Reference Data Selection

Add subtitle here

Waldo Tobler

Table of contents

1 Executive summary

Summary goes here.

2 Moprhological data

2.1 Primus

2.2 Satellite-derived building footprints

Satellite derived building footprints are becoming more widely adopted due to advanced in satellite techonolgy and computer vision algorithms. They have sucessfully been used in applications for city boundary delineation, disegregating population and others. There are multiple sources of global building footprints available - Google building footprints, Microsoft building footprints, Overture maps that combines various other sources, as well as other proprietary data.

In this study, we use the Microsoft building footprints since they fully cover our study area - Central Europe. The dataset also has good worldwide coverage and the footprints are derived in a consistent way, which ensures scalability of our method to different countries. It should be noted that our approach can also work for other consistently derived building polygons, Google Footprints for example, so long as all the data is generated from the same source.

However, the data does not come without issues. For example, in dense urban centres entire blocks can be delineated as individual buildings. Given that morphology calculations rely on precise local topological relations between neighbours, such as two buildings touching, this problem renders a whole number of possible measurements described in [@fleischmann2021measuring] meaningless. Furthermore, this issue affects even simpler calculations such as counting the number of buildings within a radius or topological neighbourhood. Other issues are that computer vision techniques sometimes miss entire buildings or misidentify building boundaries. Therefore, any approach that uses satellite-derived building footprints should be able to account for these three and potentially other problems.

2.3 Overture maps - street data

The street network is a direct download from Overture maps, a processed subset of data from Openstreetmap, which has global coverage and high quality data. Since the dataset includes multiple segments types, including footpaths, the types of segments used in the analysis are limited to … . Another type of segment that is filtered out are tunnels - the analysis strictly focuses on two dimensions and therefore undergrounds structures adversely affect the calculation of boundaries and characters.

2.4 Validation data

3 Satellite data