Code

The implementation of the methods described below is available from https://doi.org/10.5281/zenodo.5786969.

Methods

This supplement details the data sources, manipulation, and preparation process, including the main tools used throughout the research.

Software and tools

Various tools were used in order to prepare, process and visualize the data. Preparation of the network and land-use datasets took place in QGIS. The results of the preparation were handled in R to run the analysis and the outputs were in then visualized as maps in QGIS.

The main preparations in QGIS cocern road network adjustments and space syntax calculations. Place syntax tool is a QGIS plugin developed in C++ by a group of researchers from KTH School of Architecture, Chalmers School of Architecture and Spacescape AB. It was created as an extra input in Space syntax theory called Place syntax and calculates space syntax parameters as well as attraction reach parameters towards points of interest (Stavroulaki et al., 2019).[1]

This research utilizes the road center lines and thus it is easy to calculate angular and not metric choice values. The segment map is created, and calculations are performed. The distance mode is chosen as walking distance for the various radii since metric distance from each point needs to be taken into consideration. There is no need for normalization after angular choice calculations. Raw values are then firstly utilized to visualize space syntax analysis and secondly, after densifying and turning lines to points, as input to statistical calculations in R.

The land-use point dataset is validated and adjusted to the project needs in QGIS. Then the land-use points are used as an input in R. Both point patterns support the Kernel Density Estimation and subsequent Kernel Density Correlation and spatial correlation. The raster adjustments (including masking the study area), raster calculations and zonal statistics are also computed in R. All resulting layers produced are imported in QGIS in order to map them.

Data sources and description

Our study required two main datasets and two supplementary datasets. The main datasets are those used during the process itself, such as segments of the road network and the points of land-use within the study area. The supplementary datasets are datasets not utilized during the process but for pre-

[1] Documentation about the tool and examples can be found on their website https://www.smog.chalmers.se/pst.

processing and post-processing tasks. Examples include the municipality boundaries within the study area and the lakes as raster masks. All datasets, apart from land-use points, are downloaded from Kortforsyningen.[2] Land-use point data are collected from Open Street Map (OSM) through the Overpass API (see Supp Fig 1).

```
overpass_query = """[out:json];area["ISO31661"="DK"][admin_level=2]->.searchArea;(node["shop"](area.searchArea););out center;"""
response = requests.get(overpass_url,params={'data': overpass_query})
data_shops = response.json()
print('data_shops',data_shops)
```

Supplement Figure 1: Code snippet example in python for downloading shops in Denmark with Overpass API

The road network layer is described as the Road Center Lines (RCL) within the study area and contains road segments and paths. Paths can be those passing through train stations and connecting two public spaces. They were excluded since space syntax analysis must include segments with potential pedestrian and vehicular movement in the public space and paths only support pedestrian movement, potentially in private spaces.

Land-use points were used to detect active centers in the private or public sphere which are also spaces of co-presence and economic activities. Four main tags were included in the queries to the Overpass API after research conducted in OSM Wiki about their terms: *amenities*, *shops*, *offices* and *leisure*. The corresponding land use points were carefully examined afterwards and modified based on the project needs. An example of modification is the removal of features such as benches, traffic lights, barbeque sites etc. which do not represent spaces provoking co-presence or economic activity and are mainly covered by spaces in which they are placed in, such as parks.

Previous research points at the varying quality and completeness of the OSM data, often with more complete and more accurate data in areas with higher population densities and, therefore, more potential mappers (Cipeluch et al, 2010; Haklay,2010; Basiri et al., 2016). Bearing in mind these issues, OSM was still the best available data source for our study, particularly considering that it focuses on the largest urban area of Denmark.

Data preparation and process

The data was explored and prepared for road network and land-use, individually. After the preparation, both datasets are processed either individually or are combined and their results are analyzed.

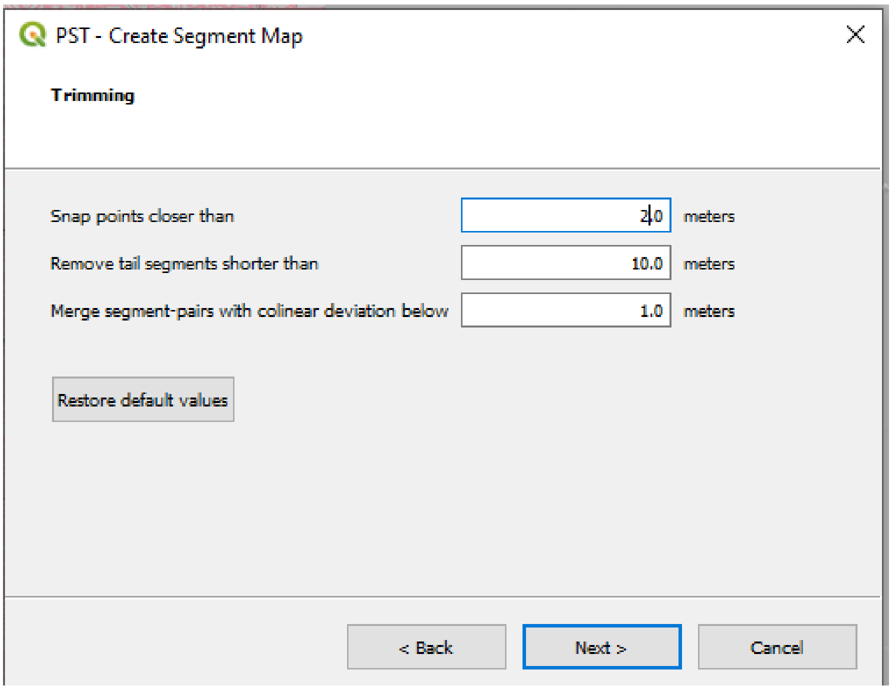
[2] See https://www.kortforsyningen.dk/.

Road network

The study area was defined by Copenhagen's finger plan where municipalities that include centers and other urban areas specified by the plan were selected. The road network intersecting a 3KM buffer around these areas was selected, and another buffer was applied to minimize the edge effect of space syntax analysis. The results are clipped to the initial extent again.

The place syntax tool[3] for QGIS is used to create a segment map of the road network (see Supp Fig 2). The segment map tool then proceeds to topological corrections such as:

1. Splitting polylines to lines
2. Splitting lines at intersections
3. Snapping end-nodes which are falling within a specific threshold
4. Removing duplicate feature lines
5. Removing zero-length features
6. Removing tails of lines by a given threshold
7. Merging two segments which are connected into one if the collinear deviation of the three nodes creating those lines is falling within a specific threshold.



Supplementary Figure 2: Place Syntax Tool and creation of segment map

The procedure described above created a topologically correct road network. Disconnected parts were fixed by using the Disconnected islands plugin.[4] It distinguished the islands and helped review them. Those parts turned out to be either segments crossing parking lots or auxiliary lanes, so they were removed.

[3] See https://www.smog.chalmers.se/pst
[4] See https://plugins.qgis.org/plugins/disconnected-islands/

Before running the analysis, line simplification was applied since natural movement supports neglecting small changes of angles both during vehicular and pedestrian movement. The QGIS simplification tool was utilized and a tolerance of 5 meters was applied.

The space syntax tool was then used to calculate the angular choice within different radii. The study focuses on pedestrian movement but in order to make a comparison on how centralities are approached, the analysis has to be repeated for various scales. Selected radii are 400 meters depicting a 5-minute walk, 800 meters for a 10-minute walk, 1600 meters for 5-minutes cycling distance, 2400 meters for 10-minutes cycling distance and 10 kilometres for the global scale which represents movement by car (Hillier, 2007; Porta et al., 2012; Porta et al., 2009, Paraskevopoulos and Photis, 2019).

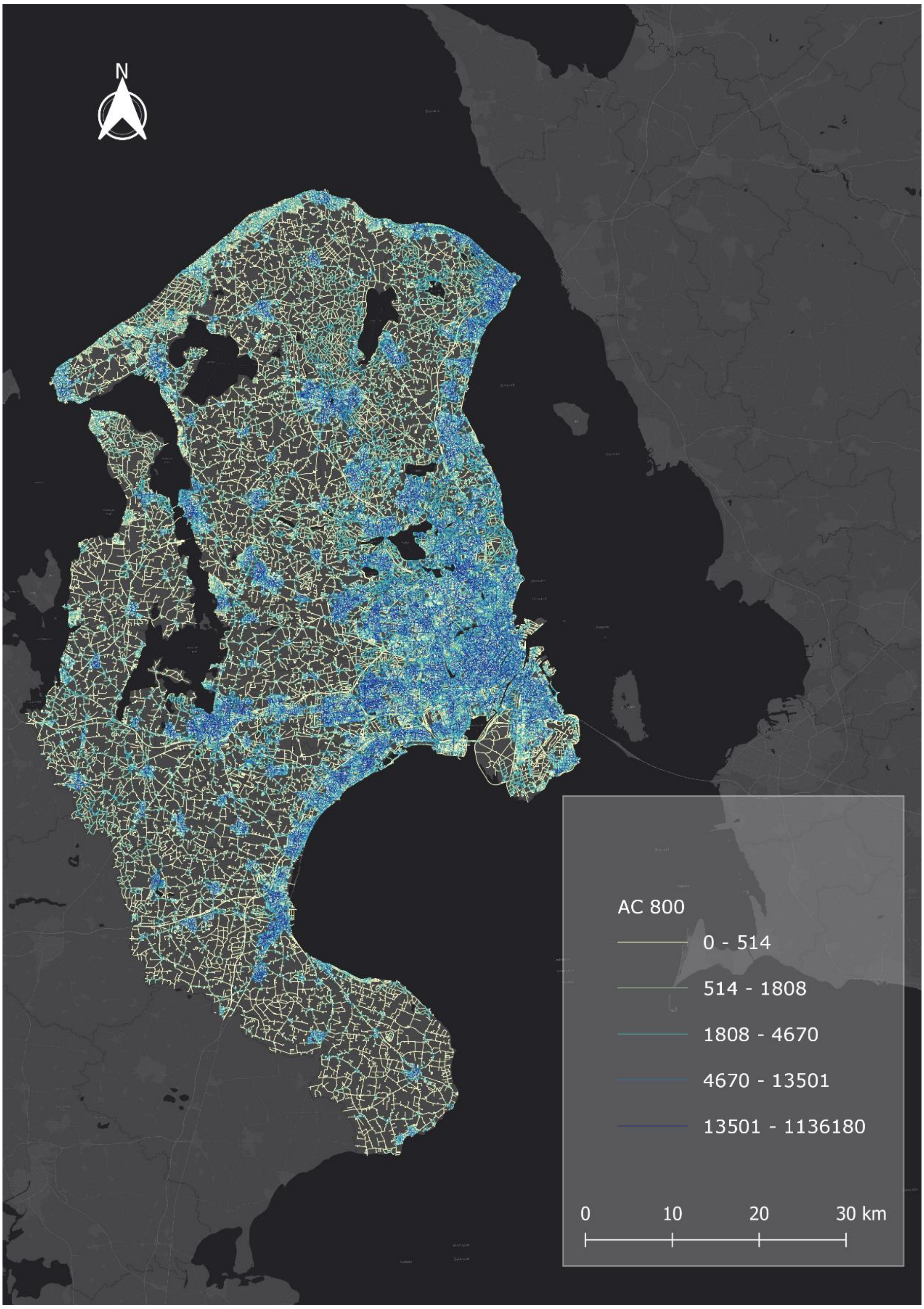
The results of the analysis were used as input to a collection of R scripts where a further manipulation was implemented in order to correlate centers created by the road network with active centers created by land-use points. Kernel Density Estimation (KDE) is applied for both datasets with the same pixel sizes so that the comparison of centers will be pixel based. Since there seems to be no R package that supports Kernel Density Estimation from lines, this function has been implemented based on ESRI's ArcGIS product documentation.[5] Documentation suggests that the algorithm calculates Density as shown in Supp Eq 1, in which L1 and L2 are standing for the two line lengths, V1 and V2 represent the population values and the area of the circle is taken into account.

$$Density = \frac{(L1*V1)+(L2+V2)}{area\_of\_circle} \quad \text{(Supp Eq 1)}$$

Our implementation densified the nodes of the road network by 20-meter interval, turned lines to points from their nodes and assigned the values calculated from space syntax analysis as attributes. A point pattern was created, and weighted KDE analysis was implemented for each population field, which in this case are the Angular Choice calculations for various radii.

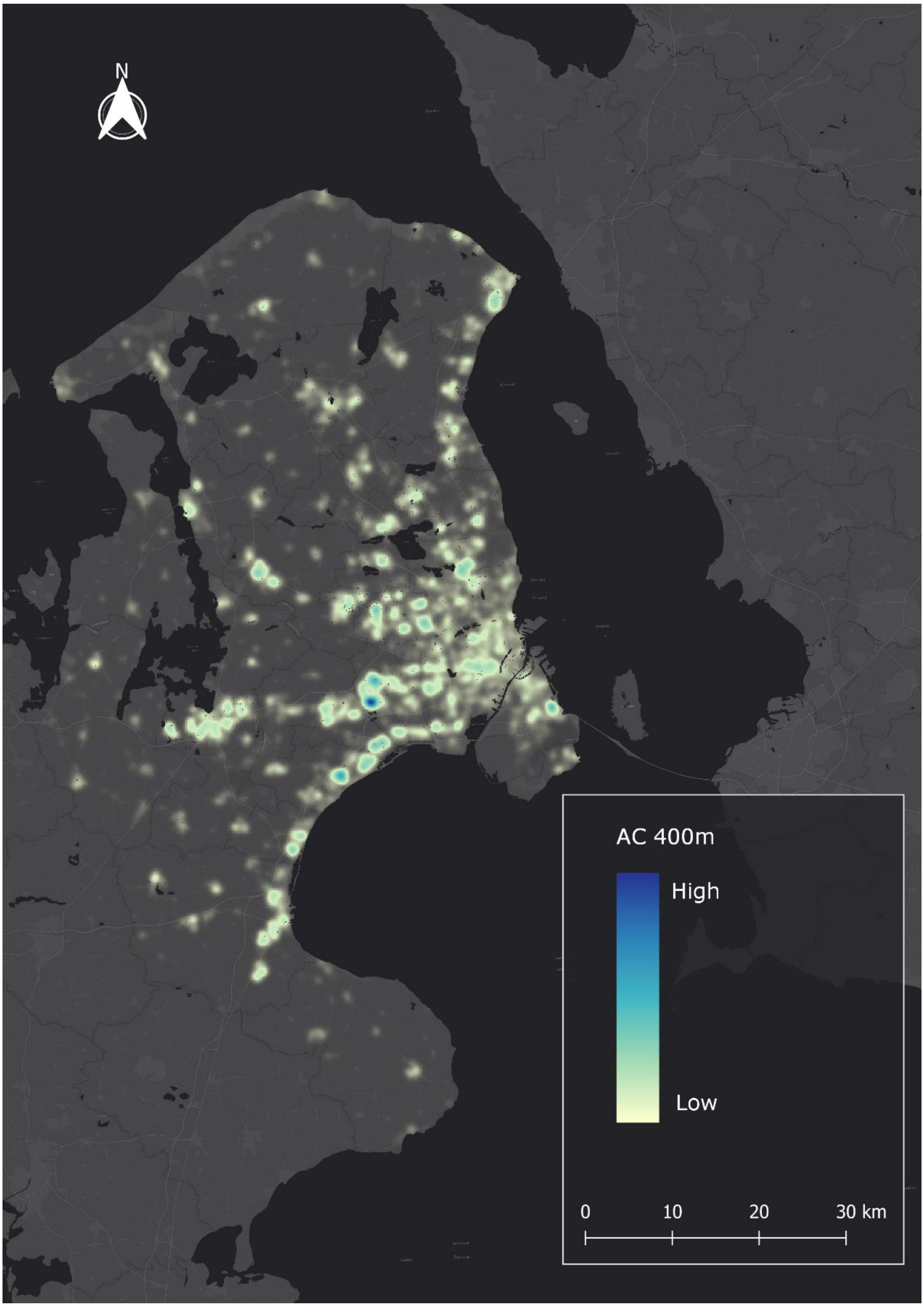
[5] See https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-line-density-works.htm

Supplemental Figures

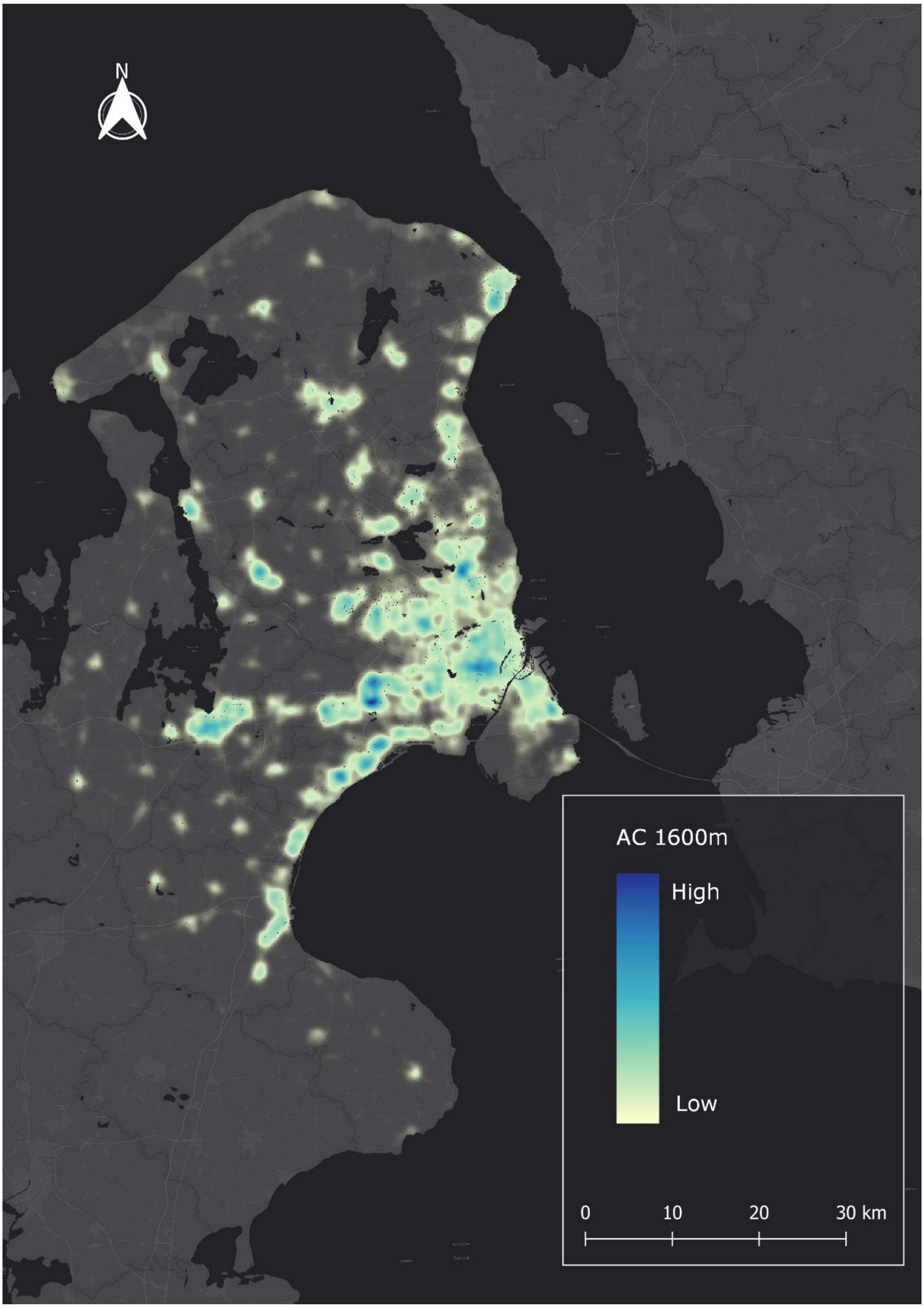


Supplement Figure 3: Angular choice calculations for 800m.

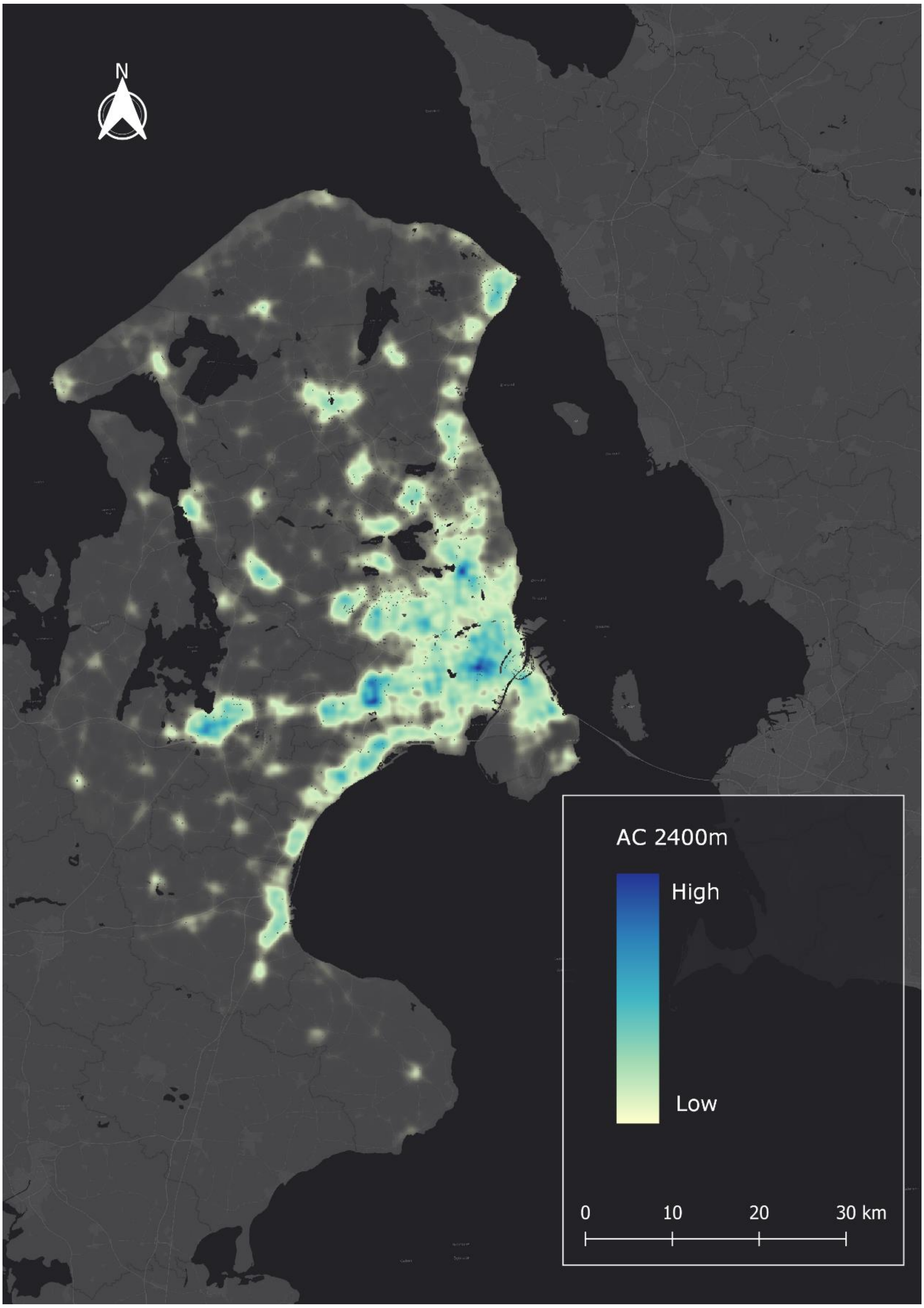
Supplement Figure 4: Angular choice calculations for 10000m.



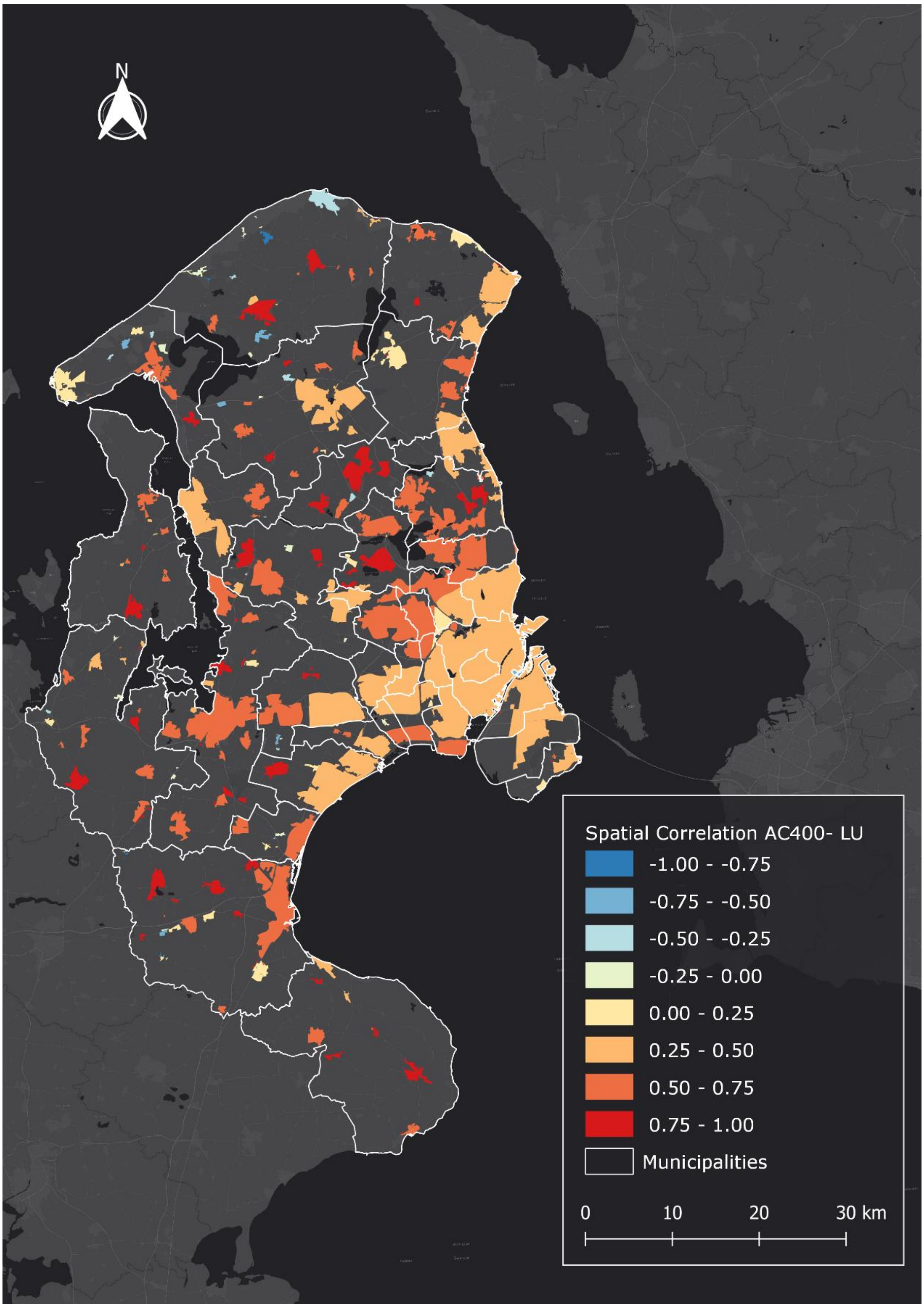
Supplement Figure 5: Angular choice KDE calculations for 400m.



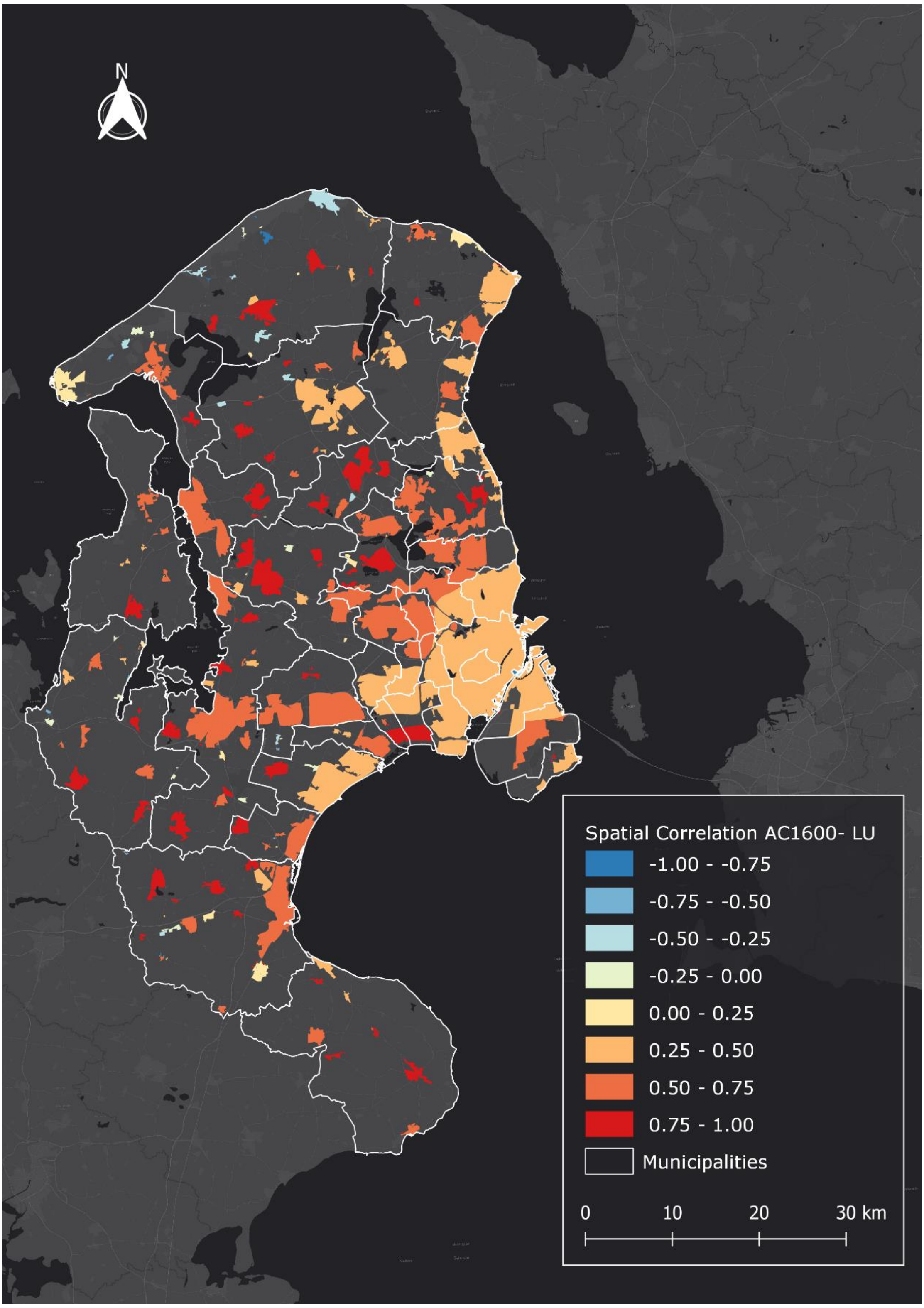
Supplement Figure 6: Angular choice KDE calculations for 1600m.



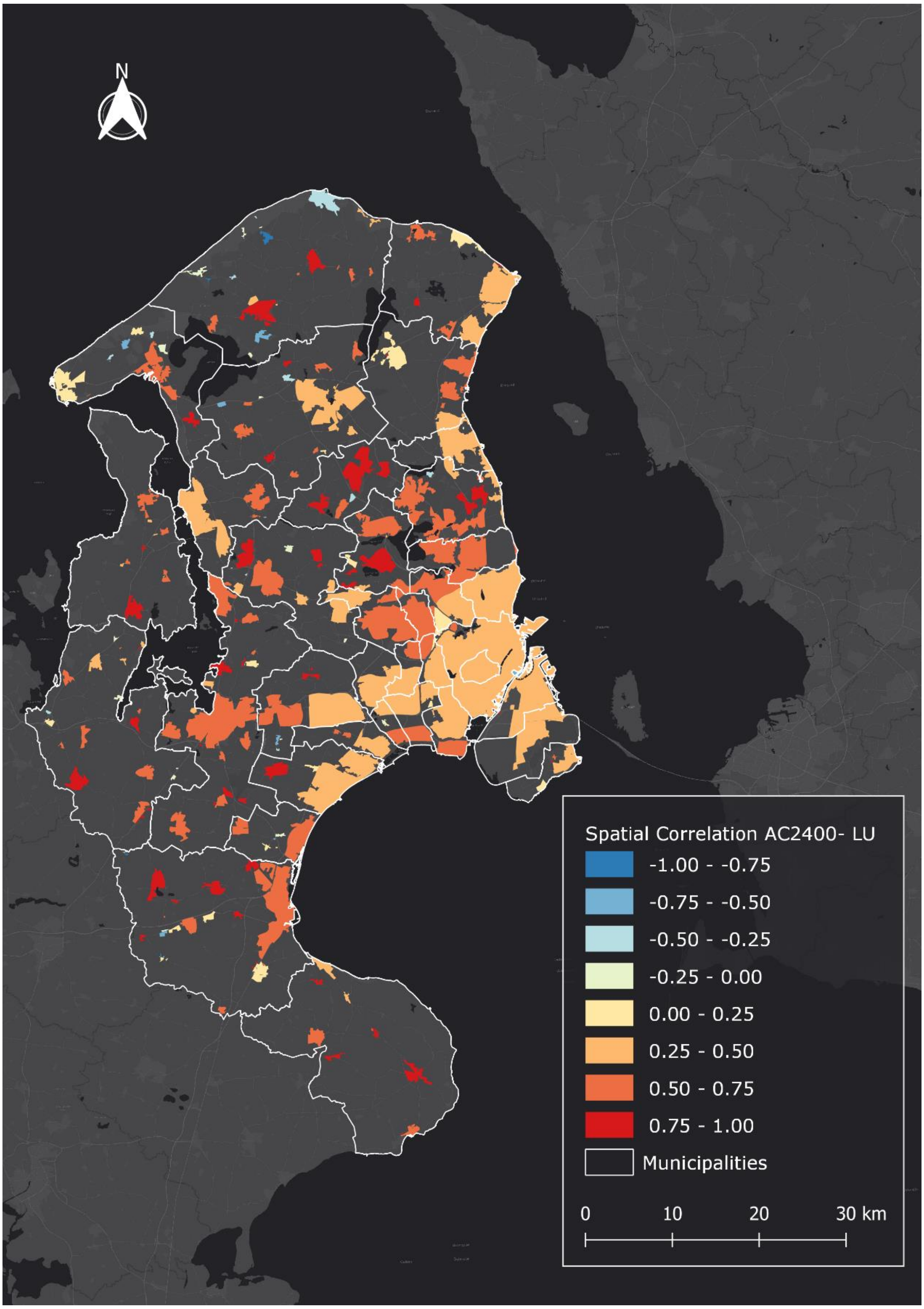
Supplement Figure 7: Angular choice KDE calculations for 2400m.



Supplement Figure 8: Spatial correlation of KDE for angular choice values and land-use KDE at municipality level for AC400.



Supplement Figure 9: Spatial correlation of KDE for angular choice values and land-use KDE at municipality level for AC1600.



Supplement Figure 10: Spatial correlation of KDE for angular choice values and land-use KDE at municipality level for AC2400.