atlas data: Romanian case study. *Management Research and Practice*, *11*(2), 21–30.

Prastacos, P., & Lagarias, A. (2016). Sprawl in European urban areas. In K. Themistocleous, D. G. Hadjimitsis, S. Michaelides, & G. Papadavid (Eds.), *Fourth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2016)* (Vol. 9688). Article 968816 (2016). Paphos, Cyprus. https://doi.org/10.1117/12.2240734

Rastogi, K., & Jain, G. V. (2018). Urban sprawl analysis using Shannon's entropy and fractal analysis: A Case study on Tiruchirappalli City, India. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII–5*, 761–766. https://doi.org/10.5194/isprs-archives-XLII-5-761-2018

Redner, J., & Srinivasan, S. (2014). *Florida vegetation and land cover 2014, final report*. Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute. [Unpublished manuscript]

Rogers, A., Castree, N., & Kitchin, R. (2013). Urban sprawl. In *A Dictionary of Human Geography* (1st ed.). Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780199599868.001.0001/acref-9780199599868-e-1982

Rubiera Morollón, F., González Marroquin, V. M., & Pérez Rivero, J. L. (2015). Urban sprawl in Spain: Differences among cities and causes. *European Planning Studies*, *24*(1), 207–226. https://doi.org/10.1080/09654313.2015.1080230

Rubiera-Morollón, F., & Garrido-Yserte, R. (2020). Recent literature about urban sprawl: A renewed relevance of the phenomenon from the perspective of environmental sustainability. *Sustainability*, *12*(16), Article 6551. https://doi.org/10.3390/su12166551

Safe Software. (2023a, January 27). *NumericRasterizer*. https://docs.safe.com/fme/html/FME_Desktop_Documentation/FME_Transformers/Transformers/numericrasterizer.htm

Safe Software. (2023b, January 27). *Safe Software: FME—data integration platform* [Feature Manipulation Engine]. https://www.safe.com/

Salvati, L., & Carlucci, M. (2014). Urban growth and land-use structure in two Mediterranean regions: An exploratory spatial data analysis. *SAGE Open*, *4*(4), 1–13. https://doi.org/10.1177/2158244014561199

Salvati, L., & Sabbi, A. (2011). Exploring long-term land cover changes in an urban region of southern Europe. *International Journal of Sustainable Development & World Ecology*, *18*(4), 273–282. https://doi.org/10.1080/13504509.2011.560453

Sandoval Félix, J., & Castañón-Puga, M. (2019). From simulation to implementation: Practical advice for policy makers who want to use computer modeling as an analysis and communication tool. In M. A. Sanchez, L. Aguilar, & A. Rodríguez (Eds.), *Applied Decision-Making* (Vol. 209, pp. 115–155). Springer International Publishing. https://doi.org/10.1007/978-3-030-17985-4_6

Schwab, J., & Horiguchi, R. (2020). *Ein offenes Werkzeug zur Messung der räumlichen Zersiedelung* [Bachelor Thesis, Hochschule für Technik Rapperswil]. https://eprints.ost.ch/id/eprint/869/1/FS%202020-BA-EP-Horiguchi-Schwab-Ein%20offenes%20Werkzeug%20zur%20Messung%20der%20r%C3%A4umlichen%20Zersiedelung.pdf

Schwarzak, M., Behnisch, M., & Meinel, G. (2014). Zersiedelung in Deutschland – erste Ergebnisse nach Schweizer Messkonzept. In G. Meinel, U. Schuhmacher, & M. Behnisch (Eds.), *Flächennutzungsmonitoring VI, Innenentwicklung – Prognose – Datenschutz (IÖR-Schriften)* (Vol. 65, pp. 213–222). Rhombos-Verlag, 2014. https://slub.qucosa.de/api/qucosa%253A7520/attachment/ATT-0/

Schwick, C., Jaeger, J. A. G., Hersperger, A., Cathomas, G., & Muggli, R. (2018). *Zersiedelung messen und begrenzen, Massnahmen und Zielvorgaben für die Schweiz, ihre Kantone und Gemeinden* (Vol. 57). Haupt Verlag. https://www.researchgate.net/publication/327670795

Schwick, C., Jaeger, J. A. G., & Kienast, F. (2011). *Zersiedelung messen und vermeiden* (No. 47; Merkblatt für die Praxis, p. 12). Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft. https://www.dora.lib4ri.ch/wsl/islandora/object/wsl%3A9141/datastream/PDF/view

Siedentop, S. (2005). Urban Sprawl—verstehen, messen, steuern: Ansatzpunkte für ein empirisches Mess- und Evaluationskonzept der urbanen Siedlungsentwicklung. *disP - The Planning Review*, *41*(160), 23–35. https://doi.org/10.1080/02513625.2005.10556903

Sim, S., & Mesev, V. (2011). Measuring urban sprawl and compactness: Case study Orlando, USA. *Proceedings of the 25th International Cartographic Conference and the 15th General Assembly of ICA*, 1–10, Paris, France. https://www.researchgate.net/publication/307228230_Measuring_urban_sprawl_and_compactness_case_study_Orlando_USA

Soulard, C. E., Acevedo, W., & Stehman, S. V. (2018). Removing rural roads from the national land cover database to create improved urban maps for the United States, 1992 to 2011. *Photogrammetric Engineering and Remote Sensing*, *84*(2), 101–109. https://doi.org/10.14358/PERS.84.2.101

Stadt Wetzikon. (2021). *Bevölkerungsentwicklung* (pp. 1–3). https://www.wetzikon.ch/verwaltung/einwohnerdienste/bevoelkerungsstatistiken/Jahresstatistik%202020.pdf/at_download/file

Stadt Wetzikon. (2023). *Wetzikon* [Photograph]. https://m.wetzikon.ch/data/private/thumbs/0/3149149.jpg

Stadt Winterthur. (2023, January 13). *Bevölkerung*. https://stadt.winterthur.ch/themen/die-stadt/winterthur/statistik/bevoelkerung

Steinmeier, C. (2013). *CORINE land cover 2000 / 2006: Final report* (pp. 1–30). Swiss Federal Institute for Forest, Snow and Landscape Research WSL. https://www.wsl.ch/de/publikationen/corine-land-cover-2000-2006-switzerland-final-report.html

Steurer, M., & Bayr, C. (2020). Measuring urban sprawl using land use data. *Land Use Policy*, *97*, 1–8. https://doi.org/10.1016/j.landusepol.2020.104799

Stys, B., Kautz, R., Reed, D., Kertis, M., Kawula, R., Keller, C., & Davis, A. (2004). *Florida vegetation and land cover data derived from 2003 Landsat ETM+ imagery* (p. 8). Florida Fish and Wildlife Conservation Commission. https://www.researchgate.net/publication/266449995_Florida_Vegetation_and_Land_Cover_Data_Derived_from_2003_Landsat_ETM_Imagery

Swiss Federal Council. (2022). *Environment Switzerland 2022*.

Terzi, F., & Kaya, H. S. (2011). Dynamic spatial analysis of urban sprawl through fractal geometry: The case of Istanbul. *Environment and Planning B: Planning and Design*, *38*(1), 175–190. https://doi.org/10.1068/b35096

Torrens, P. M. (2008). A toolkit for measuring sprawl. *Applied Spatial Analysis and Policy*, *1*(1), 5–36. https://doi.org/10.1007/s12061-008-9000-x

Torrens, P. M., & Alberti, M. (2000). Measuring sprawl. *Centre for Advanced Spatial Analysis, Paper 27*, 1–34. https://www.ucl.ac.uk/bartlett/casa/sites/bartlett/files/migrated-files/paper27_0.pdf

Tsai, Y.-H. (2005). Quantifying urban form: Compactness versus "sprawl." *Urban Studies*, *42*(1), 141–161. https://doi.org/10.1080/0042098042000309748

UK Parliament. (2023, May). *Churchill and the Commons Chamber*. https://www.parliament.uk/about/living-heritage/building/palace/architecture/palacestructure/churchill/

United States Census Bureau. (2022). Urban area criteria for the 2020 census—Final criteria. *Federal Register*, *87*(57), 16706–16715.

United States Department of Agriculture. (2022, August 1). *National resources inventory*. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/nri/

Vemaps. (2019). *Grey map of Switzerland* [Free vector maps]. https://vemaps.com/uploads/img/large/ch-03.jpg

Viator. (2023). *Grindelwald* [Photograph]. https://www.viator.com/de-CH/tours/Lucerne/Interlaken-and-Grindelwald-Day-Trip-from-Lucerne/d576-2460INTER_L

Wang, Y., Han, Y., Pu, L., Jiang, B., Yuan, S., & Xu, Y. (2021). A novel model for detecting urban fringe and its expanding patterns: An application in Harbin City, China. *Land*, *10*(8), Article 876. https://doi.org/10.3390/land10080876

Zhou, W., Jiao, M., Yu, W., & Wang, J. (2019). Urban sprawl in a megaregion: A multiple spatial and temporal perspective. *Ecological Indicators*, *96*, 54–66. https://doi.org/10.1016/j.ecolind.2017.10.035

# Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent work.
All external sources are explicitly acknowledged in the thesis.

Patrick Schenker
Zurich, 30.06.2023