sciendo

The delimitation of planning regions on the basis of functional regions: An algorithm and its implementation in Turkey

Burak BEYHAN [a] [*]

Abstract

The proper delimitation of planning regions is a critical issue in the success of regional plans, and it constitutes a rich domain of research. In this paper, it is argued that planning regions should be based on functional regions – if the main intention is to increase the driving power of the people behind the planning process. Within this context, the aims of this paper are twofold: (1) to develop an algorithm (FRGIS) for the delimitation of planning regions on the bases of functional regions, and to implement it by using the scripting facilities available in Free and Open Source Software for Geographic Information Systems (GIS); and (2) to delimit the planning regions in Turkey by using FRGIS and the script developed for this purpose, by employing the commuting flows occurring between districts in the case country (Turkey) in 2010. The results show that FRGIS is successful in terms of the formation of spatially-balanced regions having higher levels of self-containment compared to those of existing regions. Nevertheless, it is also evident from this study that a combination of the nomothetic and ideographic methods of science is inevitable if functional regions are to be employed as planning regions.

Keywords: delimitation of regions; functional region; planning region; geographic information systems; Turkey

1. Introduction

One of the problems planners encounter during their engagement in a planning activity is the delineation of boundaries of the planning region (PR)[1]. In this respect, there are several alternatives for the delimitation of a PR. The first and easiest one, which has also been widely employed in Turkey (the case study country used here), is to designate administrative regions (normative regions) as PRs. Regional plans based on normative regions, however, are actually incapable of covering the geographical extent of the communities that are in interaction with each other, which eventually decreases the driving power of the people behind the planning process. This stems from the rigid characteristic of normative regions that may "continue to exist for reasons other than those that brought them into existence" (Wirth, 1937, p. 494). Indeed, while the number and magnitude of connections among the basic spatial units (BSU) involved in a country change over time, it becomes harder and harder for political authorities to change the boundaries of normative regions according to newly-formed relations between BSUs.

As a result of this situation, NUTS 1 and NUTS 2 regions of Turkey were defined in line with the provincial boundaries without taking into account the actual regional boundaries revealed by the interactions between districts. Parallel to this, it is observed that the majority of regional plans in Turkey have also been prepared for the normative regions whereas, as Wirth (1937), Tekeli (1972) and Geray (1997) argue, local administrative areas should be seen as supple tools for the accomplishment of a more beneficial socio-economic life. Thus, particularly from a planning point of view, the delimitation of PRs according to functional linkages, such as commuting patterns defining functional regions (FR) constitutes a substantial research area in terms of the establishment of geographical unity essential for the administration and planning of regions. The delineation of FRs on the basis of commuting flows has particularly become widespread during the last decades in many developed countries (Cattan, 2002).

In this regard, there have been several studies carried out for the delimitation of FRs or PRs in Turkey, considering

[a] Department of City and Regional Planning, Faculty of Architecture, Mersin University, Mersin, Turkey (*corresponding author: B. Beyhan, e-mail: *burakbeyhan@mersin.edu.tr* / *burakbeyhanmd@gmail.com*)

[1] Many acronyms are used in this paper: for reference, they are listed at the end of the work. Readers may wish to print a copy of these acronyms, as they are spelled out usually only at first time usage in the text.

the whole country. The first one was conducted by DPT (1982) on the spatial organisation of settlement units in Turkey and covers all the districts in the country. Yet, it is not updated and it does not allow for the formation of FRs without the designation of some cities as central places. Among recent studies, only Öztürk (2009) pays attention to flow data. In his study, Öztürk (2009) uses migration flows that are not actually suitable for the delimitation of PRs. In another recent study conducted by Ecemis Kılıç (2009), employing an eclectic method, it is observed that the tendency to define PR in accordance with the normative regions continues. Overall, the studies conducted for the delimitation of PRs or FRs in Turkey suffer from an absence of considering commuting data, as well as proper methods of analysis.

Given these considerations, the main objective of this paper is to develop an algorithm (named FRGIS) for the delineation of PRs via the delimitation of FRs, and to reveal PRs in Turkey by using FRGIS. For this purpose, firstly the concept of FR is elaborated, together with the key issues and methods of delimitation of FRs as PRs, in order to present the theoretical framework of the study. In the third section, the characteristics of the commuting and spatial databases used is presented in order to demonstrate the inherent characteristics of the respective databases. In this section, FRGIS and the script created for FRGIS are also elaborated by explaining the phases and stages of this rule-based method. In the fourth section, the results of implementation of FRGIS in Turkey are presented and discussed in order to reveal the degree of success of FRGIS. Subsequently, some concluding remarks are drawn with reference to both the theoretical considerations and the findings emerging from the functional regionalisation analysis for Turkey.

2. The key issues and methods of delimitation of functional regions as planning regions

The idea of FR focusing on interconnections between BSUs rather than on the similarities between them owes much to early studies revealing the influence of urban centres over their surrounding areas (see e.g. Platt, 1928; Christaller, 1933). Parallel to these studies, the need in regional planning for the delimitation of PRs in line with FRs first emerged for the planning of growing metropolitan areas by taking into account the geographical extent where the constituent parts of the respective cities are strongly connected to each other through daily interactions such as commuting relations (see e.g. Wirth, 1942; Friedmann, 1956). Subsequently, the first serious analytical efforts made for the delimitation of FRs were reported in the 1960s. In this respect, Nystuen and Dacey's (1961) study on the use of graph theory for the delineation of FRs deserves special attention. Again within the context of a systematic application of graph theory to the analysis of spatial structures, Haggett and Chorley (1969) show how FRs can be delimited by using network analysis.

One of the first studies aimed at the delimitation of FRs with particular reference to the employment of these regions for planning purposes was conducted by Brown and Holmes (1971). In the subsequent decades, the need for the delimitation of PRs in line with FRs is also manifested in a number of other studies (e.g. Hemmasi, 1980; Keeling, 1994; Van der Laan and Schalke, 2001; Feldman et al., 2005; Klapka et al., 2014; Halás et al., 2015; Erlebach et al., 2016; Drobne, 2017) parallel to improvements of the methods of delineation of FRs and PRs. The use of FRs as PRs also

received the attention of Coombes et al. (1986), the team developing the CURDS (Centre for Urban and Regional Development Studies) algorithm for the demarcation of FRs on the base of travel-to-work-areas (TTWA).

Considering the problem of the delimitation of FRs as a problem of the grouping of BSUs so that they can form different regions excluding each other, various studies (e.g. Masser and Brown, 1975; Coombes et al., 1986; Noronha and Goodchild, 1992; Halás et al., 2015) reveal that a FR should cover a group of BSUs whose interaction with other BSUs is organised in such a way that while the total amount of flows among the BSUs within the respective FR is maximised, the flows between them and other BSUs are minimised. The main expectation from this process is the creation of self-contained regions (Coombes et al., 1986). This basic rule constitutes the main rationale behind the delimitation of FRs. Thus, a FR is not something abstract, but rather "a reflection of the spatial behaviour of individuals in a geographic space" (Halás et al., 2015, p. 1175).

In this sense, an important issue in the delimitation of FRs as PRs is the scale of the regions concerned. FRs can be delimited by using daily travel-to-work flows at two different scales (Klapka et al., 2014; Erlebach et al., 2016); micro (e.g. labour market areas – LMA) and meso (e.g. NUTS 2 regions delimited on the base of local LMAs). In particular, at the meso-scale FRs can serve as a tool for regional planning (Erlebach et al., 2016). The main objective of FRGIS is also to expose those meso-scale regions by using a two-phase model. In addition to these scales, a macro level of region can also be defined by using other types of flows in the socio-economic system.

Since a FR is characterised by a high frequency of intra-regional economic and social interactions, spatial interaction matrices can be built by using not only labour commuting flows (Coombes et al., 1986; Andersen, 2002; Cörvers et al., 2009; Landré and Håkansson, 2013), but also migration flows (Brown and Horton, 1970; Holmes and Haggett, 1977; Hemmasi, 1980; Öztürk, 2009), housing markets (Hincks and Wong, 2010), telephone communication flows (Nystuen and Dacey, 1961; Clark, 1973; Clayton, 1980; Chi et al., 2016), commodity flows (Brown and Pitfield, 1990), inter-firm relational data (Van Oort et al., 2010), retail and wholesale shopping flows (Aydemir, 1978; DPT, 1982), banking relations and newspaper circulation (Clayton, 1980). Nevertheless, there is no doubt that the commuter flows are the most suitable database for the delimitation of PRs.

Another critical issue that is particularly relevant for the delimitation of PRs is the condition of adjacency. Internal contiguity is regarded as a major criterion in the delimitation of administrative regions and PRs (Andersen, 2002; Casado-Díaz and Coombes, 2011). Thus, if a FR is delimited for planning or administrative purposes, the formation of Multi-Polygon Regions (MPR) should not be permitted. Actually, one of the advantages of using commuting data for delimitation of FRs is that due to the restriction imposed by 'friction of distance' on the patterns of movement of people, the strongest interactions tend to occur between nearby BSUs (Coombes, 2000; Casado-Díaz and Coombes, 2011). As a result of this characteristic of commuting data, FRs formed after the analysis of interaction data are less likely to be fragmented.

Based on this finding, in FRGIS, particularly in the first phase of the delimitation of FRs, regions are formed on the base of a loose contiguity constraint by using Graph

Theoretical Geodesic Distance (GGD) between BSUs instead of a direct constraint. In graph theory, the distance between two vertices in a graph is the number of edges in a shortest path connecting them. This is also known as the geodesic distance. As evident from the definition, GGD does not correspond to physical distance. Actually, GGD can be considered as a measure of degree of contiguity and used as a parameter in the delimitation of FRs. Treatment of contiguity as a criterion is a common characteristic of the methods of delimitation of FRs as PRs.

There are various methods of delimitation of FRs. And there are different classification schemes for the respective methods. Coombes (2000) classifies these methods into three groups:

1. clustering methods;

2. methods using hierarchical algorithms;

3. rule-based methods.

In contrast, Erlebach et al. (2016) and Klapka and Halás (2016) create a separate category for graph-oriented approaches that are included by Coombes (2000) within clustering methods, and they group hierarchical algorithms under the clustering methods characterised by a numerical taxonomy approach. Rule-based procedures (also known as multistage methods) are again defined as a separate category by them.

In another classification scheme, graph-oriented approaches and hierarchical algorithms are classified by Klapka et al. (2014) and Halás et al. (2015) as sub-categories of (1) clustering and (2) rule-based methods, not as separate groups. Sub-categories of these two groups of methods are defined according to whether they are based on:

a. divisive or agglomerative;

b. hierarchical or non-hierarchical; and

c. numerical or graph-theoretical, procedures.

There are also other alternative classification schemes such as the ones developed by Van der Laan and Schalke (2001), Van Nuffel (2007), Casado-Díaz and Coombes (2011), and Farmer and Fotheringham (2011). A common theme observed for the first three schemes is the distinction between inductive and deductive approaches. The last two schemes, however, still pay attention to the distinction between rule-based algorithms and hierarchical clustering methods.

In this study, parallel to Klapka et al. (2014) and Halás et al. (2015), the methods of delimitation of FRs are classified into two main groups of methods based on (1) general cluster analysis and data reduction techniques and (2) specific algorithms relying on certain rules and stages. The preference to group cluster analysis and database reduction methods together stems from their reliance on pure statistical techniques. Application of purely statistical techniques, as Coombes et al. (1986, p. 946) remark, is largely deterministic and they exclude 'fine tuning' of the parameters that may be required to define FRs.

For those methods relying on statistical methodologies, Brown and Holmes' (1971) study employing the functional distance approach based on mean first passage time can be considered as one of the first examples of general cluster analysis by using various linkage measures. Keane (1978) and Cörvers et al. (2009) also employ this approach in their studies. The Intramax procedure developed by Masser and Brown (1975) on the basis of modifications introduced to Ward's (1963) hierarchical aggregation procedure, is

another example. This approach that was further improved by Masser and Scheurwater (1978) who benefit from some graph theoretical procedures was applied in a number of subsequent studies (e.g. Brown and Pitfield, 1990; Fischer et al., 1993; Feldman et al., 2005; Öztürk, 2009; Mitchell and Watts, 2010; Drobne and Lakner, 2016).

Similar cross-fertilisations of various procedures can also be observed for the database reduction methods based on eigenvalues, such as Correspondence Analysis and Principal Component Analysis (see e.g. Clark, 1973; Clayton, 1980). As a data reduction method, Factor Analysis can also be used for the delimitation of FRs (see e.g. Illeris and Pedersen, 1968; Hemmasi, 1980; Nader, 1980). The results of database reduction methods can be further processed by employing graph theoretic techniques in order to reveal FRs (see Goddard, 1970; Clark, 1973; Van Oort et al., 2010, for the employment of graphs as supporting visuals for the delimitation of FRs). More recent and comprehensive implementations of graph theoretic methods for the delimitation of FRs can be found in Farmer and Fotheringham (2011) and Chi et al. (2016).

In the second group of methods corresponding to rule-based procedures, there are two major approaches for the delimitation of FRs by employing commuting flows:

a. delimitation of FRs around some central BSUs (e.g. CURDS algorithm used in Coombes et al., 1986; Andersen, 2002; Pálóczi et al., 2016; European Regionalisation Algorithm (ERA) used in Coombes, 2000; the local labour market approach used in Karlsson and Olsson, 2006; Drobne et al., 2010; Konjar et al., 2010); and

b. delimitation of FRs without identifying any central BSU (e.g. partly the 'commuting zone approach' used in Karlsson and Olsson, 2006; the newest version of the TTWA algorithm considering all the individual zones as proto-TTWAs in the beginning, as described in Coombes and Bond, 2008; 'commuting aggregation approach' (CAA) used in Konjar et al., 2010; and the grouping evolutionary algorithm (GEA) developed by Martínez-Bernabeu et al., 2012; Casado-Díaz, et al., 2017).

In the first approach, after the identification of central BSUs according to some pre-defined criteria, other BSUs are assigned to their respective BSUs along with the degree of interaction between them and the central BSUs. The CURDS algorithm (Coombes et al., 1986) is a typical example. It is a multi-phase and stage aggregation method in which firstly foci zones are identified from the set of BSUs. In the subsequent phase, unallocated BSUs are assigned to these foci. In the final phase, it is ensured that all the resulting regions satisfy the specified constraints. Parallel to the CURDS algorithm, in ERA (Coombes, 2000), there are up to five steps of which first three steps are in line with the first two phases of the CURDS algorithm.

In the second approach, BSUs are assigned to each other without identifying any central BSU, according to magnitude of interaction between them. For this purpose, in the newest version of the TTWA algorithm (Coombes and Bond, 2008), after ranking all proto-TTWAs (BSUs) in terms of their size and self-containment values, whether the lowest-ranked proto-TTWA satisfies the requirements to be considered as a TTWA is checked. If it dissatisfies the requirements, it is dissolved into its constituent zones and, subsequently, each zone is grouped with the proto-TTWA with which it is most strongly linked. After the calculating the size and

self-containment values of altered proto-TTWAs, the whole step is repeated to satisfy the parameters specified for the formation of LMAs.

Other methods (e.g. CAA classified under the second group of approaches) group BSUs to form FRs in several phases/steps. In CAA, after calculating the share of people commuting from one BSU to another one, in the first phase, Konjar et al. (2010) group the couple of BSUs having the highest degree of interaction with each other according to maximum share. In the second phase, the groups of BSUs are fused according to a measure of mutual dependence between them. Accordingly, groups of BSUs defined in the first phase and actually corresponding to small local LMAs, are included in the second phase as the building blocks of the analysis.

In FRGIS, parallel to the first phase, the second phase of the functional regionalisation is also based on the demarcation of regions according to the maximum share – instead of a mutual dependence measure that seems to favour the centralisation of some groups of BSUs over others. The introduction of a minimum population limit for FRs also seems to support the domination of some BSUs. The most problematic issue in the delineation of FRs around some central nodes is the dismissal of the possibility of a region without a dominant core. Indeed, it is possible to define centres on the basis of interaction flows without using additional data. Nevertheless, if FRs are delimited around some central BSUs, the designation of these centres implies that there is a strict hierarchy in the spatial organisation of settlement units. This implication contradicts with contemporary reality of the 'network model' of places, in which horizontal relationships between places of similar size are possible.

Although the majority of the early rule-based methods were actually hierarchical, as Coombes (2000) remarks, some of them have been improved to break away from being hierarchical. For example, in CURDS proto-FRs dissatisfying the population size and self-containment criteria, are dismembered and reallocated to other FRs, to meet the specified constraints (Coombes et al., 1986). Similarly, in the latest TTWA algorithm (Coombes and Bond, 2008), those FRs with minimum validity are identified and disaggregated into their constituent BSUs, and subsequently they are associated with more dominant FRs. In a similar fashion, in FRGIS, parts of FRs not satisfying the internal contiguity criteria are disassembled and they are not allowed to be members of the same FR in subsequent iterations. Thus, the number of regions produced by rule-based methods is actually not known beforehand.

In rule-based methods, a threshold level for the degree of interaction can also be introduced for BSUs to combine with each other or other centres. For example, in the local labour market approach, Karlsson and Olsson (2006) use a cut-off frequency in order to exclude the very few long-distance commuters. These threshold values can also be estimated during the process of the definition of FRs rather than setting them in advance (Halás et al., 2015). Nevertheless, in FRGIS, this threshold value required in the first phase of the functional regionalisation is predefined by observing the results of an analysis for various values. This allows the researcher to fine-tune the parameters and to select the best set of FRs by comparing the results obtained for various runs of the script (RS). A different kind of fine-tuning procedure is also available in the TTWA algorithm (Coombes and Bond, 2008) for the prevention of MPRs.

Another parameter taken into account in FRGIS is the maximum area occupied by a region. Although in many studies (e.g. those using CURDS, ERA, TTWA and GEA algorithms) the number of inhabitants is used as a constraint, in relatively large countries where the population is not distributed uniformly, the use of population as a restriction may result in undesired outcomes. For example, in Turkey, İstanbul is the largest metropolitan region, overflowing into contiguous provinces. In terms of the planning of such regions, it would be wiser to include all BSUs involved in the region. If a restriction is imposed on the total population of a FR, even the existing provincial boundaries of İstanbul may not be preserved. Actually, the regions resulting from the introduction of a population threshold in the delimitation of FRs correspond to, what Noronha and Goodchild (1992) call, equitable regions that can be taken as a unit of analysis for statistical and administrative purposes, but may not be suitable for regional planning purposes.

In this respect, for comparative purposes, the R package 'LabourMarketAreas' developed by Franconi et al. (2016; 2017) for the implementation of the newest version of the TTWA algorithm described in Coombes and Bond (2008), is also used in this study for the delimitation of FRs in Turkey. The algorithm implemented in R is based on the following components required in order for a cluster to be considered as an LMA (Franconi et al., 2016, pp. 3–4):

1. a set of parameters for thresholds on the population size of the LMA and the level of self-containment (LSC) (main parameters are $minSZ$ (minimum number of employees for an LMA), $tarSZ$ (target value for the population size of the LMA), $minSC$ (LSC that is acceptable for cluster of large sizes), and $tarSC$ (the target self containment of an area in order for a small cluster to be considered an LMA));

2. a condition of validity to form valid LMAs;

3. a measure of cohesion between a BSU and the clusters;

4. a reserve list consisting of unassigned communities; and

5. an iterative procedure selecting one community at a time, aggregating it to a different cluster, and defining the operations to be implemented.

Overall, compared with other studies reported above, in the FRGIS developed in this study there is no implication for the designation of the centres in line with the contemporary reality of a 'network model' of places. FRGIS also differs from other studies by introducing a limit for the maximum area that can be occupied by a region. It is argued that this can serve for the formation of PRs better than a population limit. Last, but not least in importance, this study contributes to existing methods for the delimitation of FRs by developing an auto-control mechanism for the internal contiguity within a FR via a graph theoretical parameter that is integrated into FRGIS in such a way that no interruption is required by the user during the running of the algorithm, as such due to the employment of GIS.

3. The databases used and the algorithm developed for delineation of FRs as PRs

The database employed is based on travel-to-work data. It has been compiled from data obtained from the Social Security Institution (SSI) of Turkey. SSI was established by the SSI Law in 2006 and it brings the Social Insurance Institution (SSK), the Social Insurance Institution for the Craftsmen and Artisans and Other Self Employers

(BAGKUR), and the General Directorate of Retirement Fund (Emekli Sandığı) under a single roof. As some of the fields included in the BAGKUR database were either not up-to-date or unavailable, the database used covers travel-to-work data (as of April 17, 2010) for the Turkish labour force registered in SSK and Emekli Sandığı. Accordingly, the total number of commuters is 10,430,994. For this study, districts (NUTS4) are designated as BSUs, which allows us to observe the formation of FRs without any strict restriction imposed at the province (NUTS 3) level. Some districts have been either split or aggregated, however, in order to prevent mismatch between the residential and work addresses. As a result of this process, the total number of BSUs is determined as 939 (see Fig. 1). It is important to note that only 34 BSUs out of 939 BSUs (3.6%) have a self-containment (share of commuters residing and working in the same BSU) above 70%.

As all the districts in Turkey are included in the commuting database, the majority of the cells are actually empty in the matrix, showing the commuting interaction between origin BSU (OB) and destination BSU (DB). Accordingly, only 165,408 cells out of 881,721 cells have values greater than 0 (18.76%), which is, as noted by Brown and Holmes (1971), an expected feature of interaction matrices showing commuting relationships in a country. Commuting flows amounting to at most 50 occupy 95.13% of these filled cells. Yet, they only represent 6.91% of commuters. Conversely, although commuting flows amounting to at least 500 occupy only 1.15% of the filled cells, they represent 84.49% of commuters. Actually, this is also an expected characteristic of the interaction matrices representing the commuting flows in a country and reflects the tendency of people to commute between BSUs located close to each other. Indeed, more than half of the commuting flows (51.27%) occur within the same BSU, and the flows occurring between the BSUs located within a GGD of 3 from each other amount to 84.22% of the flows.

According to the method of delineation employed in this study, the first step in the delimitation of FRs in a country requires the calculation of both shortest distance (SD) and GGD between BSUs. For the calculation of SD, the OSM (Open-Street-Mapping) road database for Turkey and the 'Origin-Destination Matrix' tool in Network plug-in available in gvSIG CE (Community Edition), are employed. For the calculation of GGD, the algorithm developed by Beyhan (2012, pp. 32–35) is used. Before the calculation of GGD in FRGIS, physically-disconnected BSUs (such as islands) separated from the other BSUs because of the sea or channels, have been connected to others via narrow strips in order to prevent the exclusion of these BSUs in the calculation of GGD.

The algorithm, FRGIS (Fig. 2), developed in this study is translated into a java script that can be compiled into a JAR file running as a plugin, 'Functional Regionalisation', in OpenJUMP, a Free and Open Source Software (FOSS) for Geographic Information Systems (GIS). As the plugin is built on top of OpenJUMP, it can draw on a series of functions that facilitate the analysis required to delineate FRs (e.g. geometry converter functions and spatial predicates in JTS – Java Topology Suite). For example, in addition to the calculation of GGD, the condition for the prevention of MPRs and the total area occupied by a FR before including a new BSU inside the region, is automatically controlled. Overall, it is assumed that FRs can be delimited as PRs by using a two-phase model. In the first phase, no restriction is imposed for the total area of a FR in order to unite BSUs. In this phase, the model used in FRGIS is similar to the first phase of CAA (commuting aggregation approach).

In FRGIS, it is assumed that the commuting database is organised in an edge data format showing the labels for OB and DB together with SD and the amount of commuting (AC) between them, and lastly the commuting level (CL) measured as the share of those commuting from OB to DB as a percentage of the total commuters residing in OB. It is further assumed that the rows of the respective database are sorted according to firstly CL in descending order, secondly SD in ascending order, and thirdly AC in descending order. This sorting operation helps us prioritise the parameters taken into account in the creation of FRs. Nevertheless, some restrictions are also taken into account in FRGIS in order to prevent the formation of MPRs or overly-large FRs (such as the maximum GGD (GT) that can be allowed between OB and DB, and the maximum area (AT) that can be occupied by a FR). CL is mathematically defined as follows:

$$(1) \quad CL_{ij} = \frac{\text{number of commuters from BSU}_i \text{ to BSU}_j}{\text{number of commuters residing in BSU}_i} \times 100$$

where CL_{ij} is the commuting level between BSU_i (OB) and BSU_j (DB).
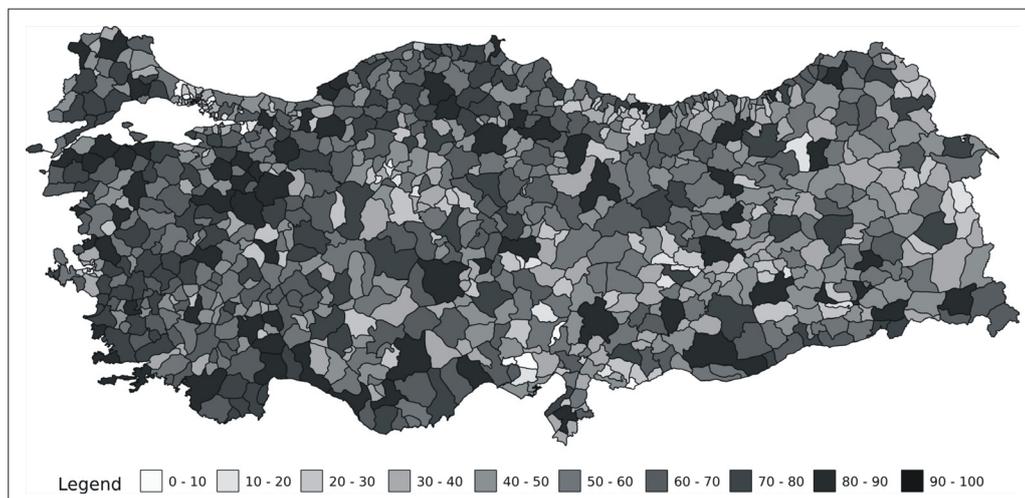

Fig. 1: The districts as BSUs according to the level of self-containment (% of total)
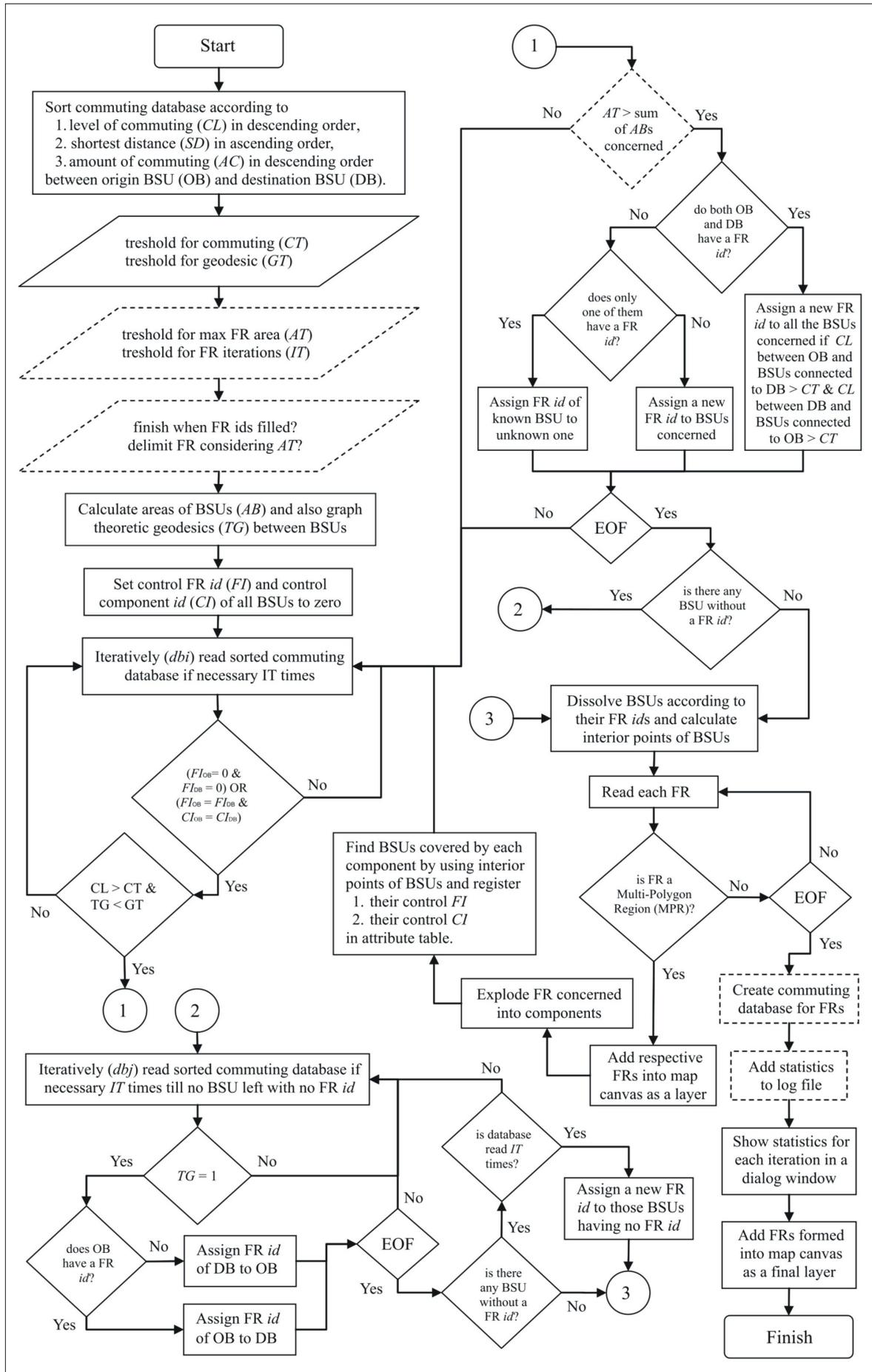*Source: author's elaboration*

Fig. 2: Algorithm (FRGIS) developed for delimitation of FRs as PRs
Source: author's elaboration

A threshold value (CT) for minimum commuting level is also required as a parameter to be defined for the formation of FRs. Together with GT, this parameter allows us to observe the tendency of BSUs to merge. In the first phase of running FRGIS, in the first stage, the database is read in sorted order, row by row, and all BSUs are allowed to freely unite with each other within the limits of a given GT and CT to form FRs. Accordingly, only those couples of BSUs whose TG (GGD between OB and DB) is less than GT and whose CL is more than CT are taken into account. As revealed in Figure 2, formation of the first FR actually starts with the union of the couple of BSUs in the database's first row showing the highest CL for any pair of OB and DB. The FR concerned enlarges with the addition of new BSUs provided that their couple in the row read is previously included in the respective FR as either OB or DB. A new FR is formed if both OB and DB of the row read are not included in the existing FRs. The formation of a FR covering n BSUs in the first stage can be mathematically expressed as follows:

$$(2) \qquad \text{first stage } FR = \bigcup_{i=1}^{n} BSU_i$$

for CL > CT and GT > TG

As some BSUs may not be involved in any FR because of CL values below CT, regardless of the phase, after the formation of FRs within the limits of a given GT and CT, in the second stage, the database is re-read in sorted order, if necessary IT (number of iterations) times, and each BSU that was not associated with a FR in the first stage is allowed to fuse with the adjacent BSU to which it is connected with the highest CL compared with CLs between it and its contiguous BSUs. Mathematically, this can be expressed for a FR covering n additional BSUs in the second stage as follows:

$$(3) \quad \text{second stage } FR = \text{first stage } FR \cup \bigcup_{i=1}^{n} BSU_i$$

for CL > other CLs and TG = 1

Thus, for each phase of running the script, FRs are actually formed by following a two-stage procedure. In the script, minimum CL (MCL) and the last line read in the database during these stages are registered, to be reported as statistics. Some statistics about BSUs connected to a FR in the second stage are also registered to the attribute table of the original map, together with the FR id (identity number) of each BSU (such as CL used to combine BSU concerned with the respective FR).

In FRGIS, each FR is also checked whether it is a MPR. If it is a MPR, each component (isolated part) in the FR is identified and BSUs covered by the FR are marked both with an component id assigned to the component covering them, and also with the FR id enveloping them. In the subsequent iterations of FRGIS for dbi (database iteration number) in Figure 2, those BSUs marked with the same FR id but different component ids are not allowed to be involved in the same FR. As MPRs are not desired, commuting statistics are not calculated and reported for the iterations involving them. It is assumed that the script will run until no MPR is formed within the given limits of IT. The main outcome obtained after RS is a report dialog window showing some statistics and a series of maps showing FRs and MPRs for each iteration, together with some basic statistics about them in the attribute table of the final FR map and original map. The results of the analysis together with the parameters used in them are also logged to a file after each RS.

4. The results of running FRGIS

Statistics obtained from various RS for the first phase are presented in Table 1. In the first three RS, BSUs are joined according to CL between only adjacent BSUs. If GT = 1, it is observed that the number of FRs is at least around 185. Nevertheless, after several other RS by relaxing the initial condition for adjacency, it is observed that the number of FRs can be reduced to 147 by allowing the couple of non-adjacent BSUs to be members of the same FR while preserving the spatial integrity of the FR in terms of prevention of MPRs and at the same time increasing LSC. From Table 1, it can be concluded that at the 12th RS for which both CT and GT are set to 4, MCL (for 'If GT = 1'), the minimum LSC, and the minimum area occupied by a FR can be maximised to better levels compared with other RS for various configurations of CT and GT. Thus, 147 FRs created after the 12th RS are considered to be the best candidates for the delimitation of PRs in the second phase with the introduction of AT (see Fig. 3).

Some statistics for the respective FRs are also given in Table 2, together with the statistics for BSUs before analysis. It is important to note that the minimum LSC increases from 2.90% (for a BSU) to 41.64% (for a FR). The average LSC also increases from 51.44% to 77.19%. In this respect, FRs that have been identified during the first phase of the regionalisation procedure partly conform to the criteria set in these recent studies. Indeed, in 117 out of 147 FRs, LSC exceeds 70% (see Tab. 2).

In the first phase, as explained above, AT is not used in the delimitation of FRs. The results of the analysis, however, show that there are huge differences between the largest and smallest FRs delimited during the first phase (Tab. 2). It is also observed that while the most crowded FR involves 21 BSUs, each of 30 FRs involves only 2 BSUs (see Tab. 3).

For both administrative and planning purposes, the area covered by a PR has critical importance owing to a number considerations: allocation of central funds to the administrative and planning units; organisation and supply of infrastructure services; and security issues. As the principal interest of this study is to identify regions that can be mainly used for regional planning purposes, one can speak of an optimum for the number of regional units and the area covered by the respective units. If it is considered that there are currently 26 NUTS 2 regions in Turkey, a NUTS 2 region roughly occupies an average area of 30,000 km^2. Thus, in the second phase, AT is used as a parameter in the delineation of PRs. In addition to this, some statistics regarding the density of population in the resulting PRs are also calculated in order to reveal some intuition about the distribution of population.

In the second phase, CT is set to 0 in order to let FRs combine with each other to form larger FRs that can serve as PRs. For the prevention of formation of MPRs, GT is mostly set to 1. From Table 4 showing the results of the second phase, it is observed that if GT = 2, LSC decreases compared with those obtained from RS for the same set of ATs when GT = 1. For the same set of ATs, the average area occupied by a PR also decreases if GT = 2, except for AT = 45,000 km^2 for which it remains the same. As expected, maximum density of a PR regularly decreases parallel to the increase in AT, which signals a balanced distribution of population in the country.

Table 4 reveals that, compared with the first phase, the results of this second phase are more successful in terms of the increase in LSC. If GT = 1, the average LSC for a PR

RS	NI	Limits		FRs	If GT = 1	MultiPol		Statistics for area of FRs (km²)				Statistics for LSC (%)			
		GT	CT		MCL	MFR	Ps	max	min	med	mean	max	min	med	mean
1	1	1	2	188	0.11	–	–	28,109.5	197.5	2,876.5	4,215.9	93.75	41.64	76.53	74.05
2	1	1	3	188	2.13	–	–	28,109.5	27.6	2,814.1	4,215.9	93.75	41.64	76.04	73.66
3	1	1	4	185	2.13	–	–	28,109.5	62.0	2,814.1	4,284.2	93.71	39.24	76.04	73.82
4	1/2	2	2	157	0.24	4	9	25,098.4	27.4	3,105.5	5,048.3	93.79	24.36	78.11	76.03
5	1/2	2	3	156	0.22	4	9	25,098.4	27.4	3,255.0	5,080.7	93.79	24.36	78.55	76.17
6	1/2	2	4	158	0.22	3	6	25,098.4	62.0	3,144.4	5,016.4	93.78	41.64	78.11	76.08
7	1/2	3	2	147/146	0.51	10	20	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.99
8	1/2	3	3	147/146	0.51	9	18	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.99
9	1/2	3	4	148/147	0.51	7	14	27,117.6	275.4	3,386.4	5,391.7	93.78	41.64	79.18	77.19
10	1/2	4	2	145/146	1.45	12	24	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.99
11	1/2	4	3	145/146	1.75/1.45	12	24	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.99
12	**1/2**	**4**	**4**	**146/147**	**1.75/1.45**	**10**	**20**	**27,117.6**	**275.4**	**3,386.4**	**5,391.7**	**93.78**	**41.64**	**79.18**	**77.19**
13	1/2	5	2	146	1.37	16	33	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.97
14	1/2	5	3	146	1.75/1.37	16	33	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.96
15	1/2	5	4	147	1.75/1.37	14	29	27,117.6	275.4	3,386.4	5,391.7	93.78	41.64	79.18	77.17
16	1/2	6	2	146	1.37	16	33	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.97
17	1/2	6	3	146	1.75/1.37	16	33	27,117.6	27.4	3,387.0	5,428.7	93.79	24.36	79.50	76.96
18	1/2	6	4	147	1.75/1.37	14	29	27,117.6	275.4	3,386.4	5,391.7	93.78	41.64	79.18	77.17

Tab. 1: The results of running the script for different limitations set for FRs. Source: author's computations
Notes: RS – 'Run of Script'; NI – 'Number of Iteration'; GT – 'Threshold for maximum GGD'; CT – 'Threshold for minimum commuting level'; FRs – 'Number of FRs formed'; MCL – 'Minimum Commuting Level' used in the formation of FRs; MultiPol statistics for MPR, MFR – number of MPRs; Ps – number of components in MPRs concerned; LSC – 'Level of Self-Containment'. For NI, "1/2" indicates that FRs are delimited within the limits of two iterations, and for the respective of RS if cell values involve a division sign "/", numerator part shows the values for the 1st iteration and denominator part shows the values for the 2nd iteration. If there is no "/" sign in the cell, the value is the same for both iterations. Since no statistics regarding area and self-containment of FRs are calculated for MPRs, for RS involving 2 iterations MultiPol shows statistics for the 1st iteration that actually result in MPRs, and statistics regarding the area and self-containment of FRs belong to the 2nd iteration.

Summary statistics for LSC, CL (%) and area (km²) covered by a BSU or FR	BSU	FR	LSC (%)	BSU		FR	
				frequency	%	frequency	%
Average area	834.7	5,391.7	0–10	28	2.98	0	0
Median area	649.7	3,386.4	10–20	91	9.69	0	0
Maximum area	5,349.9	27,117.6	20–30	149	15.87	0	0
Minimum area	5.7	275.4	30–40	189	20.13	0	0
Maximum LSC	89.91	93.78	40–50	182	19.38	3	2.04
Minimum LSC	2.90	41.64	50–60	158	16.83	10	6.80
Median LSC	51.49	79.18	60–70	108	11.50	17	11.56
Average LSC	51.44	77.19	70–80	32	3.41	46	31.29
Average CL from one BSU or FR to other	0.28	0.82	80–90	2	0.21	65	44.22
MCL taken into account (%)	–	1.45	90–100	0	0	6	4.08
Total number of BSUs or FRs	939	147	Total	939	100.0	147	100.00

Tab. 2: Comparison of statistics for FRs with those of BSUs together with the frequency and share of BSUs and FRs according to the intervals defined for the LSC. Source: author's computations
Note: Statistics regarding FRs given in this table are calculated for FRs obtained after the 12th RS

BSUs in a FR	FFR	BSUs in a FR	FFR	BSUs in a FR	FFR	BSUs in a FR	FFR	BSUs in a FR	FFR
2	30	6	11	10	6	14	3	19	3
3	22	7	7	11	7	15	3	20	1
4	17	8	8	12	3	16	1	21	1
5	12	9	8	13	2	17	2	Grand Total	147

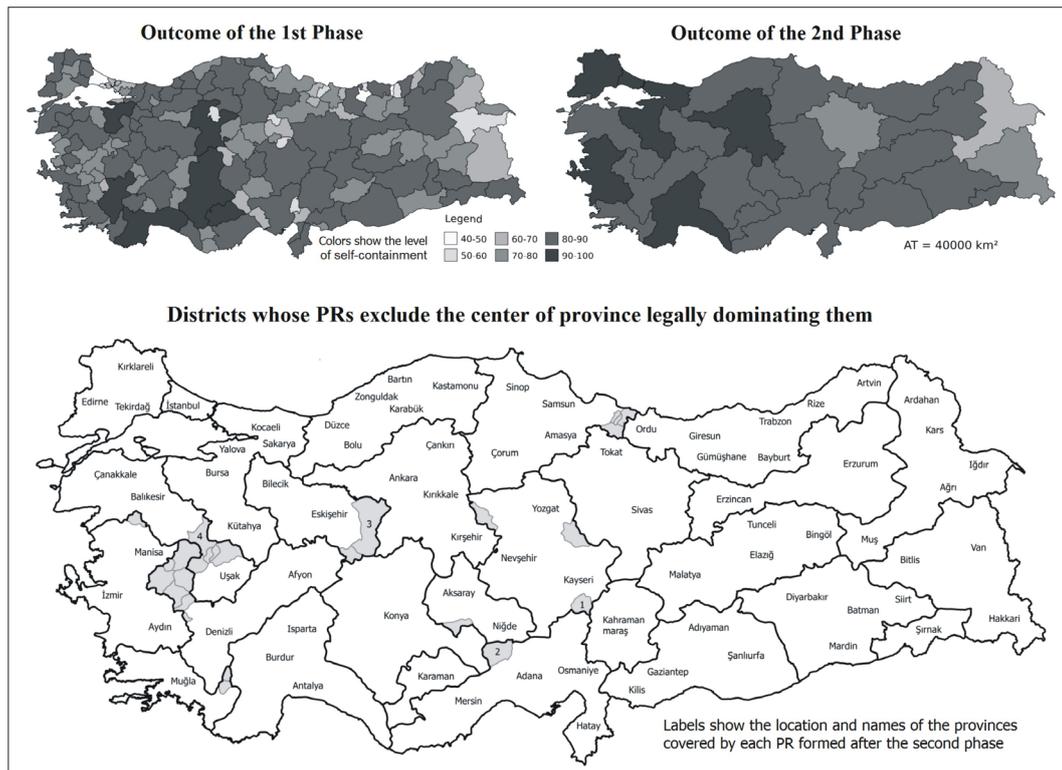*Tab. 3: Number of BSUs in a FR according to the frequency of FR (FFR) concerned. Source: author's computations*

Fig. 3: The results of the 1st and 2nd phases of functional regionalisation of BSUs
Source: author's elaboration

RS	NI	Limits		PRs	MCL	MultiPol		Statistics for area of PRs (km²)				Statistics for LSC (%)				Population per km²			
		GT	AT			MFR	Ps	max	min	med	mean	max	min	med	mean	max	min	med	mean
1	1	1	15,000	62	0.04	–	–	27,117.6	4,635.0	12,803.9	12,783.6	96.75	59.64	83.93	82.94	1,157	10	60	94
2	1	1	20,000	49	0.14	–	–	27,117.6	3,737.8	17,158.9	16,175.2	96.76	65.00	84.23	83.42	906	19	60	87
3	1	1	25,000	41	0.11	–	–	27,117.6	3,829.2	20,357.1	19,331.4	96.78	65.00	83.19	83.86	781	19	59	87
4	1	1	30,000	35	0.06	–	–	29,958.4	7,365.6	23,874.7	22,645.3	96.74	65.00	84.03	84.71	626	19	61	86
5	1/2	2	30,000	31/36	0.06/0.11	6	12	29,958.4	7,365.6	23,721.7	22,016.3	96.74	65.00	84.80	84.66	626	19	58	84
6	1	1	35,000	30	0.26	–	–	35,003.5	5,410.2	29,396.1	26,419.5	96.89	67.09	85.54	85.55	562	23	58	90
7	1/2	2	35,000	28/34	0.07/0.11	6	12	34,566.4	5,410.2	25,497.7	23,311.3	96.71	65.26	86.07	84.98	596	23	56	88
8	**1**	**1**	**40,000**	**26**	**0.17**	**–**	**–**	**39,976.6**	**7,365.6**	**35,683.3**	**30,484.1**	**96.89**	**68.91**	**87.51**	**86.16**	**562**	**26**	**63**	**90**
9	1/2	2	40,000	25/27	0.04	3	6	39,976.6	3,827.0	35,227.3	29,355.0	96.66	65.26	84.37	85.07	488	19	64	85
10	1	1	45,000	24	0.10	–	–	44,517.8	3,827.0	38,073.3	33,024.4	96.81	68.31	86.89	85.79	453	19	52	82
11	1/2	2	45,000	21/24	0.10/0.10	4	9	44,427.2	4,740.5	38,073.3	33,024.4	96.81	65.26	86.73	85.53	453	25	54	84
12	1	1	50,000	21	0.06	–	–	49,948.0	7,644.7	38,579.8	37,742.2	97.19	67.09	84.37	85.64	431	22	61	85
13	1	1	55,000	19	0.44	–	–	54,790.6	7,644.7	47,241.6	41,715.0	97.25	72.21	86.55	86.63	388	22	61	85
14	1	1	60,000	18	0.59	–	–	59,136.4	9,195.3	47,241.6	44,032.5	97.25	72.21	87.71	86.98	388	26	64	87
NUTS 2 Regions				26		–	–	59,531.1	5,212.1	30,722.8	30,484.1	94.88	68.03	86.57	86.10	2,543	26	73	186

Tab. 4: The results of running the script for different limitations set for PR formations compared with the existing NUTS 2 regions. Source: author's computations

increases continuously parallel to the increase in the area covered by a PR, up to a certain level where AT = 40,000 km² (8th RS revealing 26 PRs). The median LSC also increases, though slightly fluctuating, up to the same AT level. The minimum density observed for a PR also increases to the same AT level above which it decreases, albeit it starts to increase slightly again after this decrease. Nevertheless, commuting level between the last couple of FRs (MCL) (0.17%) decreases compared with the functional regionalisation exercise for which AT = 35,000 km². Yet, except for this, MCL for 40,000 km² is higher than both previous MCLs and also MCLs for two subsequent ATs.

A comparison of some statistics from the second phase of functional regionalisation of FRs for the 8th RS with those of NUTS2 regions (last line in Tab. 4) in Turkey reveals that the resulting PRs are more successful in terms of maintaining relatively higher LSC and a balanced distribution of not only physical space but also of population among PRs. The maximum density observed for a NUTS2 region is 2,543 people per km². This value decreases to 562 for PRs delimited after the 8th RS. All PRs formed after the 8th RS for the second phase of functional regionalisation can also be seen in Figure 3, jointly with a detailed map showing the location and names of the provinces covered by each PR,

together with the districts (grey colour) that are included in a PR excluding the centre of the province to which they are officially connected. Overall, although PRs identified in the second phase by setting AT to 40,000 km² generally overlap with the administrative regions of Turkey, there are also differences between PRs and normative regions.

For example, Tufanbeyli, officially a district of Adana, actually has greater interactions with Kayseri in terms of commuting flows and, beginning from the first phase of the functional regionalisation, it is always connected to the region dominated by Kayseri (district No. 1 in the map at the bottom of Fig. 3). In a similar fashion, Ulukışla, formally within Niğde, is connected by FRGIS to the region dominated by the Adana-Mersin metroplex (district No. 2). Likewise, Polatlı is connected by FRGIS to the region of Eskişehir, rather than Ankara that legally covers it (district No. 3).

Although these kinds of PRs are actually an expected outcome of the study, some PRs are too small (e.g. Şırnak, Karaman and Muğla, each of which covers less than 14,000 km²) compared with others. The tendency of small regions to remain isolated from main PRs does not readily mean that they can be designated as separate PRs. Indeed, isolation of these regions can be prevented by increasing AT. Another relatively unexpected outcome of this application of FRGIS is the formation of PRs having low levels of compactness. For example, the extension of FR covering mainly Çanakkale and Balıkesir includes Uşak and western districts of Kütahya in the form of a peninsula connected to them via Simav (district No. 4), whose removal leads to a MPR. In comparison, compact solutions in terms of the morphology of FRs are more likely to reflect a community of interest without creating an impression of the political gerrymandering of boundaries (Johnston and Rossiter, 1981).

Commuter statistics for the 1st phase of functional regionalisation

RS	FRs	People Residing in FR				People Working in FR			
		max	min	med	mean	max	min	med	mean
1	188	849,061	1,072	20,137	55,484	944,222	1,091	18,263	55,484
2	188	849,061	1,072	20,803	55,484	944,222	1,091	18,606	55,484
3	185	849,061	1,072	22,127	56,384	944,222	1,091	20,173	56,384
4	157	1,098,906	1,417	25,419	66,439	1,474,614	1,091	23,464	66,439
5	156	1,098,906	1,417	25,705	66,865	1,474,614	1,091	23,454	66,865
6	158	857,520	1,417	26,248	66,019	949,700	1,091	24,108	66,019
7	146	1,098,906	1,417	27,849	71,445	1,474,614	1,091	24,765	71,445
9	147	855,841	1,417	28,090	70,959	948,548	1,091	24,891	70,959
10	146	1,098,906	1,417	27,849	71,445	1,474,614	1,091	24,765	71,445
12	147	855,841	1,417	28,090	70,959	948,548	1,091	24,891	70,959
13	146	1,098,906	1,417	27,849	71,445	1,474,614	1,091	24,765	71,445
15	147	855,841	1,417	28,090	70,959	948,548	1,091	24,891	70,959
16	146	1,098,906	1,417	27,849	71,445	1,474,614	1,091	24,765	71,445
18	147	855,841	1,417	28,090	70,959	948,548	1,091	24,891	70,959

Commuter statistics for the 2nd phase of functional regionalisation

RS	PRs	People Residing in PR				People Working in PR			
		max	min	med	mean	max	min	med	mean
1	62	3,087,402	8,772	70,434	168,242	3,268,994	9,805	67,208	168,242
2	49	3,421,809	10,424	85,894	212,877	3,605,598	9,072	84,223	212,877
3	41	3,472,098	10,424	100,882	254,414	3,653,727	9,072	89,170	254,414
4	35	3,518,588	21,458	117,824	298,028	3,699,765	22,104	127,538	298,028
5	36	3,518,588	21,458	119,826	289,750	3,699,765	22,104	116,912	289,750
6	30	3,640,801	28,848	190,985	347,700	3,817,684	28,789	184,658	347,700
7	34	3,527,311	16,176	126,381	306,794	3,706,701	11,891	126,478	306,794
8	26	3,640,801	21,458	239,616	401,192	3,817,684	25,220	225,171	401,192
9	27	3,584,176	8,492	221,900	386,333	3,760,933	7,761	211,318	386,333
10	24	3,706,389	8,492	237,635	434,625	3,878,852	7,761	227,381	434,625
11	24	3,706,389	16,176	237,635	434,625	3,878,852	11,891	228,147	434,625
12	21	3,860,274	22,945	256,304	496,714	4,041,653	23,025	244,124	496,714
13	19	3,911,265	22,945	256,304	549,000	4,091,154	23,025	244,124	549,000
14	18	3,911,265	65,588	287,263	579,500	4,091,154	61,168	269,909	579,500

Tab. 5: Statistics for commuters for the 1st and 2nd phases of functional regionalisation
Source: author's computations

Again, it is considered that if AT is re-adjusted, these kinds of PRs can be prevented. Indeed, after several tests, it is observed that if AT = 41,000 km^2, Uşak and western districts of Kütahya are excluded from the PR covering mainly Çanakkale and Balıkesir, and connected to the PR covering mainly Denizli and Afyon. Overall, what is evident from this exercise is that although FRs can be processed to form PRs, some idiographic judgments based on the intuition of experts are inevitable for the designation of the final form of PRs. In order to produce these judgments, on the one hand, the outcomes of the script for various adjustments made for AT can be used, and on the other hand, social-historical factors shaping the institutional structure can be taken into account.

Comparison of the results of the proposed method of delimitation with the newest TTWA algorithm provides us with some further insights about the relevance of the restriction for the area that can be occupied by a PR. For this purpose, two additional tables are created, together with another figure (see Fig. 4) showing NUTS 2 regions and the LMAs delimited by using the R package LabourMarketAreas, designed to implement the TTWA algorithm (for the use of this package, see Franconi et al., 2017; Ichim et al., 2018; Franconi and Ichim, 2018). The first table (Tab. 5) shows commuter statistics for FRs and PRs delimited by using FRGIS. The second one (Tab. 6) shows the characteristics of LMAs delimited by using the TTWA algorithm for different values of minSZ, tarSZ, minSC and tarSC. Casado-Díaz et al. (2017) show that in the application of the TTWA algorithm in a country characterised by an unbalanced distribution of population over space, population size (employed residents) and self-containment can initially be adjusted to 5,000 and 85%, respectively.

Accordingly, in the implementation of TTWA algorithm in Turkey, minSC is initially set to 5,000 and tarSC is increased up to 100,000. Nevertheless, the minimum LSC calculated for the existing NUTS2 region in Turkey is 68.03. Thus, in this study, in the implementation of TTWA algorithm, minSC is at first set to 70% and increased up to 78% for various runs of TTWA algorithm. Yet, tarSC is at first set

to 85% and increased up to 96%. Apart from the results given in Table 6, the TTWA algorithm has also been run for a set of other configurations of minSZ, tarSZ, minSC and tarSC (see Tab. 7).

Table 5 demonstrates that compared with the second phase of functional regionalisation, commuter statistics for various RS in the first phase do not change much after the decrease of FRs to 146–147. Nevertheless, in the second phase, the numbers of people residing and working in a PR increase up to the 8th RS after which there is a slight decrease for all the measured statistics. As discussed above, compared with other RS, in the 8th RS, higher LSC is maintained with a more balanced distribution of physical space. This is particularly important when the results of the TTWA algorithm for different limitations set for LMA formations are analysed. In this regard, Table 6 reveals that, although compared with PRs identified by FRGIS relatively better minimum and average LSC values and commuter statistics are obtained from the TTWA algorithm, there are huge differences between the physical areas occupied by different LMAs. The result regarding the unbalanced distribution of physical space among LMAs delimited by the TTWA algorithm does not change for a set of other configurations of minSZ, tarSZ, minSC and tarSC (see Tab. 7 and Fig. 5).

An unbalanced distribution of physical space among the 26 LMAs delimited by the TTWA algorithm can also be seen in Figure 4. The maximum density observed for a LMA is 1,097 people per km^2 and it is high compared with the one observed for a PR. At the end of second section, it was argued that in countries where the population is not distributed uniformly, the use of population as a restriction may result in undesired outcomes. It is even claimed that if a restriction is imposed on the population of a FR, the provincial boundaries of İstanbul may not be preserved. Indeed, the TTWA algorithm based on population limitation creates two regions covering the parts of İstanbul province (see Fig. 4):

1. the European part of İstanbul, together with the other provinces in Thrace region except for the Gallipoli peninsula; and
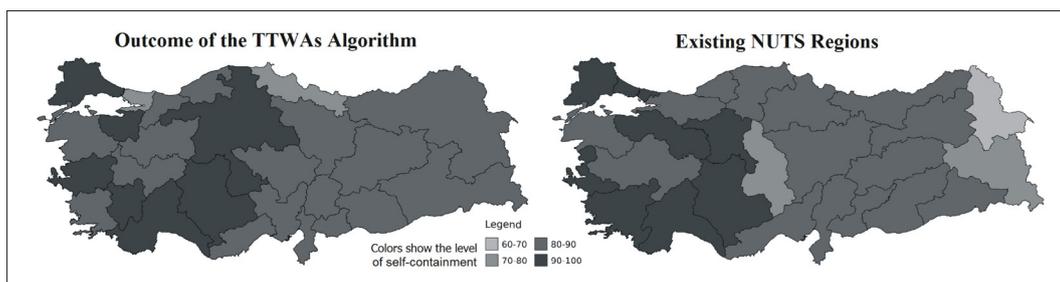

Fig. 4: 26 LMAs delimited by the TTWA algorithm compared with the NUTS 2 regions
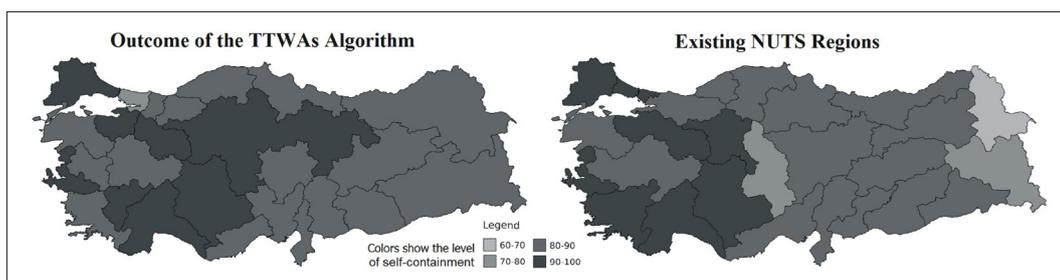Source: author's elaboration

Fig. 5: 26 LMAs delimited by the TTWA algorithm for a set of other configurations of parameters compared with the NUTS 2 regions. Source: author's elaboration

Parameters used in TTWA				Num	MultiPol		Statistics for area of LMA				Statistics for LSC (%)				Statistics for Residents				Statistics for Workers			
minSC	tarSC	minSZ	tarSZ	LMAs	MFR	Ps	max	min	med	mean	max	min	med	mean	max	min	med	mean	max	min	med	mean
70	85	5,000	50,000	82	14	41	52,762	1,085	6,720	9,666	92.97	71.37	83.80	83.58	1,544,999	12,549	55,399	127,207	1,774,283	11,640	54,184	127,207
71	86	10,000	55,000	75	13	36	52,762	1,085	7,645	10,568	92.97	71.37	84.14	83.91	1,544,999	20,895	63,243	139,080	1,774,283	20,549	58,500	139,080
72	88	20,000	62,500	66	8	24	60,993	1,085	9,742	12,009	96.31	72.13	84.56	84.55	2,764,003	24,801	73,145	158,045	2,959,589	24,678	67,200	158,045
73	90	35,000	70,000	48	10	31	87,605	2,109	12,741	16,512	96.03	74.15	85.40	85.49	2,738,357	46,156	122,160	217,312	2,927,927	46,850	111,376	217,312
74	92	50,000	80,000	39	9	29	84,199	2,861	14,814	20,323	96.37	74.15	86.03	86.46	2,771,716	56,216	162,211	267,461	2,965,861	53,908	148,852	267,461
76	94	70,000	90,000	34	8	30	98,825	3,343	15,255	23,311	96.37	76.78	86.45	86.77	2,771,716	80,844	177,824	306,794	2,965,861	74,525	164,968	306,794
78	96	90,000	100,000	26	13	44	112,765	5,871	21,984	30,484	94.41	79.35	87.22	86.79	1,925,758	117,423	228,002	401,192	2,195,715	105,002	218,091	401,192
Existing NUTS 2 in Turkey				26	–	–	59,531	5,212	30,723	30,484	94.88	68.03	86.57	86.10	2,639,699	64,926	275,332	401,192	2,805,537	54,661	266,162	401,192
PRs identified by FRGIS				26	–	–	39,977	7,366	35,683	30,484	96.89	68.91	87.51	86.16	3,640,801	21,458	239,616	401,192	3,817,684	25,220	225,171	401,192

Tab. 6: The results of running the TTWA algorithm for different limitations set for LMA formations compared with the existing NUTS 2 regions and PRs identified in the study. Source: author's computations

Parameters used in TTWA				Num	MultiPol		Statistics for area of LMA				Statistics for LSC (%)				Statistics for Residents				Statistics for Workers			
minSC	tarSC	minSZ	tarSZ	LMAs	MFR	Ps	max	min	med	mean	max	min	med	mean	max	min	med	mean	max	min	med	mean
70	85	5,000	50,000	82	14	41	52,762	1,085	6,720	9,666	92.97	71.37	83.80	83.58	1,544,999	12,549	55,399	127,207	1,774,283	11,640	54,184	127,207
70	85	10,000	60,000	76	14	38	52,762	1,085	7,489	10,429	92.97	71.37	84.10	83.91	1,544,999	21,431	62,085	137,250	1,774,283	21,892	58,248	137,250
72	86	15,000	72,500	68	8	24	69,937	1,085	8,793	11,656	96.03	72.82	84.61	84.55	2,738,357	23,909	67,802	153,397	2,927,927	21,892	65,203	153,397
72	86	25,000	85,000	58	7	19	76,433	1,085	10,287	13,665	96.03	72.82	85.09	85.06	2,738,357	32,173	86,511	179,845	2,927,927	31,240	80,357	179,845
74	88	35,000	100,000	43	12	32	97,748	2,681	13,252	18,432	96.03	77.34	86.03	86.19	2,738,357	46,156	131,654	242,581	2,927,927	46,850	121,836	242,581
74	88	45,000	115,000	39	12	31	101,395	2,970	14,003	20,323	96.03	76.83	86.20	86.70	2,738,357	50,991	144,501	267,461	2,927,927	49,501	140,949	267,461
76	90	55,000	130,000	31	10	25	113,658	4,865	16,987	25,567	93.74	77.58	87.26	86.77	1,925,758	65,411	192,728	336,484	2,195,715	65,806	178,367	336,484
76	90	65,000	145,000	28	9	29	115,885	4,865	21,962	28,307	93.74	78.06	87.40	87.02	1,925,758	100,123	206,209	372,536	2,195,715	100,456	196,359	372,536
78	90	75,000	150,000	26	10	34	119,058	4,865	22,025	30,484	93.74	79.24	87.74	87.34	1,925,758	100,123	224,875	401,192	2,195,715	100,456	223,745	401,192
Existing NUTS 2 in Turkey				26	–	–	59,531	5,212	30,723	30,484	94.88	68.03	86.57	86.10	2,639,699	64,926	275,332	401,192	2,805,537	54,661	266,162	401,192
PRs identified by FRGIS				26	–	–	39,977	7,366	35,683	30,484	96.89	68.91	87.51	86.16	3,640,801	21,458	239,616	401,192	3,817,684	25,220	225,171	401,192

Tab. 7: The results of running the TTWA algorithm for a set of other configurations of parameters used for LMA formations compared with the existing NUTS 2 regions and PRs identified in the study. Source: author's computations

2. the Asian part of İstanbul together with the provinces of Kocaeli and Yalova, previously part of İstanbul province.

A similar kind of situation can also be observed for Çukurova Region, covering mainly the Adana-Mersin metroplex but extending to Hatay. Although FRGIS delimits this region as a single PR, the TTWA algorithm creates three LMAs for this well known region in Turkey:

1. Adana,

2. Mersin, and

3. Hatay and Osmaniye, previously part of Adana province.

Overall, in the western part of the country, compared with the NUTS 2 regions, some LMAs revealed by the TTWA algorithm are smaller, sometimes forming two LMAs approximately for a single NUTS 2 region. As argued in the second section, in terms of planning for such regions, it would be wiser to include all BSUs involved in the regions concerned. Nevertheless, this does not mean that overly large regions can be considered as PRs. In this respect, comparison of the results of the TTWA algorithm with the NUTS 2 regions reveals another positive aspect of the introduction of a limitation for the area that can be occupied by a PR in a country where the population is not distributed uniformly.

Indeed, Figure 4 shows that some LMAs revealed by the TTWA algorithm for the eastern part of the country and central Anatolia are overly large compared with the NUTS 2 regions. Some of these LMAs cover huge parts of two or three NUTS 2 regions. This contrast observed in the formation of the LMAs for the eastern and western parts of the country seems to stem from the unbalanced distribution of the population in the country.

5. Concluding remarks

Compatibility of the administrative regions with FRs and subsequently with PRs is desired because the lack of compatibility usually leads to both planning problems and tensions between different administrative regions. In this context, this study used FRs based on the commuting patterns of people as PRs. In this framework, an algorithm (FRGIS) designed to run as a two-phase model is developed and used in order to delineate FRs that can serve as PRs in Turkey.

In terms of the employment of a nomothetic method of science, the results of the analysis show that FRGIS is generally successful in terms of the formation of spatially-balanced regions having higher LSC compared with those of existing NUTS 2 regions. Although the minimum and average LSC values of the resulting PRs are not as high as the results from the TTWA algorithm, compared with LMAs, the formation of regions according to the proposed method of delimitation is spatially more balanced. This is actually an expected result of the study. The distribution of population in terms of density is also more balanced in the resulting PRs. Thus, it can be argued that the objectives of the research are mostly fulfilled.

It is also concluded, however, that the formation of very small PRs or PRs having low levels of compactness, may not be prevented. This can be considered a disadvantage of the method of delimitation used in this study. It seems that alternative methods for the delimitation of FRs, however, also suffer from similar problems. The introduction of some other constraints, such as the minimum area occupied by a PR and the calculation of the morphometric characteristics

of the PR, may prevent the undesired outcomes mentioned above. Nevertheless, some idiographic judgments seem to be inevitable if FRs are to be employed as PRs. For example, in this regard, it is suggested that the outcomes of FRGIS for various adjustments made for different parameters used for delimitation of regions, as well as social-historical factors shaping the institutional and spatial structure, can be taken into account in the designation of FRs as PRs. Another disadvantage of FRGIS is the lack of a direct control on the minimum LSC. This is particularly noticeable for the outcomes in the first phase of the functional regionalisation process. Since FRGIS is actually designed to exhibit PRs, the outcomes of the first phase should be considered as an intermediate stage for designation of final FRs that can then be employed as PRs.

There are various advantages of the proposed method of delimitation of FRs as PRs in this study. First of all, it does not require any information nor have any implication for the designation of some BSUs as central places. Another advantage of the approach used in this study is its simplicity in terms of the steps of FRGIS and the preparation of the database used in the analysis. The first and second phases of FRGIS are actually based on the same stages except for the parameters used to constraint the formation of FRs or PRs. Last, but not least in importance, the employment of GIS that facilitates the spatial analysis required to delineate FRs, is another advantage of this study.

In this respect, FRGIS draws on a series of functions and libraries available in FOSS for GIS for the calculation of GGD between BSUs, the prevention of MPRs, and the calculation of the total area occupied by a FR. With the introduction of GGD as an indirect and loose control of the condition for contiguity that can easily be checked in the GIS environment, and for the automatic prevention of MPRs, no interruption is required by the user to control MPRs in the delimitation of FRs and PRs. In alternative approaches, the prevention of MPRs is a lengthy process requiring intervention of the user for each run of the script produced to reveal FRs.

Compared with other functional regionalisation studies conducted in Turkey, this is the first research project to use an algorithm (FRGIS) and script specifically developed for this purpose and drawing on the commuting flows occurring in the country. As the plugin produced in this study is based on FOSS, it may trigger further developments in this field for other countries. The database employed in this study also allows for a more accurate delimitation of the boundaries of the metropolitan regions or sub-regions. Within this context, future studies can explore the geographical extent of the metropolitan regions and improve FRGIS by adding considerations for the automatic determination of the parameters leading to the formation of optimal FRs. Future studies may also take the morphometric characteristics of the FRs into account in order to produce the formation of more compact FRs.

Appendix 1: List of abbreviations

- AC: Amount of Commuting
- AT: Threshold for maximum area that can be occupied by a PR
- BAGKUR: Social Insurance Institution for the Craftsmen & Artisans & Other Self Employers
- BSU: Basic Spatial Units
- CAA: Commuting Aggregation Approach
- CL: Commuting Level
- CT: Threshold for minimum commuting level between BSUs
- CURDS: Centre for Urban and Regional Development Studies
- dbi: database iteration number – number of iteration to read database
- DB: Destination BSU
- Emekli Sandığı: General Directorate of Retirement Fund
- EOF: End of file
- ERA: European Regionalisation Algorithm
- FOSS: Free and Open Source Software
- FR: Functional Region
- FRGIS: Algorithm for Functional Regionalisation in GIS with Auto-Prevention of MPR on the Base of GGD and Areal Restriction
- GEA: Grouping Evolutionary Algorithm
- GGD: Graph Theoretical Geodesic Distance
- GIS: Geographic Information Systems
- GT: Threshold for maximum GGD between BSUs
- id: identity number
- JTS: Java Topology Suite
- LSC: Level of Self-Containment
- LMA: Labor Market Area
- MCL: Minimum CL
- minSC: Level of self-containment that is acceptable for cluster of large sizes
- minSZ: Minimum number of employees for a cluster to be considered an LMA
- MPR: Multi-Polygon Region
- NUTS: Nomenclature des Unités Territoriales Statistiques
- OB: Origin BSU
- OSM: Open-Street-Mapping
- PR: Planning Region
- RS: Run of the Script
- SD: Shortest Distance
- SSI: Social Security Institution
- SSK: Social Insurance Institution - former Workers' Insurance Institution
- tarSC: Target self containment of an area in order for a small cluster of communities to be considered an LMA
- tarSZ: Target value for the size of the cluster in terms of occupied persons
- TG: GGD between OB and DB
- TTWA: Travel-to-work-areas

References:

ANDERSEN, A. K. (2002): Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark? Regional Studies, 36(8): 833–844.

AYDEMİR, Ş. (1978): Doğu Karadeniz Bölgesi Trabzon Alt Bölgesi (TBAB) Kentsel Etki Alanlarının Saptanması için Yöntem: Etkileşim Esası. Trabzon, Karadeniz Matbaacılık ve Gazetecilik A.Ş.

BEYHAN, B. (2012): Plugins and Scripts Developed to Conduct Space Syntax Analysis in FOSS4GIS: OpenJUMP, gvSIG, OrbisGIS, Quantum GIS, OpenEV, Thuban, MapWindow GIS, SAGA, and R Project. Mersin, Mersin University Publications.

BROWN, L. A., HOLMES, J. (1971): The delimitation of functional regions, nodal regions, and hierarchies by functional distance approaches. Journal of Regional Science, 11(1): 57–72.

BROWN, L. A., HORTON, F. E. (1970): Functional Distance: An Operational Approach. Geographical Analysis, 2: 76–83.

BROWN, P. J. B., PITFIELD, D. E. (1990): An Intramax derivation of commodity market areas from freight flow data. Transportation Planning and Technology, 15(1): 59–81.

CASADO-DÍAZ, J. M., COOMBES, M. (2011): The delineation of 21st century local labour markets areas: A critical review and a research agenda. Boletín de la Asociación de Geógrafos Españoles, 57: 7–32.

CASADO-DÍAZ, J. M., MARTÍNEZ-BERNABÉU, L., ROWE, F. (2017): An evolutionary approach to the delimitation of labour market areas: an empirical application for Chile. Spatial Economic Analysis, 12(4): 379–403.

CATTAN, N. (2002): Part I: A Synthesis of Definitions and Usage. In: Redefining territories: The functional regions (pp. 8–23). Paris, OECD Publications Service.

CHI, G., THILL, J-C., TONG, D., SHI, L., LIU, Y. (2016): Uncovering regional characteristics from mobile phone data: A network science approach. Papers in Regional Science, 95(3): 613–631.

CHRISTALLER, W. (1933): Die zentralen Orte in Suddeutschland. Jena, Gustav Fischer.

CLARK, D. (1973): The Formal and Functional Structure of Wales. Annals of the Association of American Geographers, 63(1): 71–84.

CLAYTON, C. (1980): Interdependence in Urban Systems and Its Application to Political Reorganization. Geografiska Annaler, Series B, Human Geography, 62(1): 11–20.

CLIFF, A. D., HAGGETT, P. (1970): On the efficiency of alternative aggregations in region-building problems. Environment and Planning A, 2(3): 285–294.

COOMBES, M. G. (2000): Defining locality boundaries with synthetic data. Environment and Planning A, 32(8): 1499–1518.

COOMBES, M. G., BOND, S. (2008): Travel-to-work areas: The 2007 review. London, Office for National Statistics.

COOMBES, M. G., GREEN, A. E., OPENSHAW, S. (1986): An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the1984 Travel-to-

Work Areas Revision in Britain. The Journal of the Operational Research Society, 37(10): 943–953.

CÖRVERS, F., HENSEN, M., BONGAERTS, D. (2009): Delimitation and Coherence of Functional and Administrative Regions. Regional Studies, 43(1): 19–31.

DPT (1982): Türkiye'de Yerleşme Merkezlerinin Kademelenmesi: Ülke Yerleşme Merkezleri Sistemi, Cilt I. Ankara, Devlet Planlama Teşkilatı.

DROBNE, S. (2017): Functional regions and areas: Literature review according to application fields. Geodetski Vestnik, 61(1): 35–57.

DROBNE, S., KONJAR, M., LISEC, A., PICHLER MILANOVIĆ, N., ZAVODNIK LAMOVŠEK, A. (2010): Functional Regions Defined by Urban centres of (Inter)National Importance: The case of Slovenia. In: Schrenk, M., Popovich, V. V., Zeile, P. [eds.]: Real Corp 2010: Proceedings of 15th International Conference on Urban Planning, Regional Development and Information Society, Wien, Austria, 18–20.05.2010 (pp. 295–305). [online]. [cit.10.06.2017]. Available at: http://www.geomultimedia.org/archive/CORP2010_153.pdf

DROBNE, S., LAKNER, M. (2016): Intramax and other objective functions: The case of Slovenia. Moravian Geographical Reports, 24(2): 12–25.

ECEMIŞ-KILIÇ, S. (2009): Proposals for Regional Administrative Structure and Planning in Turkey. European Planning Studies, 17(9): 1283–1301.

ERLEBACH, M., TOMÁŠ, M., TONEV, P. (2016): A functional interaction approach to the definition of meso regions: The case of the Czech Republic. Moravian Geographical Reports, 24(2): 37–46.

FARMER, C. J. Q., FOTHERINGHAM, A. S. (2011): Network-based functional regions. Environment and Planning A, 43(11): 2723–2741.

FELDMAN, O., SIMMONDS, D. C., TROLL, N., TSANG, F. (2005): Creation of a System of Functional Areas for England and Wales and for Scotland. In: Proceedings of the European Transport Conference, Strasburg, France.

FISCHER, M. M., ESSLETZBICHLER, J., GASSLER, H., TRICHTL, G. (1993): Telephone Communication Patterns in Austria: A Comparison of the IPFP-based Graph-Theoretic and the Intramax Approaches. Geographical Analysis, 25(3): 224–233.

FRANCONI, L., D'ALÒ, M., ICHIM, D. (2016): Istat implementation of the algorithm to develop labour market areas. Roma, ISTAT. Available at: https://www.istat.it/en/files/2016/03/Description-of-the-LabourMarketAreas-algorithm.pdf

FRANCONI, L., ICHIM, D. (2018): Census data for the development of functional geographies: The case of Labour Market Areas – Note by the Italian National Institute of Statistics (Istat), Rome, Italy. Economic Commission for Europe, Conference of European Statisticians, Group of Experts on Population and Housing Censuses, Twentieth Meeting, Geneva, 26–28 September 2018, Available at: https://undocs.org/ECE/CES/GE.41/2018/8

FRANCONI, L., ICHIM, D., D'ALÒ, M. (2017): Labour Market Areas for territorial policies: Tools for a European approach. Statistical Journal of the IAOS, 33(3): 585–591.

FRANCONI, L., ICHIM, D., D'ALÒ, M., CRUCIANI, S. (2017): Guidelines for Labour Market Area delineation process: From definition to dissemination. Available at: https://ec.europa.eu/eurostat/cros/system/files/guidelines_for_lmas_production08082017_rev300817.pdf

FRIEDMANN, J. R. P. (1956): The Concept of a Planning Region. Land Economics, 32(1): 1–13.

GERAY, C. (1997): Bölgesel Gelişme için Planlama ve Örgütlenme. SBF Dergisi, 52(1–4): 299–313.

GODDARD, J. B. (1970): Functional Regions within the City Centre: A Study by Factor Analysis of Taxi Flows in Central London. Transactions of the Institute of British Geographers, 49: 161–182.

GRIGG, D. (1965): The Logic of Regional Systems. Annals of the Association of American Geographers, 55(3): 465–491.

HAGGETT, P., CHORLEY, R. J. (1969): Network Analysis in Geography. New York, St. Martin's Press.

HALÁS, M., KLAPKA, P., BEDNÁŘ, M., TONEV, P. (2015): An alternative definition and use for the constraint function for rule-based methods of functional regionalisation. Environment and Planning A, 47(5): 1175–1191.

HEMMASI, M. (1980): The Identification of Functional Regions Based on Lifetime Migration Data: A Case Study of Iran. Economic Geography, 56(3): 223–233.

HINCKS, S., WONG, C. (2010): The Spatial Interaction of Housing and Labour Markets: Commuting Flow Analysis of North West England. Urban Studies, 47(3): 620–649.

HOLMES, J. H., HAGGETT, P. (1977): Graph theory interpretation of flow matrices: a note on maximization procedures for identifying significant links. Geographical Analysis, 9(4): 388–399.

ICHIM, D., FRANCONI, L., D'ALÒ, M. VAN DEN HEUVEL, G. (2018): Package 'LabourMarketAreas'. Available at: https://cran.r-project.org/web/packages/LabourMarket Areas/LabourMarketAreas.pdf

ILLERIS, S., PEDERSEN, P. O. (1968): Central Places and Functional Regions in Denmark: Factor Analysis of Telephone Traffic. Lund Studies in Geography, Series B, No. 311, Department of Geography, The Royal University of Lund, Sweden.

JOHNSTON, R. J., ROSSITER, D. J. (1981): An approach to the delimitation of planning regions. Applied Geography, 1(1): 55–69.

KARLSSON, C., OLSSON, M. (2006): The identification of functional regions: theory, methods, and applications. Annals of Regional Science, 40(1): 1–18.

KEANE, M. J. (1978): A functional distance approach to regionalization. Regional Studies, 12(3): 379–386.

KEELING, D. J. (1994): Regional Development and Transport Policies in Argentina: An Appraisal. The Journal of Developing Areas, 28(4): 487–502.

KLAPKA, P., HALÁS, M. (2016): Conceptualising patterns of spatial flows: Five decades of advances in the definition and use of functional regions. Moravian Geographical Reports, 24(2): 2–11.

KLAPKA, P., HALÁS, M., ERLEBACH, M., TONEV, P., BEDNÁŘ, M. (2014): A Multistage Agglomerative Approach for Defining Functional Regions of the Czech

Republic: The Use of 2001 Commuting Data. Moravian Geographical Reports, 22(4): 2–13.

KONJAR, M., LISEC, A., DROBNE, S. (2010): Methods for delineation of functional regions using data on commuters. Paper presented at the 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, 10–14.05.2010 [online]. [cit.12.06.2017]. Available at: http://agile2010.dsi.uminho.pt/pen/Short Papers_PDF/93_DOC.pdf

LANDRÉ, M., HÅKANSSON, J. (2013): Rule versus Interaction Function: Evaluating Regional Aggregations of Commuting Flows in Sweden. EJTIR, 13(1): 1–19.

MARTÍNEZ-BERNABEU, L., FLÓREZ-REVUELTA, F., CASADO-DÍAZ, J. M. (2012): Grouping genetic operators for the delineation of functional areas based on spatial interaction. Expert Systems with Applications, 39(8): 6754–6766.

MASSER, I., BROWN, P. J. B. (1975): Hierarchical aggregation procedures for interaction data. Environment and Planning A, 7(5): 509–523.

MASSER, I., SCHEURWATER, J. (1978): The specification of multi-level systems for spatial analysis. In: Masser, I., Brown, P. J. B. [eds.]: Spatial representation and spatial interaction (pp. 151–172). Studies in Applied Regional Science, 10, Leiden-Boston, Martinus Nijhoff.

MITCHELL, W., WATTS, M. (2010): Identifying Functional Regions in Australia Using Hierarchical Aggregation Techniques. Geographical Research, 48(1): 24–41.

NADER, G. A. (1980): An Economic Regionalization of Canada: The Validity of Provinces as Regions for the Conduct of Regional Economic Policy. Canadian Journal of Regional Science, 3(2): 117–138.

NORONHA, V. T., GOODCHILD, M. F. (1992): Modelling Interregional Interaction: Implications for Defining Functional Regions. Annals of the Association of American Geographers, 82(1): 86–102.

NYSTUEN, J. D., DACEY, M. F. (1961): A Graph Theory Interpretation of Nodal Regions. Papers, Regional Science Association, 7(1): 29–42.

ÖZTÜRK, A. (2009): Homojen ve Fonksiyonel Bölgelerin Tespiti ve Türkiye için İstatistiki Bölge Birimleri Önerisi, DPT Planlama Uzmanlığı Tezi (Planning Expertise Thesis). Ankara, DPT.

PÁLÓCZI, G., PÉNZES, J., HURBÁNEK, P., HALÁS, M., KLAPKA, P. (2016): Attempts to delineate functional regions in Hungary based on commuting data. Regional Statistics, 6(1): 23–41.

PLATT, R. S. (1928): A Detail of Regional Geography: Ellison Bay Community as an Industrial Organism. Annals of the Association of American Geographers, 18(2): 81–126.

TEKELİ, İ. (1972): Üçüncü Be – Yıllık Plan Üzerine Genel Bir Değerlendirme. Elektrik Mühendisliği, 17(192): 52–80.

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, 9(2): 201–221.

VAN NUFFEL, N. (2007): Determination of the number of significant flows in origin-destination specific analysis: The case of commuting in Flanders. Regional Studies, 41(4): 509–524.

VAN OORT, F., BURGER, M., RASPE, O. (2010): On the Economic Foundation of the Urban Network Paradigm: Spatial Integration, Functional Integration and Economic Complementarities within the Dutch Randstad. Urban Studies, 47(4): 725–748.

WARD, J. H. (1963): Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58(301): 236–244.

WIRTH, L. (1937): Localism, Regionalism, and Centralization. The American Journal of Sociology, 42(4): 493–509.

WIRTH, L. (1942): The Metropolitan Region as a Planning Unit. In: James, H. [ed.]: American Planning and Civic Annual (pp. 174–184). Harrisburg, Mount Pleasant Press, J. Horace McFarland Company.