



The Influence of the Zonation Effect on a System of Hierarchical Functional Regions

Samo Drobne, Mitja Lakner

Faculty of Civil and Geodetic Engineering, University of Ljubljana, Ljubljana, Slovenia

Abstract

Background: Hierarchical functional regions (FRs) can be calculated using data on interactions between basic spatial units (BSUs) and a hierarchical aggregation procedure. However, the results depend on the selected system of initial BSUs. In spatial sciences, this is known as the zonation effect, which is one of the effects of the Modifiable Areal Unit Problem (MAUP). **Objectives:** In this paper, we analyse the influence of the zonation effect on a system of hierarchical functional regions. **Methods/Approach:** We compared two systems of hierarchical functional regions of Slovenia modelled by the Intramax aggregation procedure using the inter-municipal labour commuting flows for the same year, but for two different initial sets of municipalities. Besides, we have introduced a new measure to compare systems of hierarchical FRs. **Results:** The results show that the zonation effect has an influence on hierarchical functional regions. The clustering comparison measure suggested here is a metric measure, which is appropriate for comparing hierarchical FRs. **Conclusions:** The zonation effect has influence on hierarchical FRs. The clustering comparison measure suggested in this paper is easy to interpret, but it should be adjusted for the number of clusterings.

Keywords: comparison, clustering measure, metric measure, SM-measure, average maximum proportion, hierarchical functional regions, municipalities, Slovenia

JEL classification: C38, C43, J40, R23

Paper type: Research article

Received: Feb 16, 2018

Accepted: Apr 21, 2018

Citation: Drobne, S., Lakner, M. (2018), "The Influence of the Zonation Effect on a System of Hierarchical Functional Regions", *Business Systems Research*, Vol. 9, No. 2, pp. 45-54.

DOI: 10.2478/bsrj-2018-0018

Acknowledgments: This work was supported by the research programme "Geoinformation Infrastructure and Sustainable Spatial Development of Slovenia" (P2-0227; 2018-2023), funded by the Slovenian Research Agency.

Introduction

In spatial sciences, the concept of functional regions (FRs) is one of the key concepts for analysing, modelling, monitoring, and predicting socio-economic structures. Brown and Hincks (2008) describe FRs as a combination of functionally complementing basic spatial units (BSUs), which have more economic interactions

with each other than with outside units. And, Johansson (1998) and Karlsson and Olsson (2006) define a FR as an area characterised by a high frequency of intra-regional economic interaction, such as labour commuting and intra-regional trade in goods and services, and an area of agglomeration of activities and transport infrastructure facilitating significant mobility of people, products, and information.

So, functional regions can be described as reasonably functioning spatial entities composed of economically and socially connected areas or basic spatial units like census units, statistical units, statistical local areas, settlements, communities, municipalities, postal zones, etc. Nevertheless, national systems of BSUs are changing as well: settlements are spreading; communities/municipalities are splitting into two or more new entities, etc. However, the change of the number (and the size and the shape) of the BSUs brought about the effects of the Modifiable Areal Unit problem (MAUP). MAUP was first described by Gehlke and Biehl (1934), but it was explained in detail not no sooner than half a century later. Openshaw (1984) proved that when different numbers, sizes, and shapes of zones are chosen in geographically based analyses, different results are generated.

MAUP includes the scale effect and the zonation effect (Openshaw, 1984). The scale effect describes the variation in results obtained when data for one set of BSUs are aggregated into larger aggregate spatial regions. On the other hand, the zonation effect is described as the variation in results obtained from different ways of subdividing geographical space at the same scale (Stillwell et al., 2014).

In this paper, we analyse the influence of the change of the system of municipalities in Slovenia in the period between 2000 and 2011 in hierarchically aggregated functional regions. In this period, the number of Slovenian municipalities changed twice. At the beginning of the analysed period, there were 192 municipalities in Slovenia. Two years later, in 2002, one new municipality was established. However, four years later, in 2006, 17 new municipalities were formed. Consequently, the number of municipalities grew from 192 to 193 and then to 210, i.e. by a total of 18 municipalities in the observed period. Therefore, it seems obvious to assume that different municipality sets generate different hierarchically aggregated FRs (with all other conditions unchanged).

Functional regions can be considered as clusterings that are compared by clustering comparison measures. Wagner and Wagner (2007) identified three sections of clustering comparison measures: measures based on counting of pairs of elements in both clusterings, measures based on summation of set overlaps, and measures based on mutual information. In this paper, we tested several clustering comparison measures already suggested in the literature about the comparison of systems of hierarchically aggregated FRs generated by different sets of municipalities. We also suggested and tested a new measure to compare various systems of hierarchical FRs.

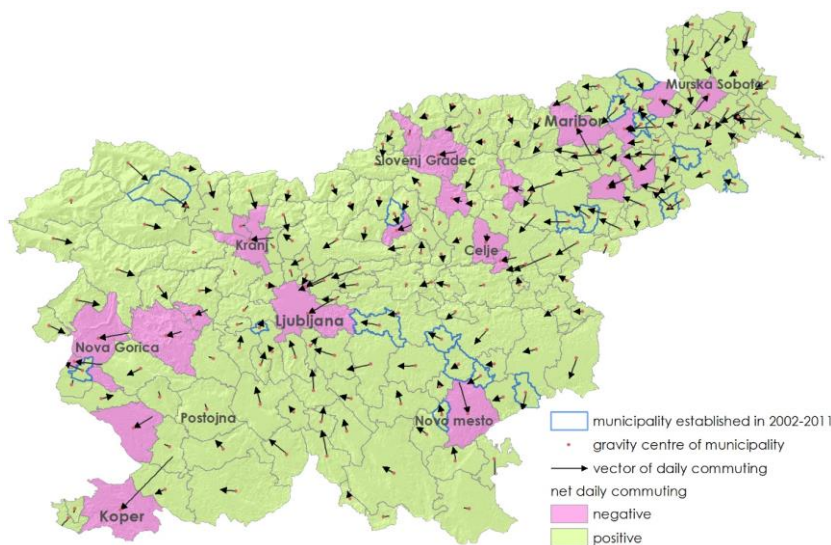
The paper is organized as follows. In the next section, we introduce the methodology, i.e. method to calculate hierarchically aggregated FRs, clustering comparison measures that have been tested, and the necessary data for the analysis. The results are presented and discussed in the following sections. The last section concludes the topic of the influence of the zonation effect on a system of hierarchical FRs and the topic of using the various clustering comparison measures to compare systems of hierarchically aggregated FRs.

Methodology

To analyse the influence of the zonation effect on a system of hierarchical FRs, two systems of hierarchical FRs were calculated for Slovenia: one considering the actual state in 2011, i.e. 210 municipalities and labour commuting flows between them (SMARS, 2016; SORS, 2016), and another one considering the same data on inter-municipal flows for 2011, but considering an older zonation from 2000 when there were 192 municipalities in the country. The analysis was done for the dimension of 210 municipalities from 2011. For this purpose, the database for 192 municipalities for 2000 was adequately expanded to the dimension of 210 municipalities for 2011. Figure 1 shows Slovenian municipalities in 2000 and 2011, respectively, and vectors of inter-municipal labour commuting for 2011. The methodology for calculations of vectors from interaction data was introduced by Jenko and Drobne (2014).

Figure 1

Municipalities in Slovenia in 2011, new municipalities established in 2002 and 2006, and vectors of inter-municipal labour commuting in 2011



Source: SMARS (2016), SORS (2016), and authors' own calculation.

Two systems of hierarchically aggregated FRs were calculated by the hierarchically aggregated procedure called the Intramax method. The Intramax method (Masser et al., 1975; Brown et al., 1990) is a popular method for modelling hierarchical FRs; there are plenty of recent studies that used this method (e.g. Drobne et al., 2014, 2015; Jaegal, 2013; Kohl et al., 2014; Koo, 2012; Landré et al., 2013; Mitchell et al., 2013), many others are listed in Drobne (2016). The hierarchical aggregation procedure, called Intramax procedure, seeks to maximise the proportion of the total interaction which takes place within the aggregations of BSUs, and thereby to minimise the proportion of cross-boundary movements in the system as a whole (Masser et al., 1980).

Systems of hierarchical FRs were compared by clustering comparison measures. For this purpose, we tested nine measures suggested in the literature, and we suggested and tested a new measure to compare systems of hierarchically aggregated FRs, as well. We analysed seven measures based on the counting of pairs of elements, which were General Rand Index (Rand, 1971), Adjusted Rand Index (Hubert et al., 1985), Fowlkes–Mallows Index (Fowlkes et al., 1983), Adjusted Fowlkes–Mallows Index (Fowlkes et al., 1983; Wallace, 1983), Mirkin Metric (Mirkin,

1996), Jaccard Index (Jaccard, 1912), Partition Difference (Li et al., 2004), one measure based on summation of set overlaps, i.e. the van Dongen Measure (van Dongen, 2000), and recently suggested measure based on mutual information, i.e. the Adjusted Mutual Information Index (Vinh et al., 2010). Definitions of the aforementioned nine measures are not explicitly provided here.

Along the nine clustering comparison measures previously introduced in the literature, we suggested a new concept of comparing FRs based on an average maximum proportion of matched municipalities in FRs (Drobne et al., 2017). Let X denote the finite set of n BSUs, in our case municipalities, $\{X_1, \dots, X_n\}$, where $|X| = n$, while FR denotes the system of N FRs $\{FR_1, \dots, FR_N\}$, of power $|FR| = N$, which consists of disjoint subsets X , so that their union is X . For a system of functional regions FR , let us assume that in each FR_i there is at least one municipality. $FR'_j = \{FR'_1, \dots, FR'_N\}$ denotes the second system of FRs consisting of the same n municipalities $\{X_1, \dots, X_n\}$. $M = [m_{ij}]$ denotes the cross-matrix with dimensions $N \times N$ of pairs FR_i, FR'_j :

$$m_{ij} = \frac{|FR_i \cap FR'_j|}{\max\{|FR'_i|, |FR'_j|\}}, \quad i, j = 1, \dots, N. \quad (1)$$

The suggested clustering comparison measure, SM -Measure, is the average value of N maximum values of matrix elements M . It expresses the average maximum proportion of matched municipalities in functional regions. If both systems of FRs, FR_i and FR'_j , are equal, then $SM = 1$. The SM -Measure was also tested by simulating and comparing 500 randomly generated partitions, where the average value of SM was 0.5, while the minimum value of SM -Measure in our test was 0.2.

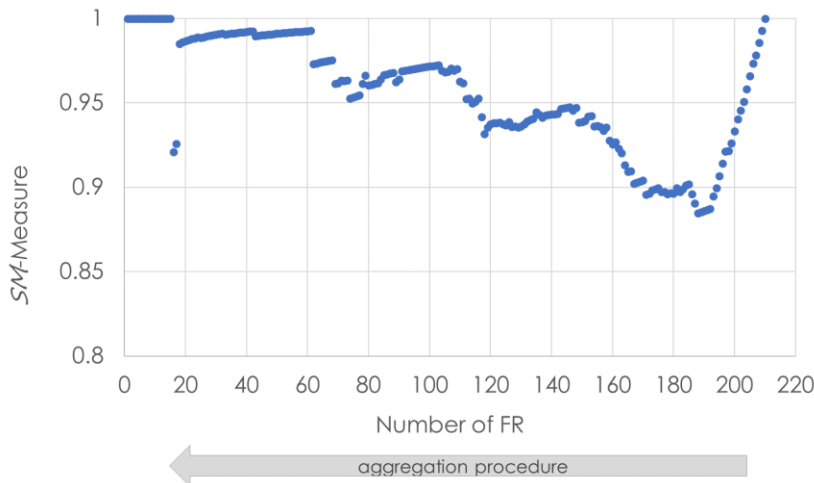
We calculated hierarchical FRs and clustering comparison measures in our own programme code in Mathematica 11.0 (Drobne, 2016; Drobne et al., 2017).

Results

The results on clustering comparison measures for comparing the systems of hierarchical FRs calculated using the same data on inter-municipal flows but for two different sets of municipalities show that the so-called zonation effect – in our case study, the change of municipalities – has an influence on the modelled FRs. It has a higher influence at the beginning of the hierarchical aggregation procedure on a higher number of FRs than at the end of the procedure on a smaller number of FRs; see Figure 2. The lowest value of SM -Measure is reached very quickly from the beginning of the aggregation procedure, just after the 22nd step that generates 188 FRs ($SM = 0.8847$). It means that the lowest proportion of matched municipalities in FRs is 88.5%. From this point, the proportion of matched municipalities in FRs in general increases. However, there are still local minimums where the differences between FRs are higher than a few aggregation steps before or after. Those are levels of 171, 118, 74, 62, and especially of 16 and 17, FRs. From Figure 2, it is obvious that, in general, the zonation effect has a much smaller influence on bigger FRs from inclusive 61 FRs in the country. However, there are two important levels of 16 and 17 FRs where the influence of the zonation effect increases suddenly: the proportion of matched municipalities in FRs decrease from 98.5% for 18 FRs to 92.6% for 17 FRs and then to 92.1% for 16 FRs. The last 15 aggregation steps generate equal systems of FRs.

Figure 2

SM-Measure for comparing the systems of hierarchical functional regions modelled by data on inter-municipal labour commuting for 2011 and for sets of municipalities in 2000 and 2011



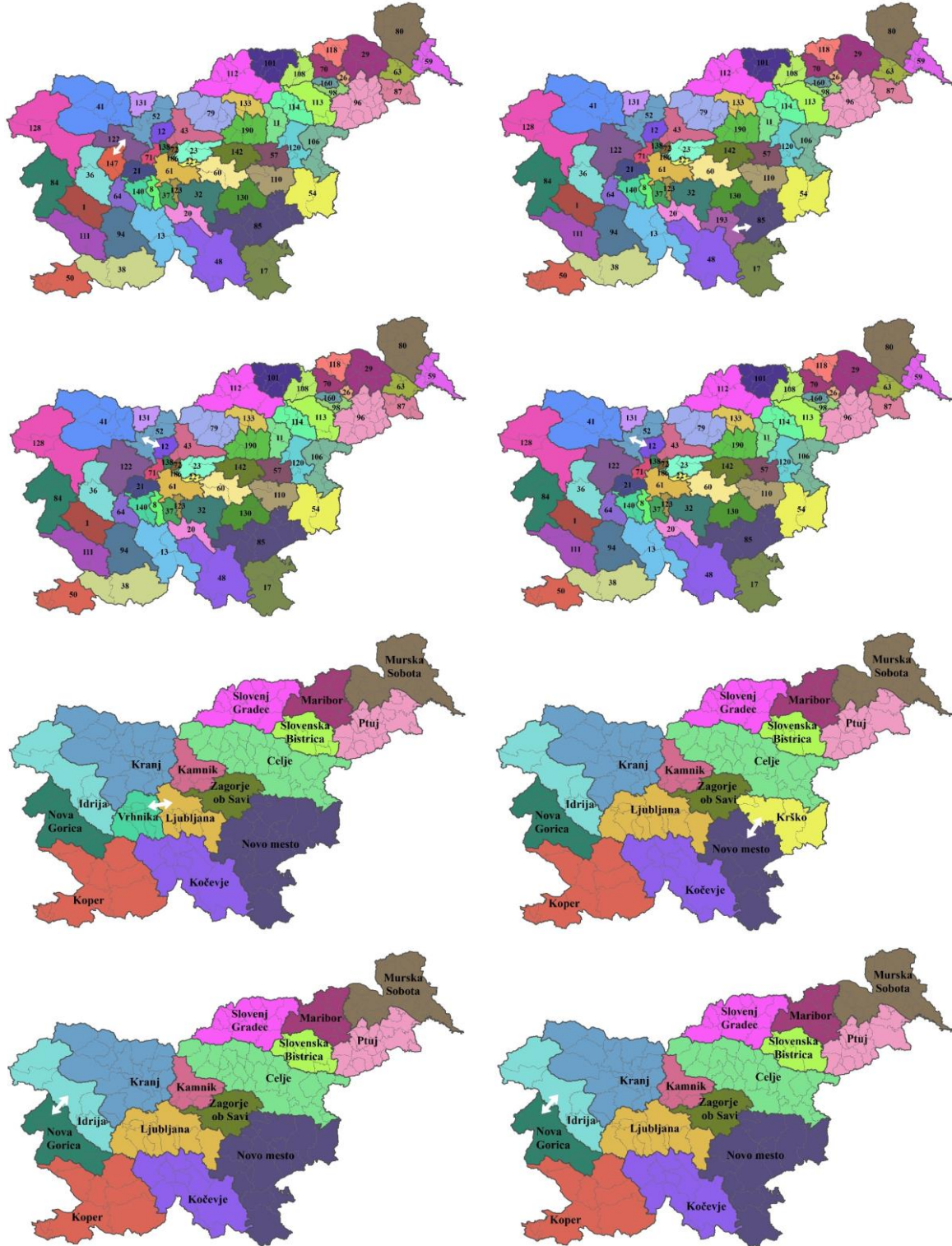
Source: Authors' work

Figure 3 shows 62 FRs before and 61 FRs after the 149th aggregating step when the proportion of matched municipalities increases suddenly and is higher than 99% up to 29 FRs, and 16 FRs before and 15 FRs after the 195th aggregating step when the proportion of matched municipalities increases suddenly and is 100% to the end of the aggregating procedure.

We analysed the influence of the change of municipalities on modelled hierarchical FRs by several clustering comparison measures. Some of them are better suited to compare systems of hierarchical FRs where each BSU should be considered as a FR. Figure 4 shows the results of tested clustering comparison measures. Rand Index counts correctly classified pairs of elements (Rand, 1971). However, Rand Index depends on both the number of clusters (FRs) and the number of elements (BSUs), as reported by Wagner and Wagner (2007); see also Fig 4a. Adjusted Rand Index is normalized for the number of clusters. But, in some cases, it may result in negative values (Meila et al., 1999), it is hard to interpret (Wagner et al., 2007), and it does not consider BSUs from the beginning of the aggregation procedure as FRs: Adjusted Rand Index at Fig. 4b starts from 0, but it should start from 1. Adjusted Fowlkes–Mallows Index (Fowlkes and Mallows, 1983; Wallace, 1983), see Figure 4d and Partition Difference (Li et al., 2004), see Figure 1g, consider BSUs from the beginning of the aggregation procedure as FRs, but, as reported by Wagner and Wagner (2007), the strong assumptions on the distribution of (Adjusted) Fowlkes–Mallows Index make it hard to interpret. Partition Difference is sensitive to cluster size and the number of clusters, and it is not normalized. The use of Mirkin Metric (Mirkin, 1996), see Figure 4e, Jaccard Index (Jaccard, 1912), see Figure 1f, and van Dongen Measure (van Dongen, 2000), see Figure 4h, show the same problems with sensitivity to cluster numbers, they are difficult to interpret, and/or not normalized. On the other hand, the AMI Index solves most of the aforementioned disadvantages, but, as reported by Romano et al. (2016), it is still hard to interpret, and, as seen from Figure 4i, it does not consider municipalities as a single FR. The last is solved by our SM - Measure, but, it has still an undesirable property, i.e. it is not independent of the number of FRs, see Figure 1j.

Figure 3

(a) 62 FRs in 2011, (b) 62 FRs in 2011 for the municipalities from 2000, (c) 61 FRs in 2011, (d) 61 FRs in 2011 for the municipalities from 2000, (e) 16 FRs in 2011, (f) 16 FRs in 2011 for the municipalities from 2000, (g) 15 FRs in 2011, (h) 15 FRs in 2011 for the municipalities from 2000

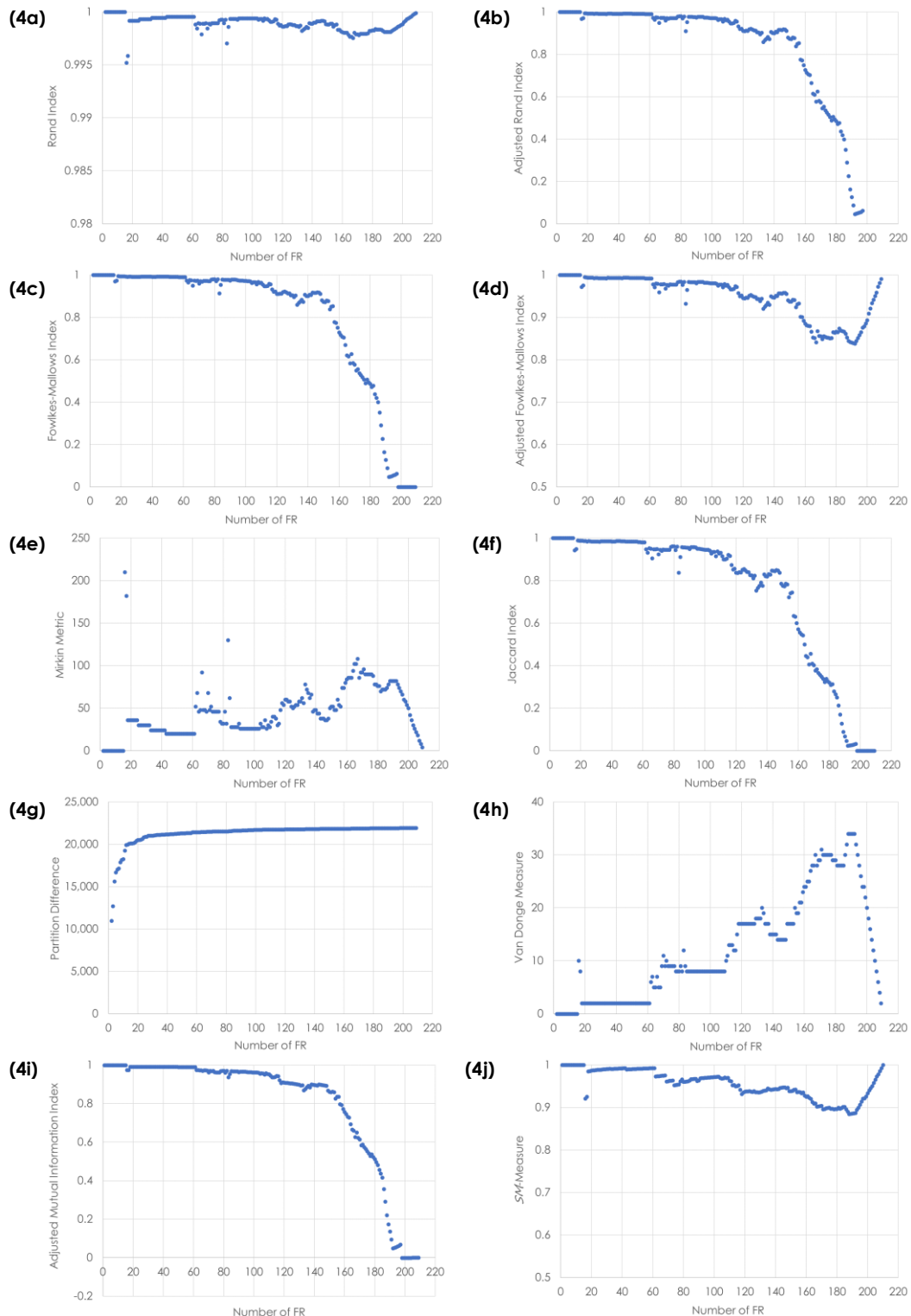


Source: Authors' work

Notes: The white arrow on the map denotes the next step of aggregation of functional regions. The codes in the Figures stand for central municipalities. Codes are explained in (SMARS, 2016).

Figure 4

Tested measures for comparing the systems of hierarchical FRs modelled by data on inter-municipal labour commuting for 2011 and for sets of municipalities in 2000 respectively 2011: (a) Rand Index, (b) Adjusted Rand Index, (c) Fowlkes–Mallows Index, (d) Adjusted Fowlkes–Mallows Index, (e) Mirkin Metric, (f) Jaccard Index, (g) Partition Difference, (h) van Dongen Measure, (i) Adjusted Mutual Information Index, (j) SM-Measure



Source: Authors' own calculation and Drobne and Lakner (2017).

Discussion and conclusions

In the paper, we analysed one of components of the so-called modifiable areal unit problem, i.e. the zonation component. It is known that a different zonation of basic spatial units can give different results when modelling larger aggregated areas (Openshaw, 1984). In the case study, we analysed the influence of different municipal systems on hierarchically aggregated functional regions for Slovenia. Our results are in line with the literature on MAUP effects. We proved that the zonation effect has an influence on hierarchical FRs, but the attitude of the influence varies according to the level of FRs. In the systems of the hierarchical aggregated FRs, there could be some levels where FRs are independent of the zone effect. If the researcher is focused on stable levels of hierarchical FRs, which do not change a lot or at all over a time, those levels of hierarchical FRs should be in the focus of the research. On the other hand, stable levels of hierarchical FRs can be considered as interesting, but not important, results, when the research is focused on changes over a time interval.

For comparison of FRs we tested several known clustering comparison measures and suggested a new measure. The *SM-Measure* suggested in this paper is a metric measure, so, it is easy to interpret, on the other hand, it is a normalized measure using the nominal [0, 1] range. Not all tested clustering comparison measures are appropriate for comparing the systems of hierarchical FRs, but the *SM-Measure* is. Our measure considers a single basic spatial unit as a functional region. The last property is especially important for evaluating systems of functional regions at the beginning of the hierarchical aggregation procedure. However, the *SM-Measure* has an unpleasant property, i.e. it is sensitive to the number of functional regions. Searching for the solution to this problem should be the focus of future research.

References

1. Brown, P. J. B., Hincks, S. (2008), "A Framework for Housing Market Area Delineation: Principles and Application", *Urban Studies*, Vol. 45, No. 11, pp. 2225-2247.
2. Brown, P. J. B., Pitfield, D. E. (1990), "The Intramax derivation of commodity market structures from freight flow data", *Transportation Planning and Technology*, Vol. 15, No. 1, pp. 59-81.
3. Drobne, S. (2016), "Model vrednotenja števila in območij funkcionalnih regij" ("A Model Evaluating the Number and Areas of Functional Regions"; in Slovene only), Doctoral thesis, University of Ljubljana, Faculty of Civil and Geodetic Engineering, Ljubljana.
4. Drobne, S., Bogataj, M. (2014), "Regions for Servicing Old People: Case study of Slovenia", *Business systems research journal*, Vol. 5, No. 3, pp. 19-36.
5. Drobne, S., Bogataj, M. (2015), "Optimal allocation of public service centres in the central places of functional regions", in 15th IFAC/IEEE/IFIP/IFORS Symposium on Information Control Problems in Manufacturing, INCOM 2015, IFAC, Ottawa, pp. 2449-2454.
6. Drobne, S., Lakner, M. (2017), "A concept of *SM-measure* to compare hierarchical clustering", in Zadnik Stirn, L. et al. (Eds.), *SOR '17 proceedings*, Slovenian Society Informatika, Section for Operational Research, Ljubljana, pp. 308-313.
7. Fowlkes, E. B., Mallows, C. L. (1983), "A Method for Comparing Two Hierarchical Clusterings", *Journal of the American Statistical Association*, Vol. 78, No. 383, pp. 553-569.
8. Gehlke, C. E., Biehl, K. (1934), "Certain effects of grouping upon the size of the correlation coefficient in census tract material", *Journal of the American Statistical Association*, Vol. 29, No. 185a, pp. 169-170.
9. Hubert, L., Arabie, P. (1985), "Comparing partitions", *Journal of Classification*, Vol. 2, pp. 193-218.
10. Jaccard, P. (1912), "The distribution of the flora in the Alpine Zone", *New Phytologist*, Vol. 11, No. 2, pp. 37-50.
11. Jaegal, Y. (2013), "Delineating Housing Market Areas in the Seoul Metropolitan Area Using

- a Geo-Computational Approach", Journal of the Association of Korean Geographers, Vol. 2, No. 1, pp. 7-20.
12. Jenko, D., Drobne, S. (2014), "Modeliranje prostorskih vzorcev delovne mobilnosti in selitev z vektorskimi polji" ("Modelling commuting and migration patterns using vector fields"; in Slovene only), in Ciglič, R., Perko, D., Zorn, M. (Eds.), Digitalni prostor, GIS v Sloveniji, 12, ZRC, Ljubljana, pp. 163-171.
 13. Johansson, B. (1998), "Infrastructure, Market Potential and Endogenous Growth", Jönköping International Business School, Jönköping (Mimeo).
 14. Karlsson, C., Olsson, M. (2006), "The identification of functional regions: theory, methods, and applications", The Annals of Regional Science, Vol. 40, No. 1, pp. 1-18.
 15. Kohl, T., Brouwer, A. E. (2014), "The Development of Trade Blocs in an Era of Globalisation", Environment and Planning A, Vol. 46, No. 7, pp. 1535-1553.
 16. Koo, H. (2012), "Improved Hierarchical Aggregation Methods for Functional Regionalization in the Seoul Metropolitan Area", Journal of the Korean Cartographic Association, Vol. 12, No. 2, pp. 25-35.
 17. Landré, M., Håkansson, J. (2013), "Rule versus Interaction Function: Evaluating Regional Aggregations of Commuting Flows in Sweden", European Journal of Transport and Infrastructure Research, Vol. 13, No. 1, pp. 1-19.
 18. Li, T., Ogihara, M., Sheng, M. (2004), "On Combining Multiple Clusterings", Proceedings of the ACM Conference on Information and Knowledge Management, Vol. 13, pp. 294-303.
 19. Masser, I., Brown, P. J. B. (1975), "Hierarchical Aggregation Procedures for Interaction Data", Environment and Planning A, Vol. 7, No. 5, pp. 509-523.
 20. Masser, I., Scheurwater, J. (1980), "Functional Regionalisation of Spatial Interaction Data: An Evaluation of Some Suggested Strategies", Environment and Planning A, Vol. 12, No. 12, pp. 1357-1382.
 21. Meila, M., Heckerman, D. (1999), "An Experimental Comparison of Model-based Clustering Methods", Proceedings of the Conference on Knowledge Discovery and Data Mining, pp. 16-22.
 22. Mirkin, B. G. (1996), "Mathematical classification and clustering", Kluwer Academic Press.
 23. Mitchell, W., Baum, S., Flanagan, M., Hannan, M. (2013). "CoffEE Functional Economic Regions", available at: http://e1.newcastle.edu.au/coffee/functional_regions/ (18 November 2015).
 24. Openshaw, S. (1984), "The Modifiable Areal Unit Problem", Geobooks, Norwich.
 25. Rand, W. M. (1971), "Objective Criteria for the Evaluation of Clustering Methods", Journal of the American Statistical Association, Vol. 66, No. 336, pp. 846-850.
 26. Romano, S., Vinh, N. X., Bailey, J., K. Verspoor (2016), "Adjusting for Chance Clustering Comparison Measures", Journal of Machine Learning Research, Vol. 17, pp. 1-32.
 27. SMARS (2016), "Digital data on territorial units of Slovenia, Data on municipalities, the Surveying and Mapping Authority of the Republic Slovenia, Ljubljana", available at: http://www.gu.gov.si/en/services/free_access_database/ (15 January 2016).
 28. SORS (2016), "Persons in employment (excluding farmers) by municipalities of residence and municipalities of workplace by sex, municipalities, Slovenia, annually), the Statistical office of the Republic Slovenia, Ljubljana", available at: <http://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=0723405E&ti=&path=../Database/Deomographics/07 labour force/05 labour force register/10 07234 working migrations/&lang=1> (15 January 2016).
 29. Stillwell, J., Daras, K., Bell, M., Lomax, N. (2014), "The IMAGE Studio: A tool for internal migration analysis and modelling", Applied Spatial Analysis and Policy, Vol. 7, No. 1, pp. 5-23.
 30. van Dongen, S. (2000), "Performance Criteria for Graph Clustering and Markov Cluster Experiments", Technical Report INS-R0012, Centrum voor Wiskunde en Informatica.
 31. Vinh, N. X., Epps, J., Bailey, J. (2010), "Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance", The Journal of Machine Learning Research, Vol. 11, pp. 2837-2854.
 32. Wagner, S., Wagner, D. (2007), "Comparing Clusterings - An Overview", Technical Report 2006-04, Faculty of Informatics, University of Karlsruhe.

33. Wallace, D. L. (1983), "A Method for Comparing Two Hierarchical Clusterings: Comment", Journal of the American Statistical Association, Vol. 78, No. 383, pp. 569-576.

About the authors

Samo Drobne is an Assistant Professor at the Department of Geodetic Engineering, Faculty of Civil and Geodetic Engineering, University of Ljubljana. His main research fields include spatial systems, geographical information systems, and spatial analysis and statistics. The author can be contacted at samo.drobne@fgg.uni-lj.si.

Mitja Lakner is an Associated Professor at the Department of Civil Engineering, Faculty of Civil and Geodetic Engineering, University of Ljubljana. His main research fields include dynamical systems, partial differential equations, and modelling. The author can be contacted at mitja.lakner@fgg.uni-lj.si.