Paper frame

**Connecting green space supply and demand: modeling the walkable environment of European cities**

Benjamin Labohm

**Introduction**

* What is the scope of the problem?
  + Improving the modeling of the walkable environment / service connecting for urban green areas in European cities based on a proximity approach
  + Provide solutions for easy-to-handle / understandable indicators based on publicly available data and software
  + Develop an approach to combine a high resolution data with an comparative approach
* What do we know about the problem?
  + Urban population is growing worldwide, green space development?
  + Living in the proximity of urban green spaces (UGS) was found to have positive effects on mental health and physical activity and to reduce mortality.
  + In the Sustainable Development Goal (SDG) 11, the United Nations have agreed to provide universal access to public green spaces by 2030.
  + Compact city paradox:
    - The more compact a city, the more people benefit from UGS, the higher the pressure on the ecological functions of the UGS
  + The accessibility of available green spaces is influenced by spatial as well as non-spatial factors
    - Accessibility of UGS vs availability of UGS (?)
    - Three perspectives have been used in past studies to tackle the problem:
    - Provision (area / person): Flow from green area to buildings (providing green space)
    - Pressure (person / area): Flow from residential buildings (population) to green spaces (pressure on / demand of green areas)
    - Proximity (min. or avg. distance): Space between supply and demand (proximity)
    - Approach from Wolff 2021: Coupling the perspectives via network characteristics
      * Modeling service connecting areas / the walkable environment to allow including the three perspective into one workflow.
      * Two promising indicators: Detour Index (DI) and Local Significance (LS).
  + Accessibility and availability of UGS in Europe have been analyzed and compared in multiple studies e.g.:
    - State of the art
      * Analysis for whole Europe:
        + Kabisch et al. 2016 green spaces Europe
        + JRC accessibility study
      * Analysis for single city:
        + Wolff 2021
        + Haase Kabisch Berlin paper
* What do we not know about the problem?
  + Research gaps (of previously mentioned studies):
    - Mostly using fixed distances
      * ‘To have or not to have access to green spaces’
    - Mostly using only one perspective
      * Not accounting for mutual dependencies of supply and demand
      * No focus on service connecting areas / walkable environment
      * A proximity perspective is necessary to account for barriers and network characteristics
    - Mostly using course resolution if on a larger scale
      * Fine resolution can reveal spatial patterns at a finer scale enabling targeted intervention
      * A finer resolution will also reduce the uncertainty that is introduced if e.g. a grid or a district aggregation is used
      * No comparable studies using e.g. a building based approach on a European scale
  + Conclusion:
    - Knowledge about UGS is important for planning and decision making to remove uncertainties in UGS provision
    - Detect mismatch between provision and demand
    - Need for a method that combines spatial and non-spatial aspects of green space accessibility / availability with a high spatial resolution that is comparable on a European scale
* What is the purpose of this paper?
  + Research questions
    - Is the approach feasible to model green space proximity in European cities via network characteristics at a high resolution using open source data and software?
    - How do data availability and resolution affect the results?
    - Can we detect mismatches between green space supply and demand on an inter-city / inner-city perspective?
  + Objective 1:
    - Create a homogeneous workflow for the modeling of service connecting areas / walkable environment (on a building level) across European cities using publicly available data and open source software:
      * Create Network characteristics for European cities:
        + Combine buffer distance and network distance (Detour index)
        + Combine population, GBS size network distance (Local significance)
  + Objective 2:
    - Find ways for effective comparisons and pattern detection both within cities  
      and between cities
    - Detect mismatch between provision and demand using the results of objective 1.

**Methods**

* Study region
  + To reduce the computation time needed for the analysis, this research will be focused on European cities with more than 500.000 inhabitants
  + This focus might be extended for future studies
  + ‘European cities’ in this study are based, partly on urban atlas city core layers and partly on ... (Manus layer specifications)
  + To account for the ability of city dwellers to leave their city if they live in proximity to a city border, GBS from a buffer of 1 km surrounding the city core will be included in the study
* Study design
  + Data
    - For the analysis of the walkable environment of European cities, we needed available and comparable data on public green spaces and residential buildings and their respective entry points.
    - Additionally, the analysis requires information on the population in each residential building and a network that connects the buildings with the green spaces.
    - We agreed on using Urban Atlas (UA) and OpenStreetMap (OSM) as our main data sources.
    - To ensure comparability across all data sources, all data was acquired from the year 2018 where possible
    - Urban Atlas (UA)
      * UA is a land use / land cover (LULC) product from the Copernicus program of the European Union.
      * The 2018 UA version provides Europe-wide comparable data for 788 Functional Urban Areas with more than 50.000 inhabitants.
      * It represents 17 urban and 10 rural LULC classes with a MMU of 0.25 and 1 ha, respectively.
      * In addition to the spatial data, UA contains information on the population for the residential land use classes.
      * Since Copernicus does not offer an API, we downloaded the latest UA version available at the time of this study (v13) by hand.
    - OpenStreetMap (OSM)
      * OSM is a community based project that provides free geospatial data.
      * The OSM community seeks to create a database of the entire planet that is free and editable.
      * For the analysis we downloaded the OSM data with the identifiers ‘building’ and ‘highway’.
      * We downloaded all OSM data within a 1 km buffer around the city core.
      * Since OSM is a community based mapping service, the latest version of OSM data is expected to have the most information.
      * We acquired the OSM data via the OSM API, which we accessed via a custom-made tool that heavily relies on the R package ‘osmdata’.
      * For information on the OSM download tool see Appendix ...
  + Data pre-processing
    - Network
      * The network, in this analysis, represents the walkable environment of a city, which connects the entry points of the UGS with those of the residential buildings.
      * To acquire information on the street network of a city, we downloaded OSM data with the identifier ‘highway’, which represents all linestrings in the OSM database that are associated with streets and paths.
      * To secure comparability across European countries, we used all OSM highway classes, except for the class ‘highways’.
      * Linestrings identified with the string ‘highway’ represent motorways which are reserved for motorized use only.
      * To ensure network connectivity and reduce overlap, we cleaned the resulting network following the tutorial on network data pre-processing and cleaning by Lucas van der Meer.
      * Further information on the network pre-processing and cleaning steps can be found in Appendix ...
    - Buildings
      * The following analysis requires information on a cities’ residential buildings, their entry points and on how many persons inhabit each building.
      * We filtered the OSM ‘building’ polygons for residential buildings (include table with excluded words: ‘university’, ‘school’ etc.)
      * We only kept those OSM building polygons that were contained inside of urban atlas residential areas (UA class code starting with: 11).
    - Building entries
      * To detect building entries, we first calculated the centroids of each building.
      * Centroids had to satisfy the constraint, that the point has to lie inside the polygon (see st\_point\_on\_surface).
      * We snapped the centroids to the closest point on the cleaned street network and assumed the resulting points to be the building entries.
    - Population per building
      * The UA dataset provides information on population mostly on a city block level.
      * We disaggregated this data to the building level by distributing the population proportionally to a buildings base area.
      * This workflow follows the assumption, that the building structure, and thus the population per base area inside one city block is similar.
      * For buildings that were contained inside UA residential polygons that erroneously did not have population values, we used the mean population per square-meter of the corresponding UA residential class in the city.
    - Green spaces
      * The last data point required for the analysis is information on publicly accessible UGS and their entry points.
      * We filtered the UA data for the classes ‘green urban areas’ (code 14100) and ‘forests’ (code 31000) to ensure that all green spaces that are used in the analysis are publicly accessible.
      * All green spaces in the UA dataset come with information on area, which we double checked for consistency.
    - Green space entries
      * To reliably detect green space entries, we intersected the outline of the UA green spaces with the cleaned network.
      * We applied different buffer sizes until each green space in a given city had an entry point.
      * We used the resulting points as entry points of the green spaces for the further analysis.
      * Further information on the method and the validation / sensitivity analysis can be found in Appendix ...
    - Network blending
      * In the process of network blending, the entry points of the residential buildings and the UGS (now called ‘nodes’) are being ‘blended’ into the network.
      * During this process, the lines (now called ‘edges’) will be broken at every node location.
      * The node location now represents the new starting / ending points of the newly created edges.
      * For more detailed information on this process see Appendix ...
  + Analysis
    - Analysis to achieve Objective 1:
      * Short summary of Objective 1
      * Modeling service connecting areas / creating workflow for index creation
      * Index building
        + We assumed the distance people will usually walk to their nearest UGS to be 500 m.
        + By making this assumption, we also limited the computational requirements for the analysis.
        + We calculated the Detour Index (DI) and the Local Significance (LS) index for each UGS inside the city core plus a buffer of 1 km.
        + We accounted for the maximum walking distance by calculating both indices for the residential buildings within a network distance of 500 m between a building entry and the nearest entry point of a UGS.
      * Detour Index (DI)
        + The DI is an indicator of barriers in a network.
        + It accounts for the detours people have to take on their way to the nearest UGS.
        + It combines the Euclidean distance, i.e. the direct connection between two points, with the network distance:

Where *DI* is the Detour Index, *d(i, j)* is the Euclidean distance between points *i* and *j*, and *nd(i, j)* is the network distance between the points *i* and *j*.

* + - * + In the case of this analysis, the two points are the entry points of a residential building and the nearest entry point of a UGS.
        + The DI can assume values between 0 and 1.
        + A DI value of 1 represents a straight line between building entry and UGS entry, while a DI value closer to 0 means that the inhabitants of the building have to take a sub-optimal route to the nearest UGS.
        + If one building has access to several UGS within a network distance of 500 m, we decided to use the mean DI value.
      * Local Significance (LS)
        + The LS is an indicator of how many people have access to a UGS.
        + The LS also takes into account the size of an UGS as well as the distance between peoples homes and the UGS entries:
        + Where *LS* is the Local Significance, *p* is the population of building *i*, *A* is the area of UGS *j* and *nd(i, j)* is the network distance between the entry points of building *i* and UGS *j.*
        + This indicator can assume infinite values.
        + A higher population and area, as well as a lower network distance lead to higher LS values.
        + We attached the LS values to each segment (edge) of the path between building and UGS entries.
        + We summed the LS values of overlapping paths from multiple buildings, leading to higher values on higher frequented edges.
        + A more detailed summary on index building can be found in Appendix …
      * Validation steps
        + Check for percentage / area of UA residential polygons covered by OSM buildings
        + Do the same for network (?)
    - Analysis to achieve Objective 2:
      * Short summary of Objective 2
      * Comparing cities / detecting mismatch
      * European level
        + Comparing distribution of DI and LS values in different cities with histograms / line plots (?)
        + Using LS / m (?)
      * City scenarios
        + Test influence of changes in different parameters
        + Lene Voigt Park Leipzig

Introduction of the LVP situation

location

surrounding city structure

population

green spaces in proximity to LVP

* + - * + Scenario 1

What if the park could be entered from all sides?

‘Cutting down all the hedges’

Creating entry points at every X meters on the outline of LVP

* + - * + Scenario 2

What if there was a population increase?

‘Growing city’

Increase population to e.g. a) max possible (95th percentile?) for each class residential, b) change all residential to high density residential?

* + - * + Scenario 3

What if the surrounding green spaces were developed to residential buildings?

Change the green spaces in the north of LVP to high density residential

**Results**

* + - Results of Objective 1
      * What was feasible? What not? Where? Why?
      * Maps / graphs showing OSM data coverage in different areas
    - Result of Objective 2
      * European level
        + Tree / graphs like in Wolff, Scheuer et al. 2020
        + Maps like in Kabisch et al. 2016
      * City scenrios
        + Show and discuss the results:

For each scenario:

2 maps: one of status quo model, one of altered parameter model

Discussion

Summary of the BIG THREE findings (1, 2, 3), i.e. the higher-level idea

Effect of new approach on results

Inter- / inner-city disparities

For each of the findings:

How does this map/relate to previous findings?

Compare to e.g.

How does this add to previous findings?

Comparability / applicability with public data

Challenges

Implications