Reproducibility review of "Developing a city-specific walkability index through a participatory approach"

Author: F.O. Ostermann ORCiD [0000-0002-9317-8291]

To cite this report use

Ostermann, F. O. (2024). Reproducibility review of “Developing a city-specific walkability index through a participatory approach”. https://doi.org/10.17605/osf.io/csb7r

This report is part of the reproducibility review at the AGILE conference. For more information see https://reproducible-agile.github.io/

Reviewed paper

Cardoso, M., Milias, V., and Harteveld, M.: "Developing a city-specific walkability index through a participatory approach AGILE GIScience Ser., 5, 2, https://doi.org/10.5194/agile-giss-5-2-2024, 2024

Summary

The paper investigates to improve on walkability indices that consider city-specifics. It employs a mixed methods approach, starting with a literature review to identify relevant elements, followed by participatory sessions with urban planners and policymakers to determine factors for the case study of Amsterdam. In a last step, the authors quantify these factors and calculate walkability scores for the study area. The first two steps in the research design are out of scope of this reproducibility review, which focuses on the computational reproducibility of the quantitative analysis (phase C in Figure 1 of the paper).

The quantitative analysis uses a wide range of data sources as input, to represent the various factors identified during phases A and B, including several official datasets from public authorities as well as OpenStreetMap. According to the others, all data sources are openly available. However, although links to data providers and concrete data sets are provided, there were several manual steps necessary to pre-process and integrate the factors with the OSM street network: The OSM street data was split into segments, and then for each segment a score for each factor was calculated, e.g., number of trees or bushes, or number of streetlights. This pre-processing and integration were done manually in QGIS. Although the spatial analysis steps for this integration are described in the paper’s Github repository (reproducible-agile/WalkabilityInAmsterdam: Files to reproduce the results of the paper: Developing a city-specific walkability index through a participatory approach (github.com)) and are fairly standard (buffer, counts, zonal statistics, spatial joins), the necessary time to collect all the datasets and reproduce these steps were out of scope for this review.

The authors provide the following data sets in the repository to execute the steps of normalization, calculation, and visualization (compare Figure 3 in the paper): Three CSV files containing the number of amenities facing a street segment, the number of obstacles per street segment, and additional information per street segment (all results of the mentioned earlier manual analysis steps). Further a GeoPackage containing walkable streets in Amsterdam, and a Python script to calculate the

walkability score. It takes the CSV files and calculates the walkability score per street segment as input and generates several walkability scores for each street segment.

The generated output could not be compared quantitatively to the output in the paper, because no output data set is provided, although the results can be interactively explored on the project's website (Info - CTstreets (miliasv.github.io)). However, a visual inspection using QGIS (after joining the output CSV with the GeoPackage as documented) shows similar patterns of walkability scores.

Below an export showing the Walkability scores using the same color scheme and classes as the upper left map in Figure 5 of the original paper:

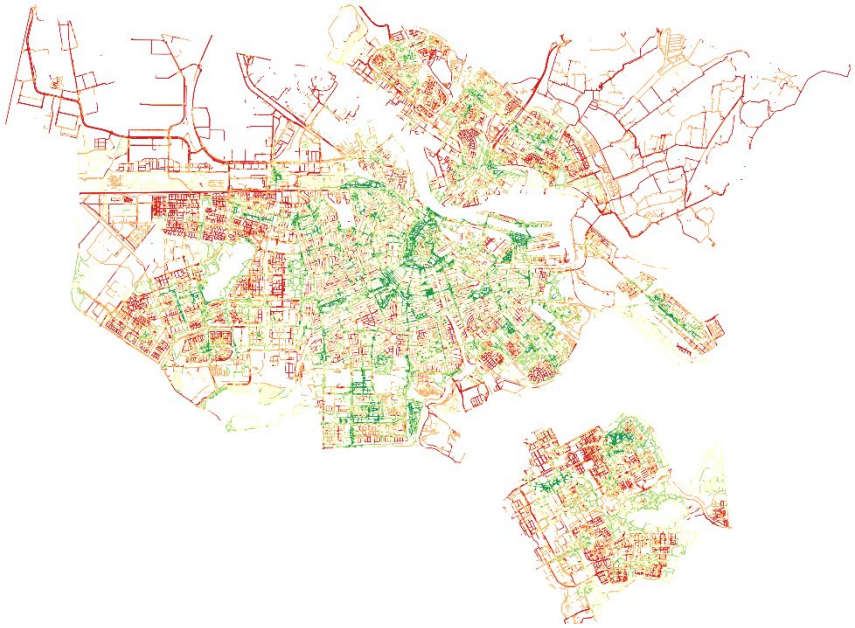

Figure 1: Output of the reproduction of the Walkability score using similar color scheme and classes as Figure 5 of the original paper

This reproducibility review was thus only able to validate part of the workflow. However, given that the input data is in principle openly available, the manual analysis steps are standard operations with documented parameters, the computations are deterministic, and all tools and software are free and open source, this review can conclude that the quantitative analysis part of the paper is reproducible.

Reproducibility reviewer notes

After manually adjusting file paths, the script ran without problems on Windows 11 and Ubuntu 22.04 VM Python 3.12.x installations, both with 16 GB RAM.