Electronic Supporting Information - Evaluating spatial material distributions: Adopting geospatial entropy definitions into resource management

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Cristina Moyaert[a], Philippe Nimmegeers[a,b,c], Bilal Mellouk[a], Dimitri Voordeckers[d], Paul De Meulenaere[e,f], Pieter Billen[a]

1. Geospatial Entropy Definitions

1.1 Area Distribution

To illustrate this concept, the entropy definition is evaluated for several cases: the cases are built in a logical way. To start, each area consists of 4 sub-areas, which are called cities. The cities can alter between two area levels: High (50 km²) and Low (5 km²). Since there are 4 cities, and 2 possibilities for each city, this results in a total of 5 possible combinations or cases to explore: LLLL, LLLH, LLHH, LHHH and HHHH, as illustrated in figure 1.

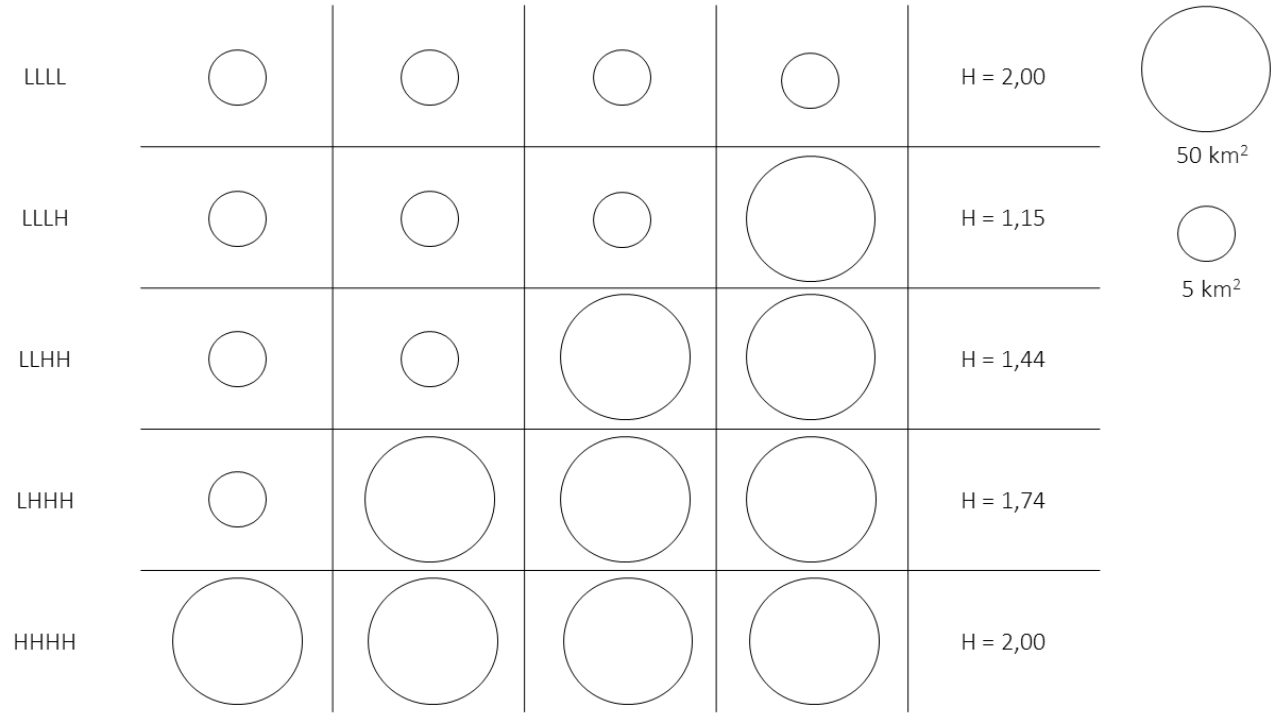

Figure 1: Illustrative overview of the possible combinations of area levels in 4 cities, along with the entropy values of the combinations. The maximum spread corresponds with maximum entropy (LLLL and HHHH) and area clustering LHHH (with only one big H) has minimum entropy.

It is clear from this figure that both case 1 and 5 match the case in which the entropy is its maximum value, namely 2. This corresponds with the case where the distribution is equal, meaning the value of all 4 sub-areas is the same. Note that the lowest entropy case is case 2 corresponding with 1 large area and 3 equally small areas (LLLH). Considering the boundaries of the case study's assumption, this is the minimum value.

1.2 Population Distribution

To illustrate this concept, the entropy of several cases is calculated: Similarly to the previous definition, an area consisting of 4 sub-areas is defined. The variable at play here is population, which is altered according to 2 levels: High and Low. The High level corresponds with an population of 100.000, and the low value corresponds with population of 10.000. Equivalently to the cases of area distribution, there are a total of 5 possible combinations: LLLL, LLLH, LLHH, LHHH and HHHH, as illustrated in figure 2.

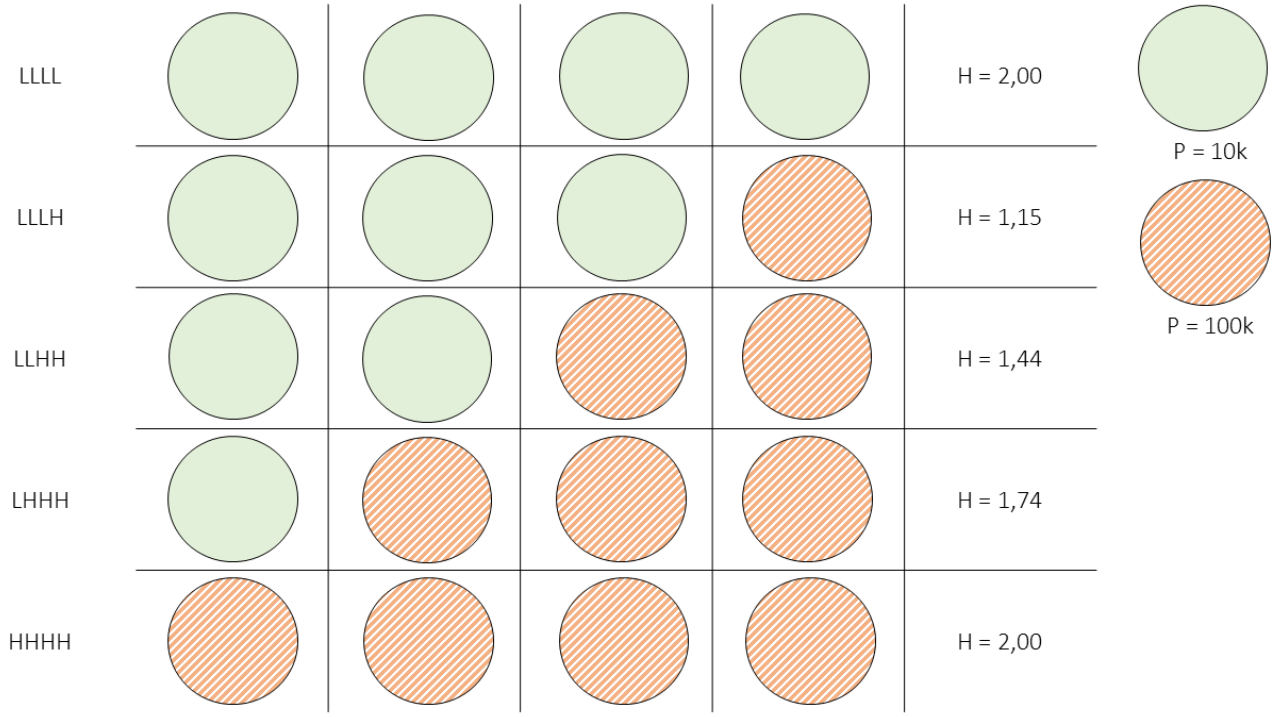

Figure 2: Illustrative overview of the possible combinations of population levels in 4 cities.

The calculated entropy for these 5 combinations/cases, is mentioned in the last row. It is clear from this figure that both case 1 and 5 match the case in which the entropy is its maximum value, namely 2. This corresponds with the case where the distribution is equal, meaning the value of all 4 sub-areas is the same.

1.3 Population Density Distribution

Following the two sections above, the example cases will be calculated by alternating the values of 4 cities according to H and L levels, following the same LLLL, LLLH, LLHH, LHHH and HHHH pattern. However, since there are two factors at play here: population and area, there are significantly more possible combinations, namely 25 (5 population combinations times 5 area combinations). An overview of all 25 combinations can be found in figure 4, and the calculated entropy for these cases can be found graphically plotted in figure 3.

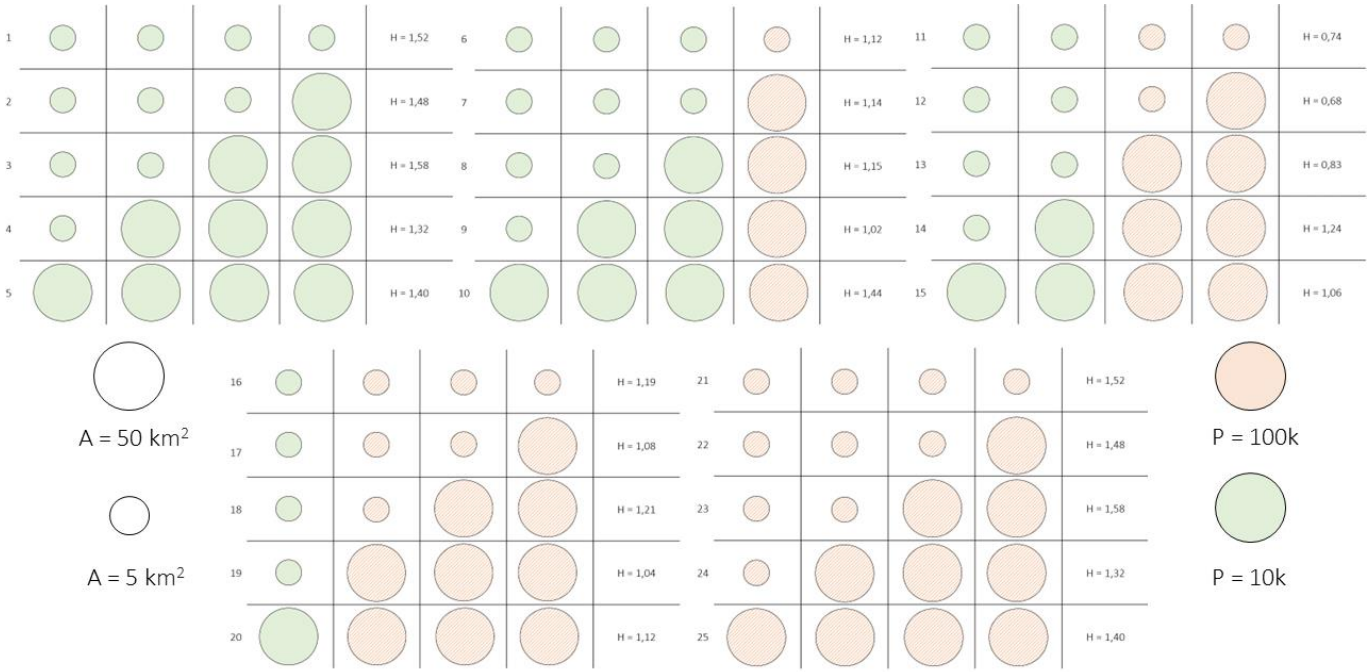

Figure 3: Illustrative overview of the possible combinations of population and area levels in 4 cities, together with their corresponding entropy values.

All values fluctuate between 2, 1.44, 1.15 and 1.74, which are the same recurring values we see in the area and population definitions entropy (definition 1 and 2). This similarity in values originates from the similarity of the probability values pi, since the used case method of alternating H and L values results in identical combinations of values.

1.4 Population Density Levels

The construction of the cases for this type of distribution is somewhat different than the previous ones. The cases are built upon an area consisting of 4 zones. Inside this area, there are 10 cities, which need to be divided into these 4 zones. The 22 possibilities combinations to make this division, are illustrated by figure 6. The calculated entropy for these 22 combinations, is graphically plotted in Figure 4.

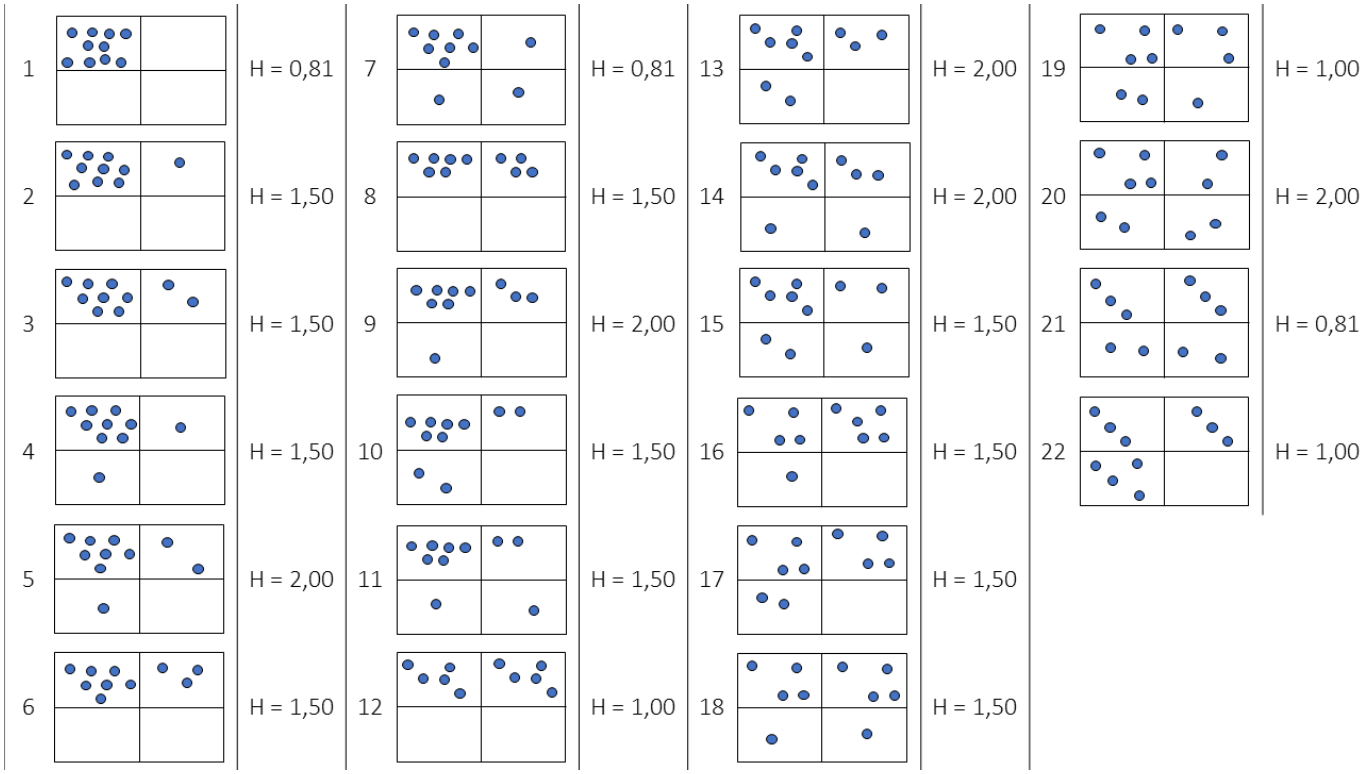

Figure 4: Illustrative overview of the 22 possible combinations to divide 10 cities over 4 areas, and their corresponding entropy values.

Most values linger around either 1,5 or 2. 1,5 is the entropy for the cases in which all cities are divided between 2 areas, so the density level 0 has 2 occurrences. the maximum value of 2 is achieved when each density level has 1 occurrence, meaning each zone has a different amount of cities.

1.5 Percentage of land development

In order to illustrate this example, case examples are set up. Their structure is similar to the previous examples: there are 4 zones, and the land development density of each zone is altered using High (80%), Medium (50%) and Low (20%) values. This gives a total case combination of 13 possibilities, which are illustrated in figure 10. The calculated entropy of the possibilities are plotted in figure 5.

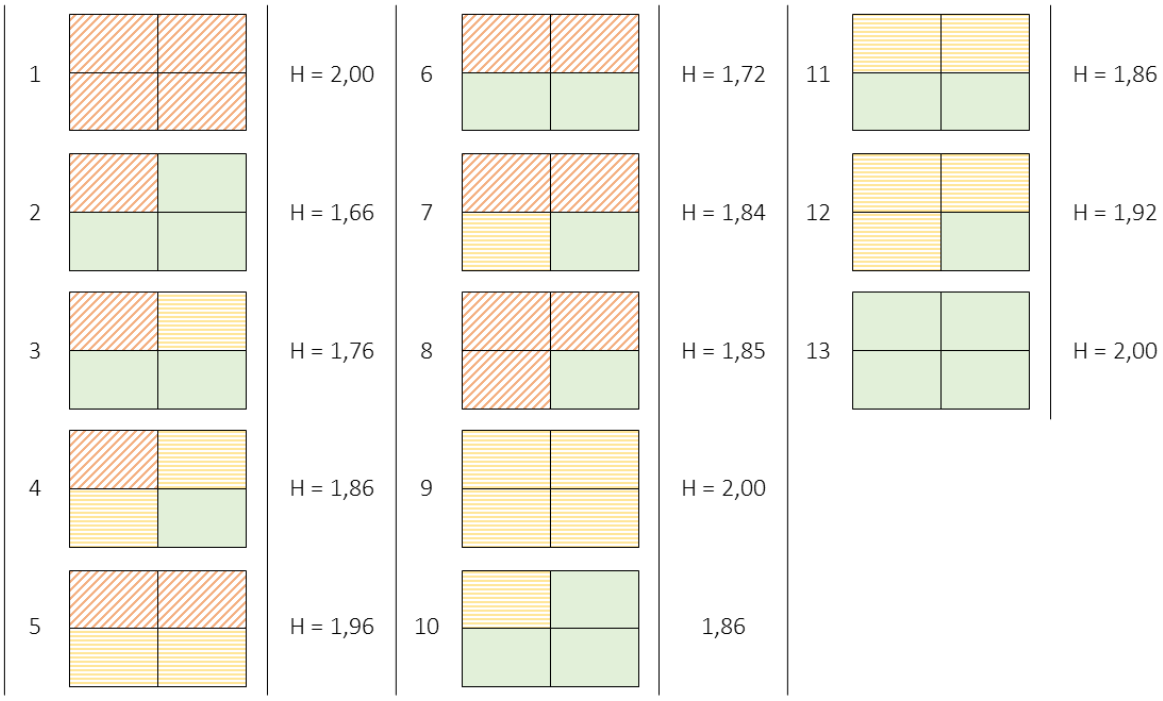

Figure 5: Illustrative overview of the possible combinations to divide 4 areas, and their corresponding entropy values. Similarly to the graph of population density levels, it is not the value of the land development that matters, but the value in relationship to its surrounding value. Due to this, all cases which have identical percentages across the 4 zones have the same entropy value. Additionally, this entropy value is also the maximum value, since it corresponds to a uniform distribution

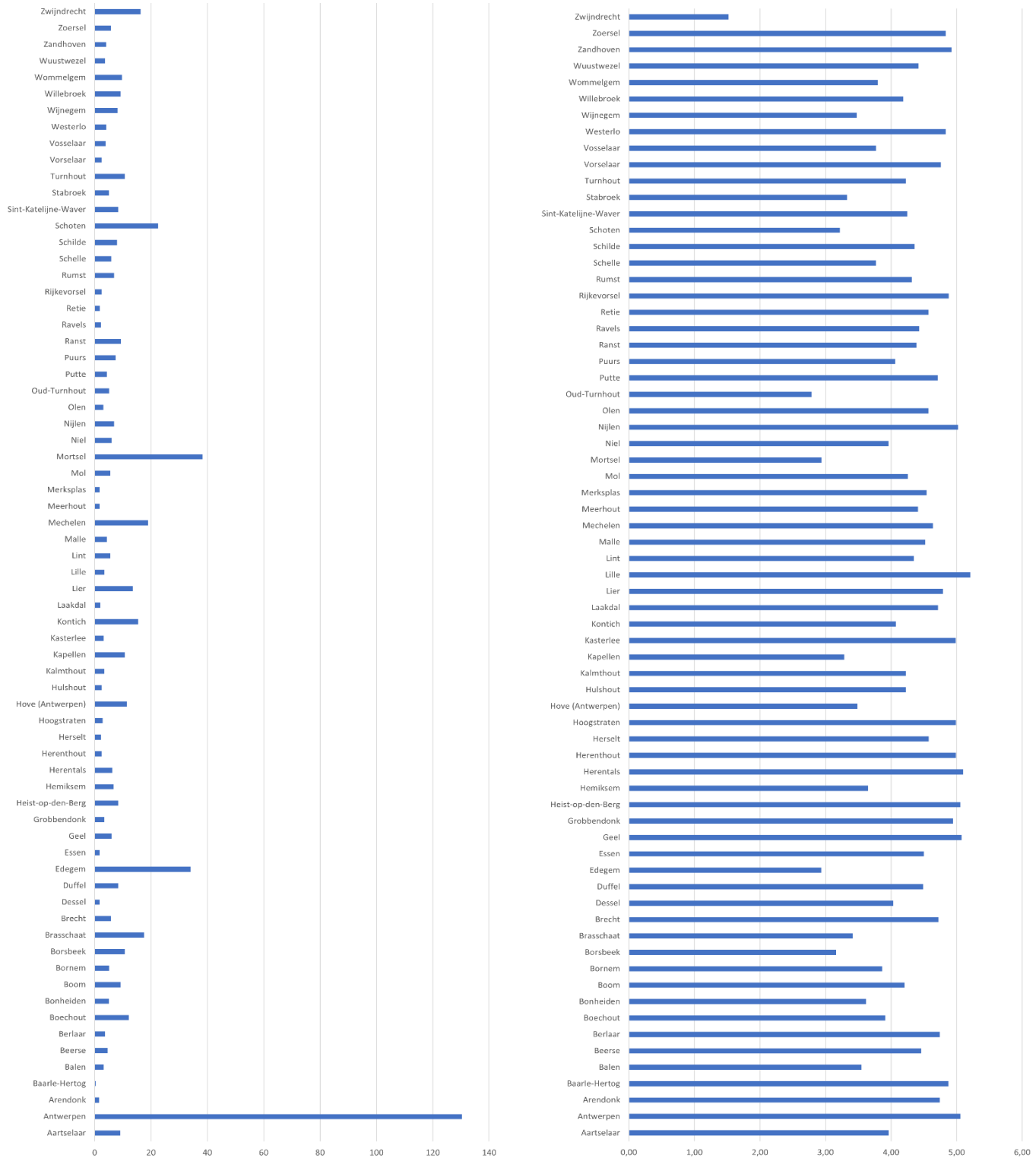

Figure 6: overview of flows of each city/town in Antwerp (left) and overview of flow entropy values of each city/town in Antwerp (right)

Author Contributions

Cristina Moyaert: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Validation, Visualization, Philippe Nimmegeers: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Validation, Visualization, Supervision. Bilal Mellouk: Writing – original draft, Conceptualization, Methodology, Formal Analysis, Bilal Mellouk: Writing – review & editing, Conceptualization, Methodology, Formal Analysis, Paul De Meulenaere: Writing – review & editing, Conceptualization, Methodology, Pieter Billen: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Validation, Resources, Supervision. All have read and agreed to the published version of the manuscript.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationship that could have appeared to influence the work reported in this paper

Acknowledgements

Philippe Nimmegeers holds a FWO senior postdoctoral fellowship (grant number: 1215523N) granted by FWO Vlaanderen/Research Foundation Flanders. This research is also part of the RESTEP FWO junior fundamental research project funded by the Research Foundation − Flanders − Grant No. G027122N.