



# Use of Constraints in the Hierarchical Aggregation Procedure Intramax

*Samo Drobne, Mitja Lakner*

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

## Abstract

**Background:** Intramax is a hierarchical aggregation procedure for dealing with the multi-level specification problem and with the association issue of data set reduction, but it was used as a functional regionalization procedure many times in the past. **Objectives:** In this paper, we analyse the simultaneous use of three different constraints in the original Intramax procedure, i.e. the contiguity constraint, the higher-inner-flows constraint, and the lower-variation-of-inner-flows constraint. **Methods/Approach:** The inclusion of constraints in the Intramax procedure was analysed by a programme code developed in Mathematica 10.3 by the processing time, by intra-regional shares of total flows, by self-containment indexes, by numbers of singleton and isolated regions, by the number of aggregation steps where a combination of constraints was applied, by the number of searching steps until the combination of constraints was satisfied, and by surveying the results geographically. **Results:** The use of the contiguity constraint is important only at the beginning of the aggregation procedure; the higher-inner-flows constraint gives singleton regions, and the lower-variation constraint forces the biggest employment centre as an isolated region up to a relatively high level of aggregation. **Conclusions:** The original Intramax procedure (without the inclusion of any constraint) gives the most balanced and operative hierarchical sets of functional regions without any singletons or isolated regions.

**Keywords:** hierarchical aggregation procedure, Intramax, constraint, functional regionalisation, functional region

**JEL classification:** C43, C44, J40, R23

**Paper type:** Research article

**Received:** Feb 12, 2016

**Accepted:** Jul 10, 2016

**Citation:** Drobne, A., Lakner, M. (2016), "Use of constraints in the hierarchical aggregation procedure Intramax", Business Systems Research, Vol. 7 No. 2, pp.5-22.

**DOI:** 10.1515/bsrj-2016-0009

## Introduction

Different actors understand the very concept of a region quite differently. In spatial sciences, a region is a delimited spatial system and an expression of an

organisational unity that differentiates it from another region (Abler et al., 1972). Haggett (1965) distinguished between formal and functional regions in general. A formal region is defined as the largest area over which a generalization remains valid (ibid.). A formal region is internally homogeneous. Formal regionalization is achieved by clustering basic data units (BDUs) at a low level (e.g., census units, statistical units, statistical local areas, settlements, communities, municipalities, postal zones) so as to minimize the between-units variance of one or more variables. In contrast to formal regions, a functional region (FR) is internally heterogeneous and causes mutual complementarity and independence. Ullman () defined a functional region as a region organized by horizontal relations in a space in a form of spatial flows or interactions between parts of the region (i.e., BDUs). So, a functional region can be understood as a generalized pattern of spatial interactions where interaction flows can vary a lot – from commuting and migration flows, journeys to school, shopping or recreation, traffic and passenger flows by land/sea/air, money flows, commodity flows, information flows, to gas/water/electricity flows, etc.

On the other hand, administrative regions are defined nominally by their borders and they are required to cover the whole of the respective territory homogeneously and to be of comparable size. In comparison with rigid administrative regions, functional regions are a product of interrelations, they are changing all the time with development of technology and with investments in space, they are quite diverse in terms of their size and population, and they may overlap as well as not fully cover the territory (Drobne and Bogataj, 2012a). According to Karlsson and Olsson (2006), a functional region is a region characterised by its agglomeration of activities and by its intra-regional (inner) transport infrastructure, facilitating a large mobility of people, products, and inputs within its borders. Smart (1974), Coombes et al. (1979), Ball (1980), Van der Laan and Schalke (2001), OECD (2002), and many others recognised the integrated labour market, in which intra-regional commuting as well as intra-regional job search and search for labour demand is much more intensive than the inter-regional counterparts, as the basic characteristic of a functional region. So, the identification and delineation of functional regions are commonly based on the conditions of local labour markets (LLMs; OECD, 2002).

Ball (1980), Casado-Díaz (2000), Andersen (2002), and others denoted that the standard administrative regions used by governments for policy making, resource allocation, and research do not provide meaningful information on the actual conditions of a particular place or region. As such, there has been a move towards the identification and delineation of functional regions.

A number of procedures for delimiting functional regions have been suggested in the literature. Farmer and Fotheringham (2011) identified three general classes of functional regionalisation procedures: hierarchical aggregation (e.g., Brown and Holmes, 1971; Masser and Brown, 1975, 1977; Slater, 1975; Cörvers et al., 2009), multistage aggregation (e.g., Coombes et al., 1986; Van der Laan and Schalke, 2001; Flórez-Revuelta et al., 2008), and central place aggregation (e.g., Karlsson and Olsson, 2006; Drobne et al., 2009, 2010a,b; Konjar et al., 2010). Besides those general classes, there are some other approaches to delineate functional regions (e.g., Farmer in Fotheringham, 2011; Fukumoto et al., 2013; Manley, 2014; Kim et al., 2015).

In this paper, we analyse the simultaneous use of three constraints in the hierarchical aggregation procedure Intramax (Masser and Brown, 1975, 1977), that is, the contiguity constraint, the higher-inner-flows constraint, and the lower-variation-of-inner-flows constraint. While the contiguity constraint is the often-used constraint in the Intramax applications, the two other here suggested and tested constraints have not been applied in this hierarchical aggregation procedure so far.

The paper is organized as follows. In the next section, we discuss the hierarchical aggregation procedure Intramax. In the third section, we introduce a methodology for analysing the simultaneous use of different constraints in the Intramax procedure. The results are presented and discussed in the following sections. The last section concludes the topic of using the various constraints in the Intramax procedure.

## Intramax

Masser and Brown (1975, 1977) developed the Intramax procedure for analysing the structure of flows in a square interaction matrix. In such a matrix, interaction flows are recorded within and between a single set of areas, zones, regions, or other BDUs. Masser and Brown (1977) emphasised two areas of application of the Intramax procedure; the first of these was seen in dealing with the multi-level specification problem and with the association issue of data set reduction, and the second in the functional regionalization procedure. The results of such a regionalization procedure are functional regions.

Intramax is a heuristic procedure and does not guarantee a global optimal solution to the partitioning problem where maximum interaction flows would stay in the regions and less would cross the regions' borders (Masser and Brown, 1977). However, the procedure seeks to maximise the intra-group shares of total interactions, which take place within the aggregations of BDUs that form the diagonal elements of the matrix (Masser and Brown, 1975). The procedure monotonically raises the internal flows of the consolidated areas by aggregating small BDUs/FRs with relatively high interconnections first.

There has been also some criticism of the Intramax procedure as a pure statistical procedure that does not allow fine-tuning of regions (e.g. Coombes et al., 1986; Casado-Díaz and Coombes, 2011; Watts, 2013). These authors reported two specific deficiencies of the hierarchical aggregation approach, which are the irreversibility of the groupings and the indeterminacy of the number of functional regions. The irreversibility of the groupings means that FRs cannot be disaggregated in the grouping procedure like in some other methods (i.e. CURDS procedure; Coombes et al., 1986; Coombes and Bond, 2008). The indeterminacy of the number of FRs means that the stop rule that defines the number of regions should be chosen arbitrarily – but this problem is not yet objectively solved in other methods.

However, Intramax's relative simplicity and its implementation in Flowmap software (de Jong and Van der Vaart, 2013) are the reasons that it has been used – even recently – for several different purposes (Drobne and Lakner, 2016): for labour market area delineation (e.g. Feldman et al., 2005; Koo, 2012; Watts, 2013; Landré and Håkansson, 2013), for housing market area delineation (Goetgeluk and de Jong, 2007; Brown and Hincks, 2008; Jaegal, 2013), for world trade block delineation (Poon, 1997; Kohl and Brouwer, 2014), for functional economic region delineation (Mitchell et al., 2007, 2013; Mitchell and Stimson, 2010; Mitchell and Watts, 2010), to identify possible administrative or statistical regions (Nel et al., 2008; Drobne and Bogataj, 2012a,b) or transport regions (Krygsman et al., 2009), in allocation analysis of services (Drobne and Bogataj, 2014, 2015), and so forth.

The Intramax procedure is a stepwise analysis. In each step of the aggregation two BDUs/FRs, whose interaction gives the highest value of the objective function, are grouped together, and the interaction between them becomes the internal (or intrazonal) interaction for the resulting FR. This new region now takes the place of the two parent BDUs/FRs in the next step of the analysis. Thus, with  $N$  basic data units, all BDUs are grouped together into one FR after  $N-1$  steps, and all interactions become intrazonal (Masser and Brown, 1975, 1977; Brown and Pitfield, 1990). The

procedure, as well as the results of the hierarchical aggregation, can be presented in a tree structure of a dendrogram.

The original objective function in the Intramax procedure, as suggested by Masser and Brown (1975), improved by Hirst (1977) and Masser and Brown (1977), and simplified by Brown and Pitfield (1990), is:

$$\max_{i \neq j} \left( \frac{t_{ij}}{t_{ij}^*} + \frac{t_{ji}}{t_{ji}^*} \right) \quad (1)$$

where  $t_{ij}$  is the observed value of the cell entry in the  $i$ th row and the  $j$ th column in interaction matrix  $T = [t_{ij}]$ , and  $t_{ij}^*$  and  $t_{ji}^*$  are the expected values that are calculated as a product of sums of the  $i$ th row and the  $j$ th column:

$$t_{ij}^* = \left( \sum_j t_{ij} \right) \left( \sum_i t_{ij} \right) \quad (2)$$

It should be noted that: (a) the standardization of the entries of the interaction matrix is not necessary, (b) the procedure maximizes the intra-regional (inner) share of total flows at each stage of the grouping process, and (c) the intra-regional (inner) flows, i.e. the values on the main sub-diagonal matrices of the partitioned matrix, should be taken into account in the row and column totals at each step of the aggregation procedure.

Masser and Brown (1975, 1977) applied the contiguity constraint, so only adjacent BDUs were considered for possible aggregation. However, Brown and Pitfield (1990) reported that the contiguity constraint had been introduced to restrict the search for potential pairings and that had served, primarily, to increase the computational efficiency of the procedure at that time (this had led to considerable savings in computer time, particularly where large data sets were involved).

Recently, Koo (2012) critically observed that the original Intramax algorithm tended to focus on the prominent flow with the greatest value of the ratio of observed flows and expected flows, rather than on maximizing the shares of inner flows. He suggested using a modified objective function in the Intramax procedure, which would focus more on the shares of inner flows. However, to arrive at acceptable results he had to apply a constraint in the procedure; actually, he applied two constraints: a contiguity constraint and an area-balanced constraint. He showed that, with respect to the inner flows, the use of a modified objective function in the constrained models gave better results than in unconstrained ones (ibid.).

## Methodology

In our research, we analysed the simultaneous use of three constraints in the original Intramax procedure (Masser and Brown, 1975, 1977). To test the use of constraints in the hierarchical aggregation procedure, we developed a programme code in *Mathematica 10.3*. Besides the contiguity constraint, we strictly applied the constraints that could be calculated solely by interaction flows. At each stage of the grouping process, we implemented the use of the objective function (1) and a chosen combination of constraints. The procedure seeks for the maximum value of function (1) until the chosen combination of constraints is satisfied.

The first constraint that we considered was the contiguity constraint,  $C$ , which ensures that only adjacent BDUs/FRs are grouped together:

$$C_{ij} = 1 \text{ when } BDU_i/FR_i \text{ and } BDU_j/FR_j \text{ are spatially contiguous,} \\ C_{ij} = 0 \text{ otherwise.} \quad (3)$$

Recently, Kim et al. (2015) reported that the implementation of the contiguity constraint could dramatically increase the complexity of the problem. Spatial contiguity is often translated into a network tree generation problem to check the validity of contiguity. Regions and their adjacency relationships are expressed as nodes and edges in terms of a graph, so that a region is verified as contiguous only if there is at least one path connecting all the spatial units within the region or if all the spatial units within the region are connected to the tree structure (ibid.). In our programme, the spatial contiguity is checked by the depth-first search algorithm as defined by Daras (2005).

The second constraint was the higher-share-of-inner-flows constraint, *HSIF*, which ensured that those BDUs/FRs were grouped together that gave a significant improvement according to intra-regional (inner) shares of total flows. *HSIF* forces seeking the maximum value of (1) until

$$SIF_{new} > \left(1 + \frac{1 - SIF_{old}}{10n}\right) \cdot SIF_{old}, \quad (4)$$

where  $SIF_{new}$  is a virtual share of inner flows after aggregation of two candidate BDUs/FRs,  $SIF_{old}$  is a share of inner flows before aggregation,  $1 + \frac{1 - SIF_{old}}{10n}$  is a parameter defined heuristically that ensures that the change of *SIF* is significant, and  $n$  is the dimension of the interaction matrix at each stage of the grouping process.

The third constraint that was tested in the Intramax procedure was the lower-coefficient-of-variation-of-inner-flows constraint, *LCVIF*, which ensured that the grouping of two BDUs/FRs gave more balanced FRs according to the (share of) inner flows. *LCVIF* forces seeking the maximum value of (1) until

$$CVIF_{new} < CVIF_{old}, \quad (5)$$

where  $CVIF_{new}$  is a virtual coefficient of variation of inner flows after aggregation of two candidate BDUs/FRs, and  $CVIF_{old}$  is a coefficient of variation of inner flows before aggregation.

We generated eight sets of FRs, namely: FRs aggregated by the original Intramax procedure without the use of any constraint (here and after original FRs) as well as seven sets of FRs modelled by the simultaneous use of constraints in the aggregation procedures. Sets of FRs are denoted by: ① – without the use of any constraint (using the original Intramax procedure), ② – with the use of the contiguity constraint, *C*, ③ – with the use of the higher-share-of-inner-flows constraint, *HSIF*, ④ – with the use of the lower-coefficient-of-variation-of-inner-flows constraint, *LCVIF*, ⑤ – with the simultaneous use of the contiguity constraint and the higher-share-of-inner-flows constraint,  $C \wedge HSIF$ , ⑥ – with the simultaneous use of the contiguity constraint and the lower-coefficient-of-variation-of-inner-flows constraint,  $C \wedge LCVIF$ , ⑦ – with the simultaneous use of the higher-share-of-inner-flows constraint and the lower-coefficient-of-variation-of-inner-flows constraint,  $HSIF \wedge LCVIF$ , and ⑧ – with the simultaneous use of all three constraints, i.e. the contiguity constraint, the higher-

share-of-inner-flows constraint and the lower-coefficient-of-variation-of-inner-flows constraint,  $C \wedge HSIF \wedge LCVIF$ .

The results of modelling FRs by using the original Intramax procedure and by using seven combinations of constraints were compared with the  $B_k^*$  index. Fowlkes and Mallows (1983) introduced a  $B_k$  index for comparing two hierarchical clusterings, but Wallace (1983) suggested an adjustment for the number of clusters that resulted in the  $B_k^*$  index. Note that, recently, Watts (2013) used an adjusted mutual information (AMI) index (Vinh, 2010; Vinh et al., 2010) for comparing FRs. However, the AMI index does not consider BDUs at the beginning of the aggregation procedure as FRs, so, by using it, we could not adequately compare sets of FRs from the first to the last step of the aggregation procedure. Using the  $B_k^*$  index solves that problem. The comparison of hierarchical clusterings was done in a programme code in Mathematica 10.3 as well.  $B_k^*$  index is metric and standardized  $[0,1]$ . It equals 1 when  $k$  clusters in each clustering correspond completely (Fowlkes and Mallows, 1983; Wallace, 1983).

The inclusion of constraints in the Intramax procedure was analysed by the processing time (PT; at computer Intel i7-4771 CPU @ 3.50GHz, RAM 16GB, WolframMark Benchmark Score: 1.78), by the share of inner flows, by size criteria that express the closure of FR, by the number of singleton regions, by notion about isolated region, by aggregation steps where a combination of constraints was applied, by searching steps until the combination of constraints was satisfied, as well as by surveying results geographically. Before performing each aggregation procedure, we stopped the local kernel.

The size criteria often used in the studies on functional regions are: job ratio, supply-side self-containment, and demand-side self-containment (Casado-Díaz, 2000; Van der Laan and Schalke, 2001; Casado-Díaz and Coombes, 2011; Landré and Håkansson, 2013). Job ratio,  $JR$ , is defined as the ratio between day-employed population in the region and resident employed population in the region:

$$JR_i = \frac{\sum_j t_{ji}}{\sum_j t_{ij}}. \quad (6)$$

Job ratio is the indicator, which, for a specific territorial unit, links the number or workplaces with the number of employed persons (according to residence). A job ratio higher than 1 indicates that there is more inflow workers than outflow ones. It is related to net in-commuting.

Goodman (1970) and Smart (1974) defined supply-side self-containment,  $SSSC$ , and demand-side self-containment,  $DSSC$ , as:

$$SSSC_i = \frac{t_{ii}}{\sum_j t_{ij}}, \quad (7)$$

$$DSSC_i = \frac{t_{ii}}{\sum_j t_{ji}}. \quad (8)$$

$SSSC$  is the share of jobs inside a region occupied by residents of this region, i.e. the internal commuting flow divided by the number of jobs;  $DSSC$  is the share of

residents employed within a region, i.e. the internal commuting flow divided by the number of employed residents (Landré and Håkansson, 2013). So, *SSSC* indicates to what extent a region offers employment to the employed living in that region, and *DSSC* indicates to what extent a region offers housing to the people working in that region (Van der Laan and Schalke, 2001).

Singleton regions (SRs) are BDUs that are aggregated just at the end of the procedure. An isolated region (IR) is a BDU that is aggregated with other BDUs/FRs very late in the hierarchical aggregation procedure. In the Intramax procedure, the largest BDU, with the highest shares of inner flows, the highest shares of in-coming flows, and most frequently with the highest shares of out-going flows, is aggregated as an isolated region very late in the procedure.

In the application, we analysed the inter-municipal labour commuting flows in 2011 in Slovenia. The dimension of the interaction matrix was  $N^2 = 210^2$ . The total number of steps to aggregate all municipalities into one region was  $N - 1 = 209$ . Out of a total of 44,100 cells in the matrix, there were 31,557 (71.56%) empty cells. 4390 (9.95%) cells recorded only 1 commuter, but only 250 (0.57%) cells' entries recorded 250 commuters or more. In 2011, there were 778,776 labour commuters in total (employed population), but only 388,376 (49.87%) of them commuted between municipalities. The maximum inflow of 109,884 labour commuters (28.29% of all inter-municipal flows) terminated in the largest employment centre of Slovenia, i.e. in the capital Ljubljana, while the outflow from Ljubljana was 16,027 labour commuters (4.13% of all inter-municipal flows).

## Results

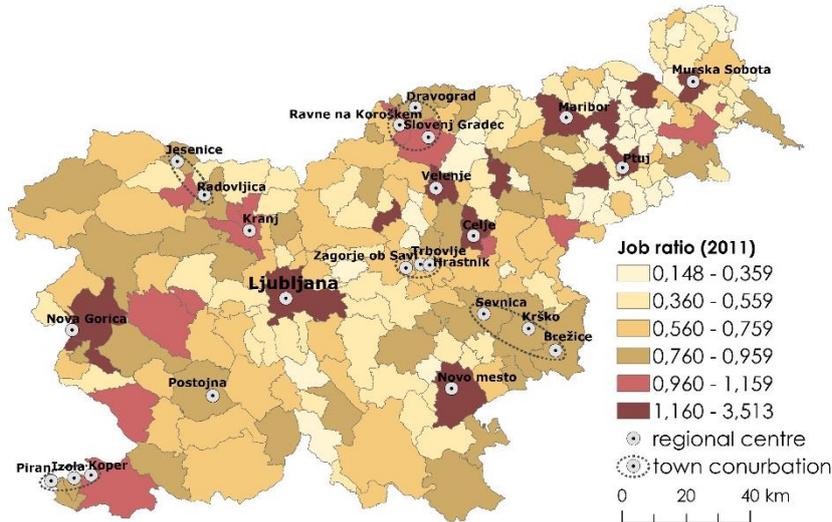
In Slovenia, a municipality, of which job ratio is more than 0.96, is labour-oriented, and others are residential-oriented (SORS, 2016). Figure 1 shows the job ratio for 210 municipalities in Slovenia in 2011. On the map, the Slovenian regional centres are denoted as defined in the Spatial Development Strategy of Slovenia (SPRS, 2004). The most workplaces and economic activities in Slovenia are concentrated in the (wider) urban areas of Ljubljana, Maribor, Celje, coastal conurbation Koper-Izola-Piran, followed by Kranj, Novo mesto, Velenje, and Nova Gorica (ibid). Figure 1 shows that there are some regional centres of Slovenia that are not labour oriented; those are some towns in the town conurbations like Dravograd, Piran and Izola, whole conurbations Zagorje ob Savi-Trbovlje-Hrastnik, Krško-Brežice-Sevnica, Jesenice-Radovljica, and urban centre Postojna.

The generated set of 2 to 209 FRs modelled by using the original Intramax procedure without the use of any constraint (1) and sets of FRs generated by the use of combinations of constraints (2–8) were compared by using the  $B_k^*$  index. Results show that FRs generated with the use of the higher-share-of-inner-flows constraint, *HSIF*, and its combinations generate FRs that differ the most from the original Intramax FR. The FRs modelled by using the lower-coefficient-of-variation-of-inner-flows constraint, *LCVIF*, or by its combination with the contiguity constraint,  $C \wedge LCVIF$ , differ less. However, FRs calculated by using the contiguity constraint only differ the least from the original Intramax FR in general; see Figure 2a where the sum of deviation of  $B_k^*$  index from 1 is shown. Figure 2b shows the differences between sets of FRs in more detail. The results of functional regionalization are equal for 1 and 2, except for 207 FRs.  $B_k^*$  index for 1–4 and for 1–6 show that FRs are equal up to 20 FRs where the largest employment centre of Slovenia, i.e. Ljubljana, is aggregated with other FRs. After that, *LCVIF* and  $C \wedge LCVIF$  give equal results up

to 4 FRs, from where  $C \wedge LCVIF$  gives equal FRs as the original Intramax procedure without the use of any constraints, while using just  $LCVIF$  gives spatially discontinuous FRs up to the last aggregation step. Anyhow, FRs modelled by the use of the higher-share-of-inner-flows constraint,  $HSIF$ , differ the most from the original Intramax FRs.

Figure 1

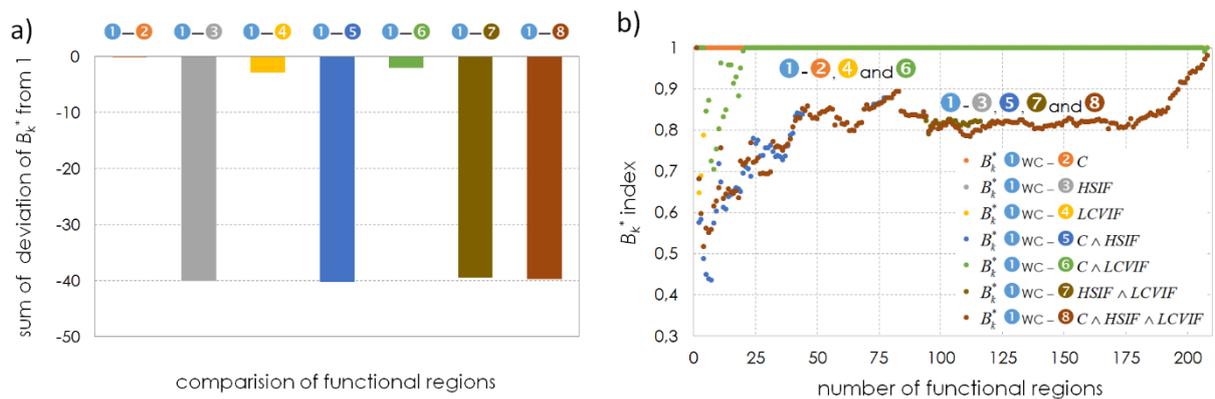
Job ratio in Slovenian municipalities in 2011



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

Figure 2

Comparison of functional regions: (a) general difference calculated by  $B_k^*$  index and (b)  $B_k^*$  index for the original Intramax functional regions and functional regions generated by the simultaneous use of constraints



Note: For better readability, the Y-axis at Figure 2b is limited at  $B_k^* = 0.3$ .

Source: Authors' own calculation.

Table 1 shows the statistics on modelling FRs in the hierarchical aggregation procedure Intramax with the simultaneous use of constraints. The fastest result is obtained without using any combination of constraints and the slowest one by using all three constraints simultaneously ( $C \wedge HSIF \wedge LCVIF$ ). In terms of single constraints,  $HSIF$  loads the processor the most and  $C$  the least.

Using  $HSIF$  and  $C \wedge HSIF$  gives singleton regions (BDUs that are aggregated just at the end of the aggregation procedure). Here, singleton regions are municipalities

with a very weak interaction (a very small relative number of outgoing and ingoing flows), mostly located at the country border (see also Figure 5b).

The occurrence of isolated region is not desired at the later stages of the aggregating procedure. Constraint *LCVIF*, and its combination with the contiguity constraint,  $C \wedge LCVIF$ , force Ljubljana to stay an isolated region even in the system of 13 FRs, while the original Intramax procedure aggregates it in the 190<sup>th</sup> aggregation step that gives 20 FRs. Using the *LCVIF* constraint also generates big FRs at the end of the aggregation procedure, which are spatially discontinuous for 2 to 4 FRs.

Table 1

Statistics on modelling functional regions using the Intramax procedure and constraints

Simultaneous use of constraints	PT (sec)	NSR	IR(Lj)	TNAS	FAS	LAS	NAS	TNSS	Geography
1 WC	13.70	0	21(0/0)	0	-	-	-	-	OV
2 C	13.97	0	21(0/0)	1	3	3	0	1	OV
3 HSIF	33.92	6	35(8/6)	203	1	203	0	4219	OV (but SR)
4 LCVIF	14.53	0	13(0/0)	16	190	208	3	63	OV (but IR and SD)
5 $C \wedge HSIF$	34.59	6	35(8/6)	203	1	203	0	4273	OV (but SR)
6 $C \wedge LCVIF$	14.31	0	13(0/0)	60	(3)190	206	(189)3	60	OV (but IR)
7 $HSIF \wedge LCVIF$	33.92	0	36(14/0)	208	1	208	0	4769	NOV
8 $C \wedge HSIF \wedge LCVIF$	35.45	0	36(14/0)	208	1	208	0	4823	NOV

Notes: WC – without constraint; SD – spatial discontinuity; C – contiguity constraint; HSIF – higher-share-of-inner-flows constraint; LCVIF – lower-coefficient-of-variation-of-inner-flows constraint; PT (sec) – processing time in seconds (together with reading data and calculating statistics); NSR – number of singleton regions; IR(Lj) – the notion about Ljubljana as an isolated region at the highest possible step of aggregation as a number of total FRs (the number of other IRs / NSR); TNAS – the total number of aggregating steps where a combination of constraints has been applied; FAS – the first step of aggregation where the combination of constraints has been applied; LAS – the last step of aggregation where the combination of constraints has been applied; NAS – the number of aggregating steps in the group of applied constraints (between FAS and LAS) where no constraints were applied; TNSS – the total number of searching steps (of the highest values of the objective function) until the combination of constraints was satisfied; Geography – a short notation about the geographical results of functional regionalization; OV – operationally valid; NOV – non-operationally valid.

Source: Authors' own calculation.

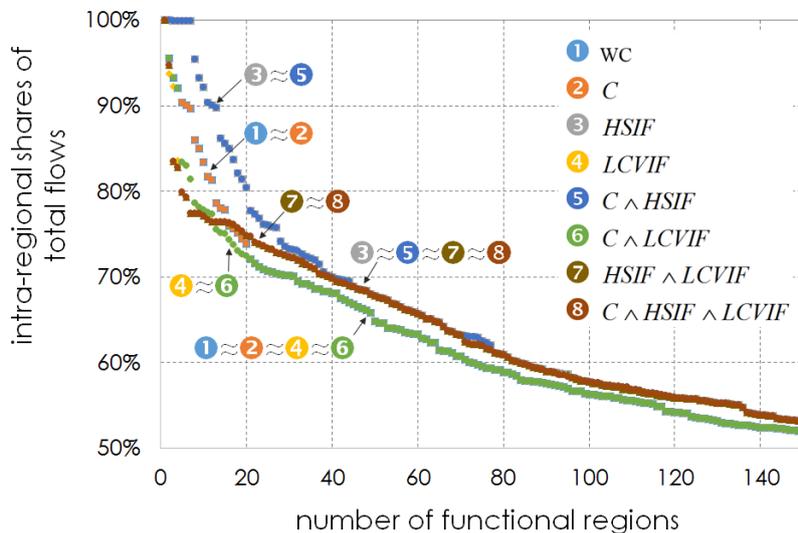
By registering a number of aggregation steps where a combination of constraints has been applied, we measured the deviation of the analysed procedure from the original Intramax one. Low deviation yields better aggregation results while two BDUs/FRs with higher relative interaction are amalgamated. The lowest deviation is obtained using solely the C constraint: in this case, two small, adjacent municipalities are forced to aggregate together just one step earlier (in the 3<sup>rd</sup> step) than in the

original procedure (in the 4<sup>th</sup> step). However, the results of the 5<sup>th</sup> step of aggregation are equal for both procedures. A small deviation from the original procedure is shown also when using *LCVIF* where only 16 steps of aggregation were forced no earlier than at the 190<sup>th</sup> step (moreover, later, 3 aggregation steps did not use *LCVIF*). The use of the  $C \wedge HSIF$  constraint that should ensure higher-share-of-inner-flows – which should be the main goal of the Intramax method – gives the maximum deviation from the original procedure. Here, 4219 (!) searching steps had to be done in the whole aggregation procedure of modelling 209 to 2 FR to satisfy the constraint. Consequently, the use of *HSIF* influences the results/statistics in combination with other constraints.

One of the most important attributes of FRs is the inner share of total flows. For this reason, the performance of constraints has been measured by the share of inner flows at each aggregation step; see Figure 3. In all cases, the use of the *HSIF* constraint gives the best results: it gives FRs with the highest share of inner flows. When *HSIF* is combined with *LCVIF* and  $C \wedge LCVIF$ , the share of inner flows becomes lower for a small number of larger regions (from 44 to 2 FRs). For a high number of small FRs (from 209 up to 21 FRs), performing the Intramax procedure without the constraint(s), the use of *C*, *LCVIF*, and  $C \wedge LCVIF$  constraints gives the same results. For 20 FRs and bigger ones, the original objective function as well as the use of the contiguity constraint only give FRs with higher inner shares of total flows than if using the *LCVIF* constraint.

Figure 3

Intra-regional shares of total interactions in relation to the number of functional regions (Intramax, simultaneous use of constraints, intra-municipal commuting, Slovenia, 2011)



Note: For better readability, X-axis is limited at 150 FRs.

Source: Authors' own calculation

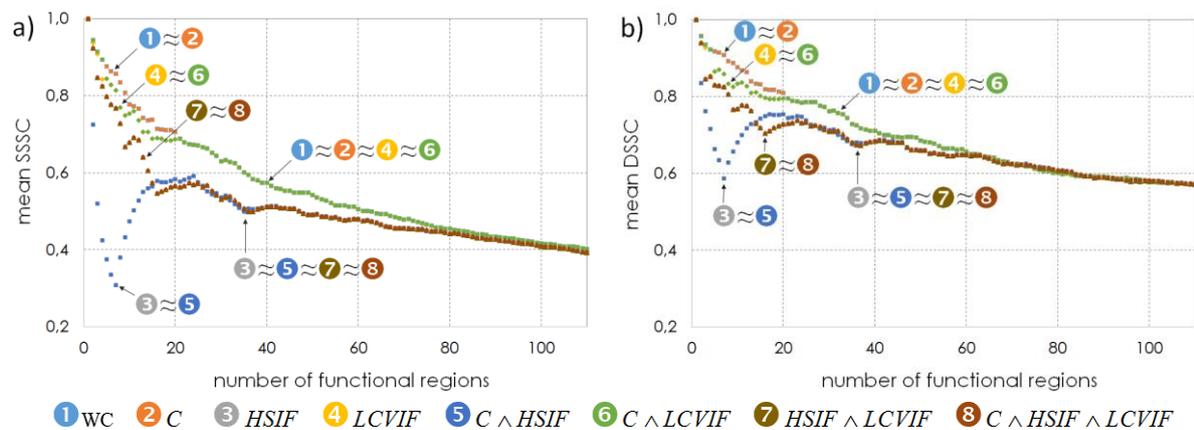
Besides inner flows, an important attribute of a FR is also its closure: the capability to what extent a region offers employment to the employed living in that region, *SSSC*, and the capability to what extent a region offers housing to the people working in that region, *DSSC*. Figures 4a and 4b show the mean *SSSC* and mean *DSSC* in relation to the number of FRs. For small FRs (more than 80 FRs) in the country, mean *SSSC* is very similar no matter what kind of constraints were used. For

bigger FRs, mean *SSSC* differs significantly depending on the use of constraints: original Intramax FRs as well as FRs generated by using *C*, *LCVIF* or their combination,  $C \wedge LCVIF$ , give more closed FRs than the use of the *HSIF* constraint or its combinations,  $C \wedge HSIF$ ,  $HSIF \wedge LCVIF$ , or  $C \wedge HSIF \wedge LCVIF$ . So, forcing regions with very high intra-regional shares of total interactions does not result in real, closed FRs. Even more, the closure of regions is worse.

The variation of mean *DSSC* is lower than the variation of mean *SSSC*, but relations are more similar than those for mean *SSSC*: the best results are obtained using the original Intramax procedure and the worse FRs are generated by using the *HSIF* constraint and its combinations.

Figure 4

Closure of functional regions: (a) mean supplied-side self-containment (*SSSC*) and (b) mean demand-side self-containment (*DSSC*) in relation to the number of functional regions (Intramax, simultaneous use of constraints, intra-municipal commuting, Slovenia, 2011)



Note: For better readability, points are connected and X axes are limited at 110 FRs.

Source: Authors' own calculation

## Discussion

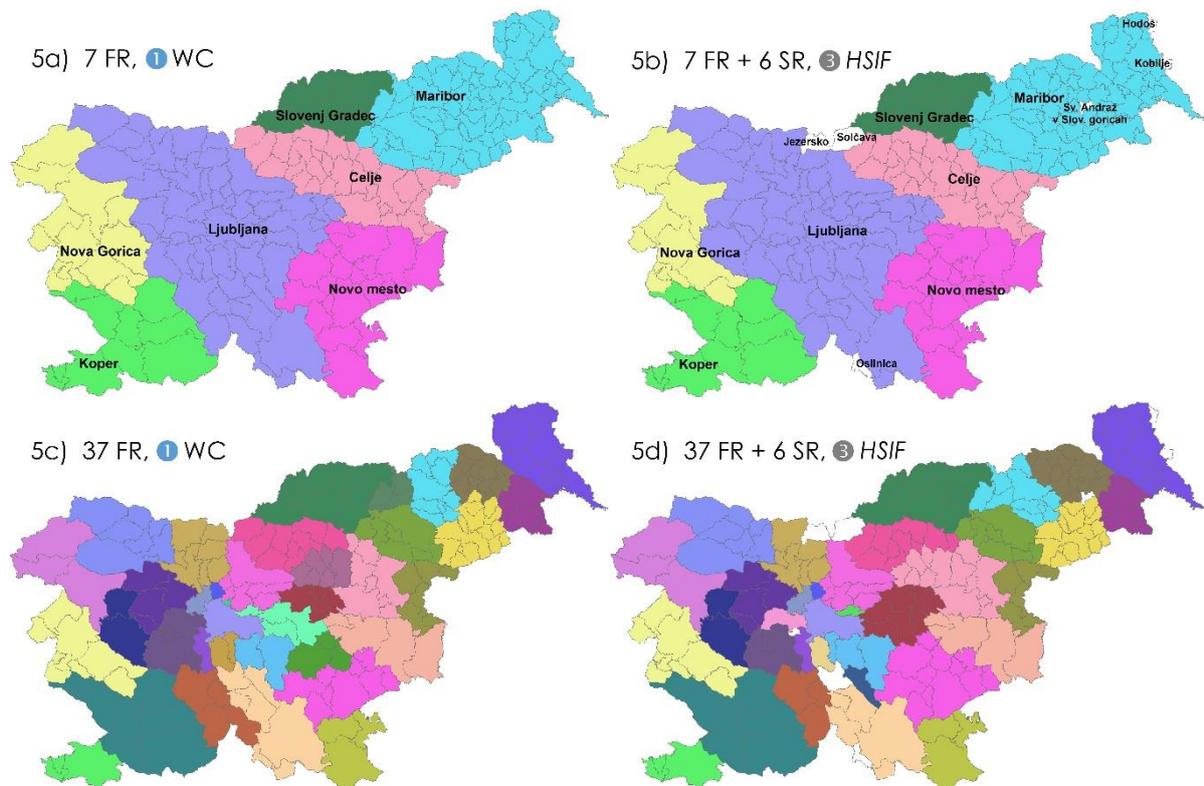
As already noted, the hierarchical aggregation procedure Intramax has some limitations: it does not guarantee a global optimal solution to the functional regionalization problem, it is an irreversible procedure and it does not allow directly controlling for the number of regions (Casado-Díaz and Coombes, 2011). The original Intramax procedure also lacks from delineating large urban areas that are disaggregated into smaller urban and adjoining suburban/rural FRs, as has been already shown in the literature (e.g., Masser and Scheurwater, 1980; Coombes et al., 1986; Feldman et al., 2005; Landré and Håkansson, 2013). This is shown by our result in Figure 5c. Otherwise, we can conclude that the Intramax procedure gives balanced FRs that are operationally valid. When using labour commuting data, the use of the spatial continuity constraint is not needed (the only difference is shown at the very early stage of the procedure). The Intramax procedure also solves the problem of small singleton regions that are aggregated in the early stages of the aggregation procedure. Consequently, there is no need for additional (subjective) decisions on the aggregation of singleton regions.

The use of the higher-share-of-inner-flows constraint gives better results than the original Intramax procedure regarding the share of inter-regional flows, but it

delineates singleton regions. For a small number of larger FRs, and if ignoring singleton regions, the original Intramax procedure and the procedure with the use of the *HSIF* constraint give similar results (compare Figures 5a and 5b), but smaller FRs are different (compare Figures 6a and 6b where 37 FRs (plus 6 singletons) are chosen arbitrarily).

Figure 5

(5a) 7 and (5c) 37 original Intramax functional regions (1 WC), (5b) 13 functional regions (7 functional regions plus 6 singleton regions), and (5d) 43 functional regions (37 functional regions plus 6 singleton regions) calculated with the use of the higher-share-of-inner-flows constraint (3 *HSIF*); inter-municipal labour commuting flows, Slovenia, 2011

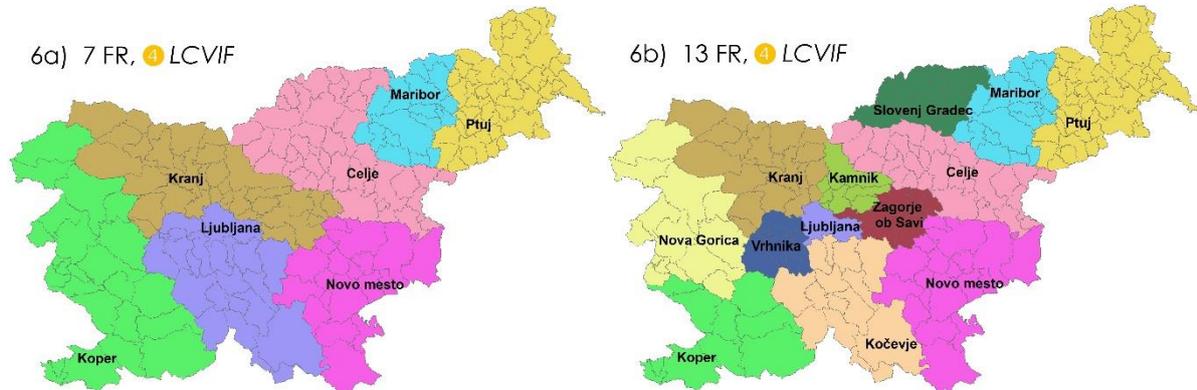


Note: In Figures 5b and 5d, the white colour denotes singleton regions.  
Source: SORS (2016) and authors' own calculation.

The *LCVIF* constraint forces to aggregate FRs with a similar share of intra-regional flows. For this reason, a BDU with a much bigger population than others (in our case, the capital Ljubljana) stays isolated for many aggregation steps before it is amalgamated with other FRs (see Figure 6b where Ljubljana is still an isolated region in the system of 13 FRs). The use of the *LCVIF* constraint delineates the metropolitan area of Ljubljana into small FRs as well (compare Figures 6a and 6b). The second important disadvantage of using *LCVIF* is the spatial discontinuity for 2 to 4 FRs – but this can be solved by using the combination  $C \wedge LCVIF$ . The combinations of  $HSIF \wedge LCVIF$  and  $C \wedge HSIF \wedge LCVIF$  constraints give geographically unexpected and unacceptable results for a small number of large FRs where the metropolitan area is disaggregated to ensure equal areas; see Figures 7a and 7b.

Figure 6

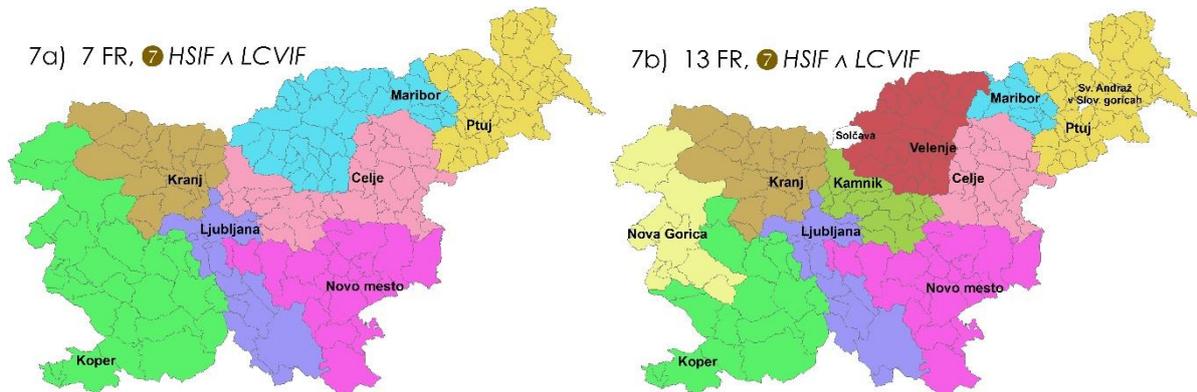
(6a) 7 functional regions and (6b) 13 functional regions; both calculated with the use of the lower-coefficient-of-variation-of-inner-flows constraint ( $\textcircled{4}$  *LCVIF*); inter-municipal labour commuting flows, Slovenia, 2011



Source: Authors' own calculation.

Figure 7

(7a) 7 and (7b) 13 functional regions calculated with the combination of the higher-share-of-inner-flows constraint and the lower-coefficient-of-variation-of-inner-flows constraint ( $\textcircled{7}$  *HSIF*  $\wedge$  *LCVIF*); inter-municipal labour commuting flows, Slovenia, 2011



Source: Authors' own calculation.

## Conclusion

In this paper, we have demonstrated the simultaneous use of three constraints in the hierarchical aggregation procedure Intramax. While the use of the spatial contiguity constraint was included in the procedure already from the very beginning (Masser and Brown, 1975), this is the first time that the other two constraints have been considered in the hierarchical aggregation procedure.

The comparison of the functional regions modelled by using a combination of three constraints (the spatial contiguity constraint, the higher-inner-flows constraint, and the low-variation-of-inner-flows constraint) and the functional regions modelled by using the original Intramax procedure, without using any constraints, was done using the  $B_k^*$  index (Fowlkes and Mallows, 1983; Wallace, 1983). In general, the results obtained by using the constraints differ from the aggregations derived by using the original Intramax procedure. The functional regions modelled by contiguity constraint differ the least, but the most different results are generated by the use of

the low-variation-of-inner-flows constraint. The original Intramax procedure gives fragmented large urban areas, but the low-variation-of-inner-flows constraint even more strictly delineates the metropolitan area into fragmented pieces.

The results show that, when using data on labour commuting, there is no need to include the contiguity constraint in the procedure. The use of the higher-inner-flows constraint generates singleton regions, and the lower-variation constraint forces a big basic data unit, as an isolated region, up to a relatively high level of aggregation.

We conclude that the original Intramax procedure delineates the most balanced and operative hierarchical sets of functional regions. Even more, it gives the most self-contained regions – which is one of the basic attributes of a functional region. It does not generate singleton regions as well – so, there is no need for (subjective) decisions on the aggregation of singleton regions.

## References

1. Abler, R., Adams, J. S., Gould, P. (1972), "Spatial organization", London, Prentice-Hall.
2. Andersen, A. K. (2002), "Are commuting areas relevant for the delimitation of administrative regions in Denmark?", *Regional Studies*, Vol. 36, No. 8, pp. 833–844.
3. Ball, R. M. (1980), "The use and definition of Travel-to-Work areas in Great Britain: some problems", *Regional Studies*, Vol. 14, No. 2, pp. 125–139.
4. 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.
5. Brown, L. A., Holmes, J. (1971), "The delimitation of functional regions, nodal regions, and hierarchies by functional distance approaches", *Journal of Regional Science*, Vol. 11, No. 1, pp. 57–72.
6. 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.
7. Casado-Díaz, J. M. (2000), "Local labour market areas in Spain: A case study", *Regional Studies*, Vol. 34, No. 9, pp. 843–856.
8. Casado-Díaz, J. M., Coombes, M. G. (2011), "The delineation of 21st century local labour market areas: a critical review and a research agenda", *Boletín de la Asociación de Geógrafos Españoles*, Vol. 57, pp. 7–32.
9. Coombes, M. G., Bond, S. (2008), "Travel-to-work areas: the 2007 review", London, Office for National Statistics, 58 pp, available at [http://www.istat.it/it/files/2014/12/final\\_TTWA\\_report.doc](http://www.istat.it/it/files/2014/12/final_TTWA_report.doc) (January 28, 2016).
10. Coombes, M. G., Dixon, J. S., Goddard, J. B., Openshaw, S., Taylor, P. J. (1979), "Daily urban systems in Britain: from theory to practice", *Environment and Planning A*, Vol. 11, No. 5, pp. 565–574.
11. Coombes, M. G., Green, A. E., Openshaw, S. (1986), "An efficient algorithm to generate official statistical reporting areas: the case of the 1984 travel-to-work-areas revision in Britain", *Journal of the Operational Research Society*, Vol. 37, No. 10, pp. 943–953.
12. Cörvers, F., Hensen, M., Bongaerts, D. (2009), "Delimitation and coherence of functional and administrative regions", *Regional Studies*, Vol. 43, No. 1, p. 19–31.

13. Daras, K. (2005), "An information statistics approach to zone design in the geography of health outcomes and provision", PhD dissertation, Newcastle, University of Newcastle.
14. De Jong, T., Van der Vaart, N. (2013), "Manual Flowmap 7.4.2", Faculty of Geographical Sciences, Utrecht University, The Netherlands, available at [http://flowmap.geo.uu.nl/downloads/FM742\\_Manual.pdf](http://flowmap.geo.uu.nl/downloads/FM742_Manual.pdf) (January 28, 2016).
15. Drobne, S., Bogataj, M. (2012a), "A method to define the number of functional regions: An application to NUTS 2 and NUTS 3 levels in Slovenia" = "Metoda opredelitve števila funkcionalnih regij: Aplikacija na ravneh NUTS 2 in NUTS 3 v Sloveniji", *Geodetski vestnik*, Vol. 56, No. 1, pp. 105–150.
16. Drobne, S., Bogataj, M. (2012b), "Evaluating functional regions", in Babić, Z. et al. (Eds.), 14th International conference on operational research, Trogir, Croatia, September 26–28, 2012, *Croatian operational research review*, Vol. 3, pp. 14–26, available at <http://hrcak.srce.hr/file/142254> (January 28, 2016).
17. 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.
18. Drobne, S., Bogataj, M. (2015), "Optimal allocation of public service centres in the central places of functional regions", *IFAC-PapersOnLine* Vol. 48, No. 3, pp. 2362–2367.
19. Drobne, S., Lakner, M. (2016), "Intramax and other objective functions", *Moravian Geographical Reports*, Vol. 24, to be appeared.
20. Drobne, S., Lisec, A., Konjar, M., Zavodnik Lamovšek, A., Pogačnik, A. (2009), "Functional vs. administrative regions: case of Slovenia", in Vujošević, M. (Ed.), *Thematic Conference Proceedings*, Vol. 1, Belgrade, Institute of architecture and urban & spatial planning of Serbia, pp. 395–416.
21. Drobne, S., Konjar, M., Lisec, A., Pichler Milanović, N., Zavodnik Lamovšek, A. (2010a), "Functional regions defined by urban centres of (inter)national importance – the case of Slovenia", in Schrenk, M. (Ed.), Popovich, V. V. (Ed.), Zeile, P. (Ed.), *Liveable, healthy, prosperous cities for everyone*, Real Corp 2010, *Proceedings 2010, 15th International conference on urban planning, regional development and information society*, May 18 – 20, 2010, Wien, Austria, pp. 295–304.
22. Drobne, S., Konjar, M., Lisec, A. (2010b), "Razmejitev funkcionalnih regij Slovenije na podlagi analize trga dela = Delimitation of functional regions of Slovenia based on labour market analysis", *Geodetski vestnik*, Vol. 54, No. 3, pp. 481–500.
23. Farmer, C. J. Q., Fotheringham, A. S. (2011), "Network-based functional regions", *Environment and Planning A*, Vol. 43, No. 11, pp. 2723–2741.
24. Feldman, O., Simmonds, D., Troll, N., Tsang, F. (2005), "Creation of a system of functional areas for England and Wales and for Scotland", paper presented at European Transport Conference, October 3–5, 2005, Strasbourg, France, available at: <http://abstracts.aetransport.org/paper/index/id/2284/confid/11> (January 28, 2016).
25. Flórez-Revuelta, F., Casado-Díaz, J. M., Martínez-Bernabeu, L. (2008), "An evolutionary approach to the delineation of functional areas based on travel-to-work flows", *International Journal of Automation and Computing*, Vol. 5, No. 1, pp. 10–21.
26. 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.
27. Fukumoto, J., Okamoto, Y., Ujiie, A. (2013), "A modularity approach to the delineation of functional regions from spatial interaction data", in *Proceedings of*

- The 13th World Conference on Transportation Research in Rio de Janeiro, Brazil, July 15–18, 2013, COPPE – Federal University of Rio de Janeiro, Brazil, available at <http://www.wctrs-society.com/wp/wp-content/uploads/abstracts/rio/selected/3377.pdf> (January 28, 2016)
28. Goetgeluk, R., de Jong, T. (2007), "What about the spatial dimension of subsidiarity in housing policy?", paper presented at ENHR 2007 International Conference on 'Sustainable Urban Areas', June 22, 2007, Rotterdam, Netherlands, available at <https://www.yumpu.com/en/document/view/26672955/what-about-the-spatial-dimension-of-subsidiarity-in-housing-policy> (January 28, 2016).
  29. Goodman, J. F. B. (1970), "The definition and analysis of local labour markets: some empirical problems", *British Journal of Industrial Relations*, Vol. 8, No. 2, pp. 179–196.
  30. Haggett, P. (1965), "Locational network analysis in human geography", London, Arnold.
  31. Hirst, M. A. (1977), "Hierarchical aggregation procedures for interaction data: a comment", *Environment and Planning A*, Vol. 9, No. 1, pp. 99–103.
  32. 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.
  33. 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.
  34. Kim, H., Chun, Y., Kim, K. (2015), "Delimitation of functional regions using a p-regions problem approach", *International Regional Science Review*, Vol. 38, No. 3, pp. 235–263.
  35. 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.
  36. 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, p. 25–35.
  37. Konjar, M., Lisec, A., Drobne, S. (2010), "Methods for delineation of functional regions using data on commuters", in Painho, M. (Ed.), Santos, M. Y. (E.), Pundt, H. (Ed.), *Geospatial thinking, Proceedings of the 13th AGILE International Conference on Geographic Information Science*, May 10–14, 2010, Guimarães, Portugal, pp. 1–10, available at <http://www.agile-online.org/Conference Paper/CDs/agile 2010/ShortPapers PDF/93 DOC.pdf> (January 28, 2016).
  38. 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.
  39. Krygsman, S., de Jong, T., Nel, J. (2009), "Functional transport regions in South Africa: an examination of national commuter data", paper presented at the 28th Annual Southern African Transport Conference, July 6-9, 2009, Pretoria, South Africa, pp. 144–154, available at [http://repository.up.ac.za/bitstream/handle/2263/11952/Krygsman\\_Functional%202009%29.pdf?sequence=1&isAllowed=y](http://repository.up.ac.za/bitstream/handle/2263/11952/Krygsman_Functional%202009%29.pdf?sequence=1&isAllowed=y) (January 28, 2016).
  40. 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.

41. Manley, E. (2014), "Identifying functional urban regions within traffic flow", *Regional Studies, Regional Science*, Vol. 1, No. 1, pp. 40–42.
42. Masser, I., Brown, P. J. B. (1975), "Hierarchical aggregation procedures for interaction data", *Environment and Planning A*, Vol. 7, No. 5, pp. 509–523.
43. Masser, I., Brown, P. J. B. (1977), "Spatial representation and spatial interaction", *Papers of the Regional Science Association*, Vol. 38, No. 1, pp. 71–92.
44. 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.
45. Mitchell, W., Stimson, R. (2010), "Creating a new geography of functional economic regions to analyse aspects of labour market performance in Australia", Working Paper No. 10-09, Centre of Full Employment and Equity, Newcastle, November 2010.
46. Mitchell, W., Watts, M. (2010), "Identifying functional regions in Australia using hierarchical aggregation techniques", *Geographical Research*, Vol. 48, No. 1, pp. 24–41.
47. Mitchell, W., Bill, A., Watts, M. (2007), "Identifying functional regions in Australia using hierarchical aggregation techniques", Working Paper No. 07-06, Centre of Full Employment and Equity, Newcastle, November 2007.
48. Mitchell, W., Baum, S., Flanagan, M., Hannan, M. (2013), "CofFEE Functional Economic Regions", AURIN project, Centre of Full Employment and Equity, Darwin, Australia, available at [http://e1.newcastle.edu.au/coffee/functional\\_regions/](http://e1.newcastle.edu.au/coffee/functional_regions/) (January 28, 2016).
49. Nel, J. H., Krygsmany, S. C., de Jong, T. (2008), "The identification of possible future provincial boundaries for South Africa based on an Intramax analysis of journey-to-work data", *Orion*, Vol. 24, No. 2, pp. 131–156.
50. OECD (2002), "Redefining territories – the functional regions", Paris, OECD Publishing.
51. Poon, J. P. (1997), "The cosmopolitanization of trade regions: global trends and implications, 1965–1990", *Economic Geography*, Vol. 73, No. 4, pp. 390–404.
52. Slater, P. B. (1975), "A hierarchical regionalisation of Russian administrative units using 1965–1969 migration data", *Soviet Geography*, Vol. 16, No. 7, pp. 453–465
53. SMARS (2016), Digital data on territorial units of Slovenia, Data on administrative services, Surveying and Mapping Authority of the Republic Slovenia, Ljubljana, available at <http://www.gu.gov.si/en/> (January 30, 2016).
54. Smart, M. W. (1974), "Labour market areas: uses and definition", *Progress in Planning*, Vol. 2, No. 4, pp. 239–353.
55. SORS (2016), Indeks delovne mobilnosti, občine, Slovenija, 2011 (Job ratio, municipalities, Slovenia, 2011, in Slovene only), Statistical office of Republic Slovenia, Ljubljana, available at <http://www.stat.si/obcine/Vsebinska.aspx?leto=2013&ClanekNaslov=TrgDelaMigracij&Indeks> (January 30, 2016).
56. SPRS (2004), "Spatial development strategy of Slovenia", Ministry of the Environment, Spatial Planning and Energy, Ljubljana, available at [http://www.mop.gov.si/fileadmin/mop.gov.si/pageuploads/publikacije/drugo/en/sprs\\_eng.pdf](http://www.mop.gov.si/fileadmin/mop.gov.si/pageuploads/publikacije/drugo/en/sprs_eng.pdf) (January 30, 2016).
57. Ullman, E. L. (1980), "Geography as spatial interaction", Seattle, University of Washington Press.
58. Van der Laan, L., Schalke, R. (2001), "Reality versus policy: the delineation and testing of local labour market and spatial policy areas", *European Planning Studies*, Vol. 9, No. 2, pp. 201–221.

59. Vinh, N. X. (2010), "Information theoretic methods for clustering with applications to microarray data", PhD dissertation, Sydney, The University of New South Wales.
60. Vinh, N. X., Epps, J., Bailey, J. (2010), "Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance", *Journal of Machine Learning Research*, Vol. 11, pp. 2837–2854.
61. 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.
62. Watts, M. (2013), "Assessing different spatial grouping algorithms: an application to the design of Australia's new statistical geography", *Spatial Economic Analysis*, Vol. 8, No. 1, pp. 92–112.

## About the authors

Samo Drobne is a Senior Lecturer 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](mailto:samo.drobne@fgg.uni-lj.si).

Mitja Lakner is an Assistant 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](mailto:mitja.lakner@fgg.uni-lj.si).