

## Article

# Delimitation of Functional Regions Using a $p$ -Regions Problem Approach

International Regional Science Review

2015, Vol. 38(3) 235-263

© The Author(s) 2013

Reprints and permission:

[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)

DOI: 10.1177/0160017613484929

[irx.sagepub.com](http://irx.sagepub.com)

Hyun Kim<sup>1</sup>, Yongwan Chun<sup>2</sup>, and Kamyung Kim<sup>3</sup>

## Abstract

Various spatial data analyses have been used for the identification of functional regions. Functional regions are identified by grouping many areal units into fewer clusters to classify the areal units in terms of similar properties, as well as to constrain the spatial contiguity of the areal units in each cluster. This article proposes a spatial optimization model, called the  $p$ -functional regions problem, to solve a regionalization problem by considering geographic flows. The magnitude of geographic flows, such as journey-to-work, is widely considered a good indicator of functional relationships between areas so that regionalization models incorporating various criteria, such as the maximum intraregion flows or the total inflows from other units, may be used to identify the  $p$  regions. We also propose an *analytical target reduction* approach to enhance the model tractability in generating optimal solutions to large problems and to demonstrate the effectiveness of the optimization model using journey-to-work data from Seoul (South Korea) and South Carolina (the United States).

## Keywords

functional regions, spatial optimization,  $p$ -regions problem, geographic flows

<sup>1</sup> Department of Geography, University of Tennessee at Knoxville, Knoxville, TN, USA

<sup>2</sup> School of Economic, Political and Policy Sciences, University of Texas at Dallas, Richardson, TX, USA

<sup>3</sup> Department of Geography Education, Kyungpook National University, Daegu, Korea

## Corresponding Author:

Kamyung Kim, Department of Geography Education, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 702-701, Korea.

Email: [kamyungkim@knu.ac.kr](mailto:kamyungkim@knu.ac.kr)

## Introduction

The concept of functional regions provides a framework for investigating the structure of geographic space. A *functional region* is characterized by the agglomeration of economic activities at a regional center that focuses people's movements from the surrounding vicinities or hinterland. The center in a functional region is often referred to as a central place or central business district (CBD), the size of which, in terms of economic activity, depends on the range of goods and services provided (Berry and Garrison 1958). A functional region is generally described in terms of the structure of spatial interactions, which helps delineate regions naturally by maximizing the difference between within-region and between-region interactions (Noronha and Goodchild 1992). Thus, the pattern of geographic flows is viewed as an effective indicator for understanding how people's economic activities are structured over space (Taaffe, Gauthier, and O'Kelly 1996). For example, daily commuting flows are generally manifested in the spatial structure of urban activity: because population flows represent the economic activities of regions, many zoning systems have used flow data to design zones (see Masser and Brown 1975; Masser and Scheurwater 1980; Rosing and ReVelle 1986; Noronha and Goodchild 1992; Alvanides, Openshaw, and Duke-Williams 2000).

This article is motivated by the practical and empirical need in geography and planning to identify functional regions and delineate their hinterland. Theoretically, a region may be composed of a single core functional unit and other units, or the region can develop multiple centers as economic activities diversify and are extended through the evolution of transportation systems. Functional centers may serve as growth poles that structure changes in labor markets and service areas in urban systems (Coombes et al. 1979; Schwanen, Deileman, and Dijst 2003). Finding evidence of multiple functional centers is key to understanding central place theory. The delineation of subregions with identified functional centers is usually accomplished using the population size of and spatial interactions among units because the influential area of a labor market is formed according to the level of access from the vicinity to the center (Berry and Garrison 1958). These functional regions can be captured at geographic scales ranging from CBDs within a city (e.g., Song and Kim 2015) to daily urban systems at a national level (Green 1958; Coombes et al. 1979; van der Laan 1998; Schwanen, Deileman, and Dijst 2003).

Delimiting functional regions usually involves aggregating small areal units into fewer clusters, which are usually based on spatial contiguity and maintaining similar properties between the areal units in each cluster (Brown and Holmes 1971). More formally, the problem can be defined as the identification of a set of  $p$  groups that are aggregated into  $n$  areal units while optimizing a predefined objective function with a given set of criteria or constraints. Objective functions can be formulated to minimize the dissimilarity of areal units or to maximize their similarity. The criteria for achieving the best representation may include contiguity, compactness, or

socioeconomic homogeneity (Hess et al. 1965; Morrill 1981; Williams 1995). In this regard, the delimitation of functional regions can be directly interpreted as a special case of the  $p$ -regions problem (Duque, Church, and Middleton 2011).

An important issue, from a planning perspective, is how to effectively formulate the regionalization problem mathematically. Traditionally, location-allocation formulations have been emphasized, beginning with Hess et al. (1965). The authors applied location-allocation formulations to political districting problems because the model structure of location-allocation problems is similar to that of regionalization. All demands (from the areal units) are forced to be exclusively assigned to facilities (centers) and the objective function is similarly prescribed based on spatial efficiency, for example, the sum of distances (similarity) from areal units to centers is minimized (Hess et al. 1965; Zoltners and Sinha 1983; Murray and Estivill-Castro 1998). However, the location-allocation approach does not always guarantee the spatial contiguity of the regions (Mehrotra, Johnson, and Nemhauser 1998; Cova and Church 2000). Recently, several researchers have included an explicit contiguity requirement as a constraint in spatial regionalization optimization models (Shirabe 2009; Salazar-Aguilar, Ríos-Mercado, and Cabrera-Ríos 2011; Duque, Church, and Middleton 2011). While the contiguity constraints have produced contiguous regions in previous research, the location of a center in each region may not be structured. If a regionalization problem involves identifying urban or regional systems or delimiting areas such as labor markets or commuting sheds, then the problem should be formulated such that the functional centers of the regions in the solution is to be determined. In addition, it is important to determine how many functional regions exist in the regionalization problem. Although the determination of an ideal number of regions or functional centers can be modeled as an endogenous process in an optimization model, the computational complexity is high even for a problem with a small numbers of areal units, often resulting in intractable solutions (e.g., see Shirabe 2009; Duque, Church, and Middleton 2011). Alternatively, if intractability is a problem, an appropriate number of regions can be determined by exploring all solutions for  $p$ , which can be simply prescribed by explicitly formulating the constraints in the model.

Accordingly, the purpose of this article is to develop a spatial optimization scheme to model the  $p$ -functional regionalization problem and to suggest an efficient method for the exact solution of the problem. The proposed model simultaneously determines a given number of functional centers and delimits their boundaries, while explicitly incorporating contiguity constraints. The article begins with a review of mathematical models for functional regionalization. The second section reviews functional regionalization based on existing mathematical models. The third section provides the model specifications for delimitating functional regions and the fourth section presents the preliminary results of the model. The fifth section introduces an efficient solution approach called *analytical target reduction* (ATR) to find an optimal solution for larger sized regionalization problems. Concluding remarks are made in the final section.

## Background

A vast amount of literature on districting and regionalization exists in operations research. Zoltners and Sinha (1983), Williams (1995), and Duque, Romas, and Suriñach (2007) have written review articles in the field. Here, we focus on studies on functional regionalization with contiguity constraints. Functional regionalization was initially motivated by interest in spatial systems, urban structures, and related regional policy and planning (e.g., Berry and Garrison 1958; Masser and Brown 1975). Functional regions are generally defined based on a greater magnitude of interactions or connections among spatial units within a region than with units outside the region (Brown and Holmes 1971). Therefore, the delimitation of functional regions involves the aggregation of small areal units into larger clusters based on spatial interaction flows such as journey-to-work and interstate migration. In addition, spatial interactions are considered a proxy for the “functional distances” that separate the areal units (Brown and Horton 1970). Early studies that used spatial interactions to delimit functional regions include those by Brown and Holmes (1971), Masser and Brown (1975), and Masser and Scheurwater (1980). Although many conceptual principles were developed in these studies, the operational approaches were restricted due to computational limitations on linear programming at that point in time.

In general, the delimitation of functional regions can be considered a special type of regionalization. A simple approach is to use conventional clustering or partitioning methods instead of formulating mathematical models (see Openshaw 1973; Masser and Brown 1975; Fischer 1980; Masser and Scheurwater 1980). Such methods often require additional processes for ensuring the spatial contiguity of the outcome regions (Openshaw and Wymer 1995). These regionalization methods have also been extended to achieve compactness of the outcome regions (Hess et al. 1965; George, Lamar, and Wallace 1997). Multiobjective models have also been developed (Wise, Haining, and Ma 1997), some of which contain constraints to ensure the spatial contiguity of the outcome regions (e.g., Zoltners and Shinha 1983; Duque 2004). However, because the addition of constraints for spatial contiguity can dramatically increase the complexity of the problem, exact optimization solutions have been limited in practice to small  $n$  cases. Alternatively, relaxed solution frameworks can be considered if the locations of the centers are known (Shirabe 2009) or if the seeds of regions can be prespecified (Duque, Church, and Middleton 2011; Salazar-Aguilar, Ríos-Mercado, and Cabrera-Ríos 2011). Heuristic methods, such as greedy algorithms (Openshaw 1977), simulated annealing, and Tabu search (Openshaw and Rao 1995; Bozkaya, Erkut, and Laporte 2003; Bozkaya et al. 2011; Duque, Anselin, and Rey 2012), are typically employed when global optimal solutions are unattainable because of the complexity of the model. However, heuristic approaches may require contextual knowledge of the areas studied to achieve the best results.

As previously discussed, the main issues in functional regionalization include (1) the objectives of the model relative to the spatial context being modeled, (2) spatial

contiguity, and (3) the tractability of the model in terms of providing an optimal solution. The objective function of regionalization problems is a critical component of the model formulation. The objective function must reflect the spatial context which is being modeled by the optimization process. The effectiveness of a model is normally evaluated by the extent to which the regions identified by the model have more socioeconomic attributes within a regional grouping than between that group and other groupings (Gordon 1999; Cova and Church 2000; Shirabe, 2009). For example, if a model is concerned with spatial efficiency in delineating regions, then the objective function can be formulated to minimize the sum of all distances from the functional spatial unit (i.e., the centroid of the spatial units) to other spatial units within the region (Hess et al. 1965; Hess and Samuels 1971; Murray and Estivill-Castro 1998). By contrast, if the model addresses spatial interactions such as geographic flows, the goal is to focus on interregional dissimilarity relative to the intraregional intensity of flows. The objective function is then to maximize intraregional flows or to minimize interregional flows to determine the regionalization of the spatial units (Plane 1982; Rosing and ReVelle 1986; Alvanides, Openshaw, and Duke-Williams 2000). Many studies have formulated various types of objective functions, but few have discussed the implications of the findings based on such objective functions. For example, many zoning process models delineate political districts or traffic zones using socioeconomic data (e.g., demographic attributes) and spatial interactions (e.g., commuting flows). Here, the concern is the interpretation of where and why a particular geographic area is identified as a functional center, as well as how well the model captures the existing spatial context of the area being studied (Morrill 1981; Williams 1995). In classical location theories, a region is characterized by a functional center (central place), which possesses high-frequency interactions with and its hinterland (Christaller 1966; Lösch 1938). If functional centers are the inflow-dominant units within the delineated region, then these units should represent an administrative unit, a population center, or a service center that provides a major source of job opportunities. By contrast, outflow-dominant units can be characterized as residential or suburban areas where people commute to selected functional centers (Yeates 1963).

Second, as a geographic criterion in regionalization problems, spatial contiguity is regarded as an important condition for eliminating the possibility that a region is made up of spatially discontinuous clusters. In general, a region consists of areal units with homogenous characteristics and is distinguished from other regions. However, a region can be fragmented when there are strong interactions between distant spatial units. For example, in air transportation, major passenger flows are often formed among the dominant functional regions (i.e., hub cities) based on some hierarchical structure, but the spatial units of a grouped region are not necessarily spatially contiguous (Nystuen and Dacey 1961; Wheeler and Mitchelson 1989). In contrast, if a region should be delineated based on geographic flows, such as commuting flows in ground transportation, then the delimitation process for the regions

involves a contiguity restriction, and a functional center is nested within its surroundings with no fragmentation (Berry and Garrison 1958). The process of identifying the functional regions with contiguity can also be applied to administrative boundaries. For example, metropolitan statistical areas are grouped as population core areas surrounded by counties, which have social and economic ties that are measured by commuting flows to workplaces (White House 2009). In this context, the contiguity requirement is prescribed explicitly by a constraint in a regionalization model (Garfinkel and Nemhauser 1970; Mehrotra, Johnson, and Nemhauser 1998; Duque, Romas, and Suriñach 2007). From a modeling perspective, spatial contiguity is often translated into a network tree generation problem to check the validity of contiguity. Areal units and their adjacency relationships are expressed as nodes and edges in terms of a graph, such 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. However, spatial contiguity is a challenging requirement because deriving an exact formula for enforcing the contiguity of spatial units substantially increases the number of variables and constraints.

Spatial contiguity has been modeled using mixed integer programming (MIP) and much effort has been expended on generating efficient constraints based on graph theory. For example, Zoltners and Shinha (1983) proposed a set of constraints using a tree structure to reflect the shortest path adjacency. The underlying logic was that areal units in a region maintain their connectedness in a tree structure when arbitrarily chosen seed units of the region at the  $k$ th level are allocated to a seed of the region when the predecessor of the seed, that is, a tree at the  $(k-1)$ th level, is also allocated to the seed. Shirabe (2005, 2009) used flow-based contiguity constraints to construct a flow-based network for each region. The constraints build  $p$ -partitioned tree networks for a given number of areal units. Suppose a network is formed for each region, after which a sink in the constraints becomes the root of a shortest path tree. If a path has only one sink within a region and receives unit flows from the other nodes (spatial units), then the region is contiguous. Although this prescription can effectively reduce the number of variables and constraints, the model remains limited to relatively small problems unless the number of sinks consisting of regions is known *a priori*. Salazar-Aguilar, Ríos-Mercado, and Cabrera-Ríos (2011) proposed an integer quadratic programming model with subtour elimination constraints that guarantee territory connectivity or contiguity. This model is based on the main principle of first solving a relaxed version of the problem without contiguity constraints. If the solution contains subtours, then the appropriate subtour elimination constraints are added, and the problem is solved again. This process is repeated until a feasible solution is found. However, the obtained feasible solution may not be a global optimum.

Third, a regionalization model, particularly when formulated in terms of MIP with contiguity constraints, may be non-tractable due to complexity. Duque, Church, and Middleton (2011) have presented three strategies for ensuring

contiguity, along with the corresponding MIP models: (1)  $\text{Tree}^{\text{PRM}}$ , which adapts subtour elimination breaking constraints similar to those given by Salazar-Aguilar, Ríos-Mercado, and Cabrera-Ríos (2011); (2)  $\text{Order}^{\text{PRM}}$ , which uses ordered-area assignment variables to prevent cycles and ensure contiguity; and (3)  $\text{Flow}^{\text{PRM}}$ , which uses unit flow constraints similar to those given by Shirabe (2005, 2009). Each strategy has its own distinct strengths in solving problems, but any  $p$ -regions problem that is formulated in the form of MIP belongs to a class of *NP*-hard problems, which means that the solution is ensured only for a limited number of spatial units regardless of the type of contiguity constraints. Any MIP model containing a set of constraints related to tree-, order-, or flow-based structures quickly reaches non-tractability with a small increase in the number of constraints and decision variables. For example, the model  $\text{FLOW}^{\text{PRM}}$ , which is formulated based on Shirabe's model (2005), is known to be the smallest model in terms of the number of constraints, but the computational effort required using commercial optimization software reveals that the model is tractable with optimality only for fewer than twenty-five units with a given level of  $p$  (for further details, see Duque, Church, and Middleton 2011). The difficulty may be due to the constraints that are imposed to prevent cycles when partitioned trees of each region are formed. A method for prespecifying the seeds of the regions (i.e., the initial target seeds) in the model may be introduced to improve the model's tractability by reducing the complexity of the model through reducing the search space of the  $p$ -regions problems (Duque, Church, and Middleton 2011). However, such an approach can be validated only when the areas to be districted are known *a priori*. For this reason, many researchers have used efficient heuristics to solve large problems in practice by sacrificing the assurance of global optimality (e.g., Duque, Anselin, and Rey 2012).

## Problem Specification

### *Formulating the $p$ -functional Regions Problem (PFRP)*

Many districting or regionalization problems focus mainly on constructing districts or regions by grouping basic areal units without necessarily identifying the regional center. However, this article proposes a PFRP, in which functional centers are determined simultaneously with contiguous functional regions. The basic structure of the PFRP combines the principle of location-allocation models with the contiguity constraints inspired by Shirabe (2009) and Duque, Church, and Middleton (2011). As discussed, the location-allocation approach has been widely adopted in modeling districting problems because the basic areal units or demands are exhaustively and exclusively assigned (see Hess et al. 1965; Zoltners and Sinha 1983; Murray and Estivill-Castro 1998), which is essential for identifying centers of regions within a set of flow-based contiguity constraints. The following notation is used in the PFRP:

$i$  = index of basic spatial units,  $i = 1, \dots, n, \in I$ ;

$k$  = index of spatial units selected as functional centers of the region,  $k = 1, \dots, n$ ;

$p$  = the number of regions to be districts;

$$c_{ij} = \begin{cases} 1, & \text{if spatial units } i \text{ and } j \text{ are adjacent} \\ 0, & \text{otherwise;} \end{cases}$$

$N_i = \{j | c_{ij} = 1\}$ , a set of spatial units that are adjacent to  $i$ ;

$a_{ik}$  = the amount of geographic movement from spatial unit  $i$  to regional center  $k$ .

The following decision variables are included in this model:

$$y_{ik} = \begin{cases} 1, & \text{if spatial unit } i \text{ is included in region } k \\ 0, & \text{otherwise;} \end{cases}$$

$$w_{ik} = \begin{cases} 1, & \text{if spatial unit } i \text{ is chosen as a sink of region } k \\ 0, & \text{otherwise;} \end{cases}$$

$f_{ijk}$  = the amount of unit flow from  $i$  to  $j$  in region  $k$  ( $f_{ijk} \geq 0$ ).

Here,  $a_{ik}$  indicates the magnitude of the actual spatial interactions between the basic areal units and the regional centers, such as commuting and migration, to determine the functional regions, whereas  $f_{ijk}$  denotes the conceptual flows between basic spatial units a contiguous region. Based on this notation and the corresponding decision variables, the PFRP can be modeled as follows:

$$\text{Maximize} \sum_k \sum_i a_{ik} y_{ik} \quad (1)$$

Subject to:

$$\sum_k y_{ik} = 1, \quad \forall i, \quad (2)$$

$$w_{ik} \leq y_{ik}, \quad \forall i, k, \quad (3)$$

$$\sum_i w_{ik} \leq 1, \quad \forall k, \quad (4)$$

$$\sum_k \sum_i w_{ik} = p, \quad (5)$$

$$f_{ijk} \leq y_{ik}(n - p), \quad \forall i, j \in N_i, k, \quad (6)$$

$$f_{ijk} \leq y_{jk}(n - p), \quad \forall i, j \in N_i, k, \quad (7)$$

$$\sum_{j \in N_i} f_{ijk} - \sum_{j \in N_i} f_{jik} \geq y_{ik} - (n - p)w_{ik}, \quad \forall i, k, \quad (8)$$

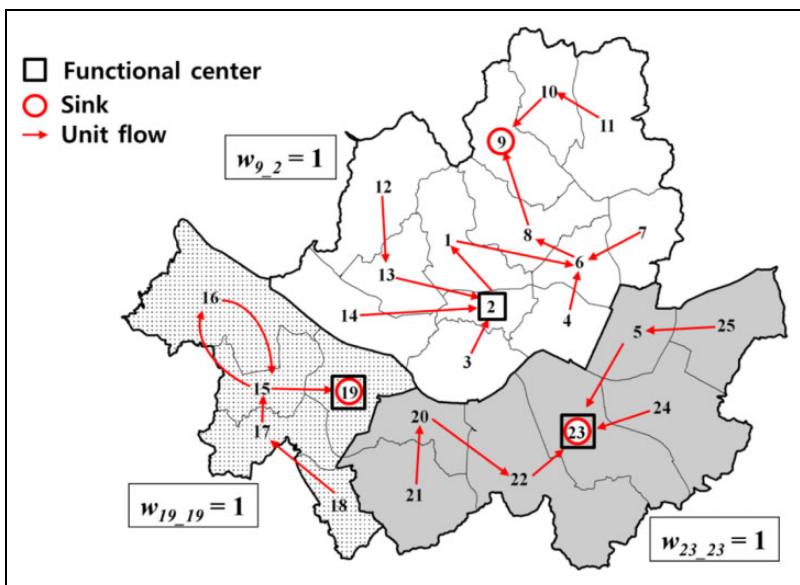
$$y_{ik} = \{0, 1\}, \quad \forall i, k, \quad (9)$$

$$w_{ik} = \{0, 1\}, \quad \forall i, k, \quad (10)$$

$$f_{ijk} \geq 0, \quad \forall i, j \in N_i, k. \quad (11)$$

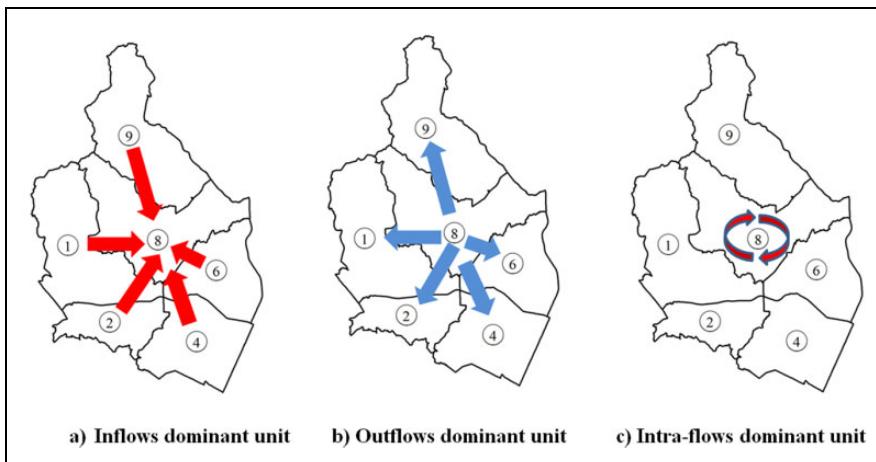
The objective function maximizes the spatial interactions from the assigned areal units to a center within a region; the objective function is defined to capture the relationship between the center and its hinterlands instead of the relationship between all the spatial units assigned to the region because the PFRP determines the functional centers and delineates their sphere of influence simultaneously. Constraint (2) ensures that each spatial unit  $i$  is assigned to only one region  $k$ . These constraints require that all the spatial units are exclusively and exhaustively allocated to regions. Constraint (3) dictates that a sink can be assigned to the region  $k$  only when the spatial unit  $i$  is assigned to the same region. Constraint (4) requires that each region contains only one sink. Constraint (5) states that the total number of sinks is  $p$  ( $p < n$ ). Constraints (4) and (5) together allow only  $p$  regions to be delimited. Constraints (6) and (7) ensure that unit flows can exist between the spatial units  $i$  and  $j$  if and only if both units are adjacent to and assigned to the same region  $k$ . The contiguity requirement for regionalization is established by constraint (8), which enforces the unit flows between the areal units assigned to the same region. If the areal unit  $i$  is not a sink ( $w_{ik} = 0$ ), the areal unit  $i$  must supply at least one unit flow. If the areal unit  $i$  is a sink, the sink can have a nonnegative net flow up to  $(n - p - 1)$ . Here “ $n - p - 1$ ” represents the largest number of areal units that can be assigned to a region. Constraints (9) and (10) impose integer restrictions on the location and allocation decision variables. Constraint (11) ensures that  $f_{ijk}$  is nonnegative.

Although the PFRP has a similar structure to the models of Shirabe (2009) and Duque, Church, and Middleton (2011) from a modeling perspective, there are several important differences. First, the objective function constructs  $p$  regions by maximizing the spatial interactions from the areal units to a functional center within a region without violating the contiguity requirement. Most minimization models do not necessarily require the number of regions to be specified, whereas the PFRP requires that  $p$  is explicitly specified as a constraint (see constraint [5]) because the PFRP is a maximization model. The similarity of this constraint to  $FLOW^{PRM}$  may appear trivial, but combining this constraint with constraints (2), (3), and (4) narrows the solution space. The computational difficulty is thus reduced. The second difference between the PFRP and other models is that the former locates the functional centers through the term  $w_{ik}$ . Figure 1 shows that the decision variables  $w_{ik}$  in the PFRP function differently in finding a solution for  $p = 3$  compared to the same



**Figure 1.** Unit flows in optimal solution for  $p = 3$ .

variables in other models, for example,  $FLOW^{PRM}$ . The index  $k$  denotes an areal unit that is selected as a functional center. By contrast, in other models,  $k$  indicates the number of  $p$ -partitioned regions that are imposed as ordered numbers ( $k = 1, 2, 3, \dots, p$ ; see Shirabe 2009; Duque, Church, and Middleton 2011). The index  $k$  in  $w_{ik}$  in the PFRP ranges from 1 to  $n$ , instead of from 1 to  $p$ , because the model should select  $p$  centers from  $n$  areal units, as well as delineating the boundaries of the  $p$  regions. If  $k$  has a value between 1 and  $p$ , then the PFRP cannot track the location of a region's functional center. The main rationale of this treatment is to identify the sink and the functional center for each region during the regionalization process. This notational treatment allows the variable  $w_{ik}$  to be interpreted in a different manner in the PFRP. For example,  $w_{9\_2} = 1$  in Figure 1 indicates that areal unit 9 is selected as a sink where all unit flows come together from all areal units within a region centered on areal unit 2. By contrast, comparable variables in other models usually represent areal unit 9 as a sink of the second delineated region among  $p$  regions. The variables  $w_{19\_19}$  and  $w_{23\_23} (=1)$  indicate that areal units 19 and 23 are both functional centers and sinks for a region. Note that the selected functional center in the PFRP is not necessarily identical to the sink of the unit flow in each region because constraint (8) determines the sink of a region among the areal units in the region. In addition, considering that the goal of the objective function is to designate the selected spatial unit  $k$  as the functional center that maximizes spatial interactions from other areal units  $i$  within a region, the selected spatial unit  $k$  represents a central



**Figure 2.** Types of dominant spatial unit in geographic flows.

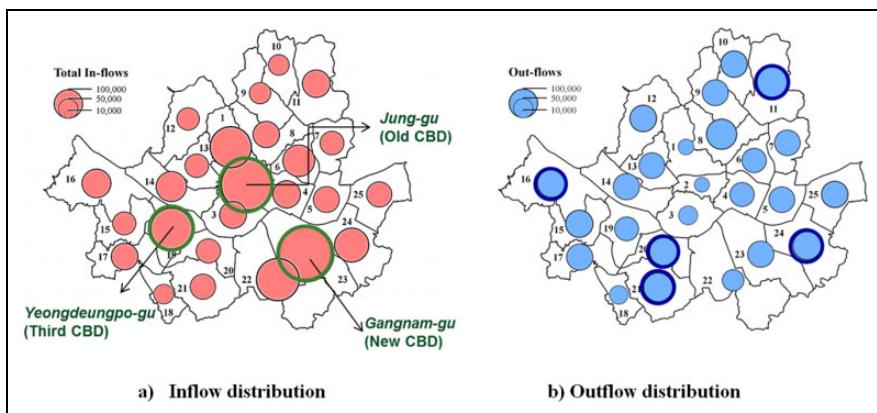
place or business center (i.e., a CBD) where geographical flows from other centers are highly concentrated.

The logic of the PFRP in aggregating areal units into a set of functional regions can be explained by determining the dominant flows of the geographic inflows, outflows, and intraflows. As illustrated in Figure 2, if an areal unit (unit 8 in Figure 2a) receives substantial inflows from other units with small outflows, then this inflow-dominant unit is very likely to be the region's functional center. By contrast, if an areal unit has dominant outflows to some vicinity (Figure 2b), then this areal unit is likely to be part of the region's hinterlands. As a special case, if an areal unit has substantial intraflows with small inflows and outflows (Figure 2c), then the areal unit is considered a self-sustaining region that is not included in any other region. The main characteristic of the daily commuting flow is that geographic proximity in journey-to-work flows is a critical factor in delineating the space of economic activities and that most job opportunities are usually more accessible in a central areal unit that acts as a functional center and whose influential space diminishes with distance (Kim et al. 2012).

## Results: Seoul Case Study

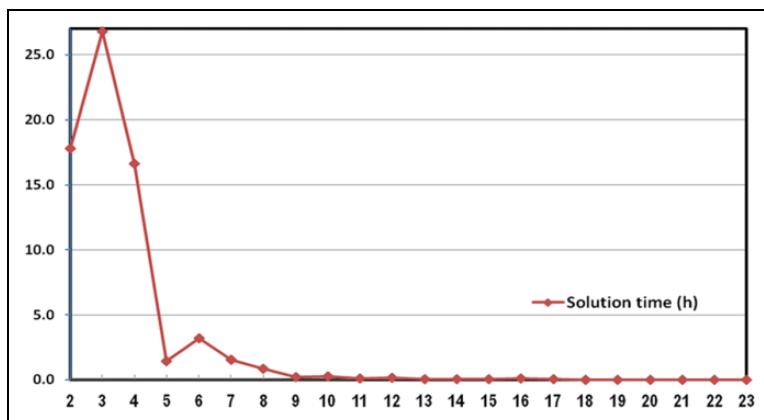
### Data

We apply the PFRP to Seoul, South Korea. We consider Seoul for several reasons. First, the number of spatial units is appropriate for examining the solvability of  $p$ -regions problems. Seoul consists of twenty-five wards ("gu" in Korean). From previous research, an optimal solution is difficult to find in most districting models for a problem size of more than twenty-five areal units; problems involving sixteen units



**Figure 3.** Daily commuting flows in Seoul (2005).

at a particular  $p$  level are often even harder to solve if there is a contiguity requirement in the model (Shirabe 2009; Duque, Church, and Middleton 2011). Second, identifying functional regions in Seoul has been an important issue in geography, regional science, and planning because since the 1970s, the city's labor and housing markets have changed, along with periodic changes in economic development plans. Historically, Seoul has evolved from a single central region into multiple functional regions over the last thirty years. Therefore, the regionalization of the city with an appropriate number of regions has been recognized as a critical task, not only for the effective administration and zoning of the city's labor market but also for the strategic provision of Seoul's transportation infrastructure (Song and Kim 2015). However, few studies have employed mathematical models to district and identify functional regions and their influential areas in terms of labor and services. To this end, the PFRP results can facilitate the decision-making process in terms of designating current CBDs and subcenters of Seoul as new centers of economic activity. For geographic flows, we consider the daily origin–destination commuting flows among Seoul's twenty-five wards for the year 2005. Figure 3 shows the spatial pattern of total inflows (Figure 3a) and outflows (Figure 3b) for these twenty-five areal units. Areal units 2 (Jung-gu, a conventional CBD), 23 (Gangnam-gu, a new CBD), and 19 (Yeongdeungpo-gu, a third CBD) are recognized as Seoul's three major CBDs because these areal units host many firms and thus receive substantial inflows from other areal units. By contrast, areal units 11, 16, 20, 21, and 24 show dominant outflows. With rapid advances in transportation systems and fluctuations in the housing market, these units have been developed as traditional residential areas (Song and Kim 2015). This raises the question of which units act as CBDs and how their influential boundaries can be delineated in terms of economic service areas of each unit. Figure 3 shows the locations of the three functional centers identified by the PFRP, which supports the findings of Song et al. (2012) who have identified the same CBDs in Seoul.

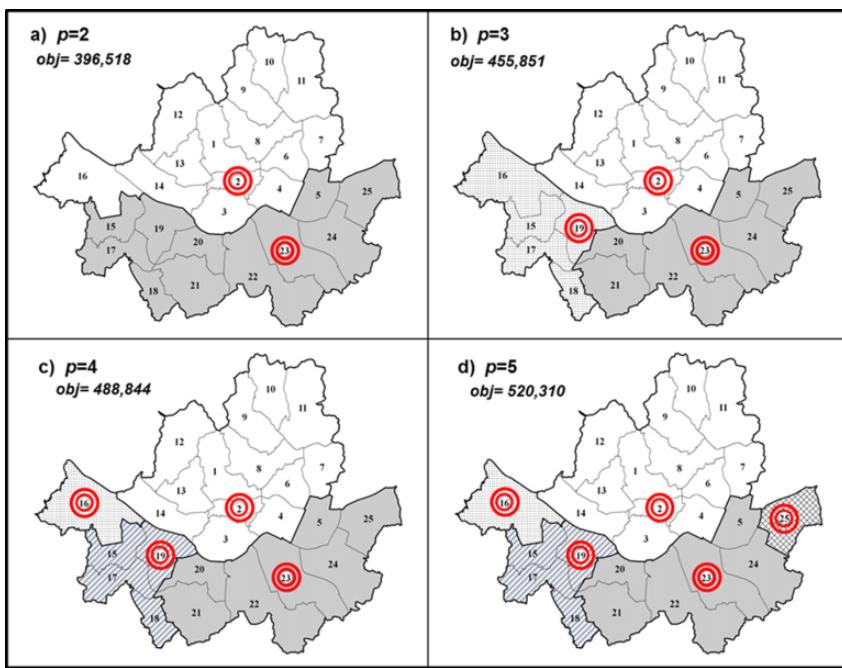


**Figure 4.** Solution times (in hours) for instances from  $p = 2$  to  $p = 23$  ( $n = 25$ ).

### Computational Results

The test cases in this study, from  $p = 2$  to  $p = 23$ , are successfully solved with optimality. To solve the different cases, we use a CPLEX 12.1 with Intel's Xeon Quad Core 3.2 GHz/3 GB RAM based on the Windows 7 operating system; the PFRP results of the problem cases are visualized using ESRI ArcGIS 9.3. Figure 4 shows the distribution of solution times for all cases to investigate the actual computational effort needed to solve the PFRP. The worst computational performance for the PFRP is observed at lower levels of  $p$  (2, 3, and 4), taking approximately sixteen to twenty-six hours; however, optimal solutions are found for the remaining 87 percent of the cases in less than thirty minutes. In fact, 50 percent of the cases are solved in just a few minutes. This indicates that the PFRP can outperform other models in practice for moderate-sized regionalization problems regardless of the level of  $p$ .

Figure 5 shows the optimal locations of the functional centers and districted regions for  $p$  ranging from 2 to 5. Here, we make the following observations. First, given the flow patterns in Figure 3, the interactions between areal units may be influenced by the behavior of commuters in Seoul who tend to live near CBDs to avoid a long commute to work. Note that the locations of the functional centers identified by the PFRP show how CBDs have emerged as Seoul has developed economically. Figure 5a, for  $p = 2$ , identifies Seoul as a city with two distinct functional centers: (1) Jung-gu (unit 2), a conventional CBD, and (2) Gangnam-gu (unit 23), a newly established CBD. At  $p = 3$ , another regional CBD, Yeongdeungpo-gu (unit 19) is added to the western region of Seoul by dividing the area covered by Gangnam-gu at  $p = 2$ . The areal unit 19 is currently recognized as a subfunctional center in the southern region of Seoul, which is consistent with this districting result. Note that the major residential regions (units 11, 16, and 24) shown in Figure 3b are nested within the three CBDs. At  $p = 4$  and 5, the spatial units 16 and 25, respectively, are newly



**Figure 5.** Optimal solutions of Seoul ( $p = 2\text{--}5$ ).

added, implying that these spatial units emerge as subfunctional centers from previous regionalization results, whereas the conventional CBD (unit 2) maintains its dominant position as a CBD regardless of the level of regionalization. This result suggests that labor markets become fragmented through the emergence of multiple centers.

Second, regarding computational aspects, the inclusion relationship between selected sets of centers is observed at different levels of  $p$ . In addition, a set of optimal solutions for functional centers at  $p - 1$  is found as a nested set of functional centers at  $p$ , or functional centers at a specific level of  $p$  can be a predefined set at  $p - 1$ . Table 1 implies that the optimal locations of the regional centers at  $p = 2$  (centers 2 and 23) are included at  $p = 3$  and that the set of functional centers at  $p = 3$  becomes part of the optimal locations at  $p = 4$ . In other words, an optimal solution at  $p$  becomes a candidate set of centers at lower levels of  $p$  ( $p - 1$ ,  $p - 2$ , ..., and 2). In the test cases, this inclusion relationship is observed from  $p = 2$  to  $p = 20$ , with increasing  $p$ , in all cases except at  $p = 15$ . The nested-set relationship can be used to predict which units are potential candidates for functional centers. This characteristic motivates a potential solution strategy that may guarantee global optima based on the optimal solutions that are identified at a selected level of  $p$ . As indicated by the three columns on the right-hand side of Table 1, there is

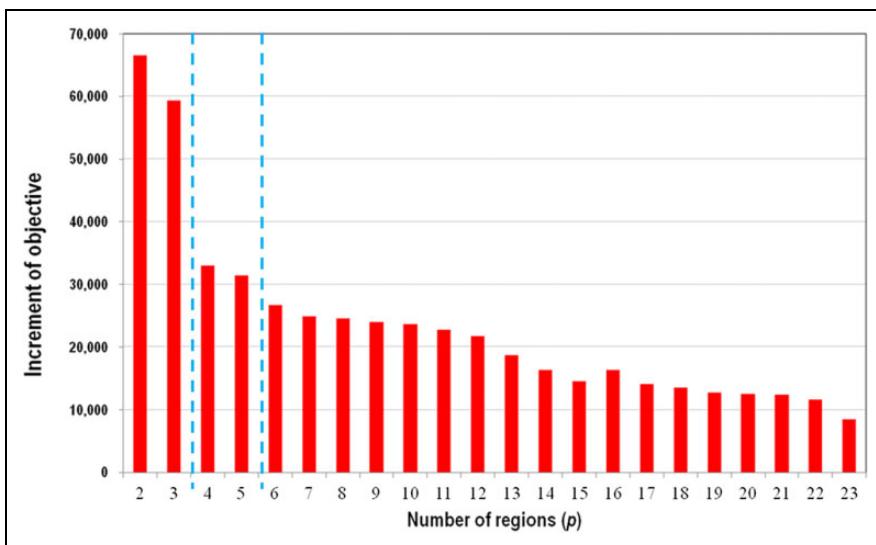
**Table 1.** Computational Results for  $p = 2$  to  $p = 20$ .

$n$	$p$	Objective	Regional centers	Solution time (seconds)	Iterations	Nodes
25	2	396,518	2, 23	64,079	277,273,010	1,128,647
25	3	455,851	Solution set at $p = 2 + 19$	96,416	358,923,930	1,835,554
25	4	488,844	Solution set at $p = 3 + 16$	59,878	258,330,152	1,365,187
25	5	520,310	Solution set at $p = 4 + 25$	5,121	24,597,894	152,677
25	6	547,029	Solution set at $p = 5 + 6$	11,520	54,995,215	504,360
25	7	571,863	Solution set at $p = 6 + 7$	5,506	78,369,039	574,614
25	8	596,464	Solution set at $p = 7 + 12$	3,037	46,693,117	378,875
25	9	620,456	Solution set at $p = 8 + 21$	799	12,749,791	125,034
25	10	644,104	Solution set at $p = 9 + 24$	867	13,569,523	137,718
25	11	666,833	Solution set at $p = 10 + 11$	369	5,551,374	63,123
25	12	688,582	Solution set at $p = 11 + 14$	504	10,509,338	141,060
25	13	707,236	Solution set at $p = 12 + 17$	109	1,494,960	26,149
25	14	723,520	Solution set at $p = 13 + 4$	89	762,827	12,489
25	15	738,098	Solution set at $p = 13 + 8, 9$	108	1,132,701	17,323
25	16	754,382	Solution set at $p = 15 + 4$	428	5,417,135	130,538
25	17	768,465	Solution set at $p = 16 + 18$	132	1,644,992	28,416
25	18	782,006	Solution set at $p = 17 + 5$	23	227,759	5,144
25	19	794,759	Solution set at $p = 18 + 3$	19	228,578	6,028
25	20	807,217	Solution set at $p = 19 + 15$	7	33,287	1,180

some computational difficulty associated with finding optimal solutions of the PFRP at lower levels of  $p$ . If any case with a lower  $p$  value suffers from computational complexity, then the search for the optimal solution at  $p$  can be initiated from the optimal solution sets at a higher level of  $p$  because the elimination of any candidates can tighten the solution space and thus help ensure global optima.

Third, an ideal number of regions for delimiting the functional regions can be determined based on changes in the objective values. As summarized in Table 1, the objective value increases with  $p$ , indicating that the efficiency of the spatial interactions increases as regions are more regionalized in the model. Furthermore, the results in Figure 6 suggest that increments in the objective values with  $p$  may indicate natural breaks for effectively delimiting the functional regions. Significant changes are observed when  $p$  changes from 3 to 4 and 5 to 6, indicating that an ideal regional district corresponds to  $p = 3$  or  $p = 5$  in the case of Seoul. As discussed in the previous section, the regionalization process in Seoul during the last three decades is well represented by both districting results.

For  $p = 3$ , the three CBDs reflect the mid-2000s, when the southern and northern regions of Seoul developed as distinct economic areas, and the southern region was divided into two functional districts, namely functional centers 23 (a new CBD) and 19 (the third CBD; Song et al. 2012). As long as regions are districted in a nested structure with an increase in  $p$ , a discussion on the hierarchy of  $p$  regions is possible. For example, at  $p = 5$ , spatial units 2 and 23 are the top-tier centers, followed by spatial unit



**Figure 6.** Increments of objective values with  $p$  (Seoul). Note: The increment is calculated by the difference in the objective value between at  $(p - 1)$  and  $p$ .

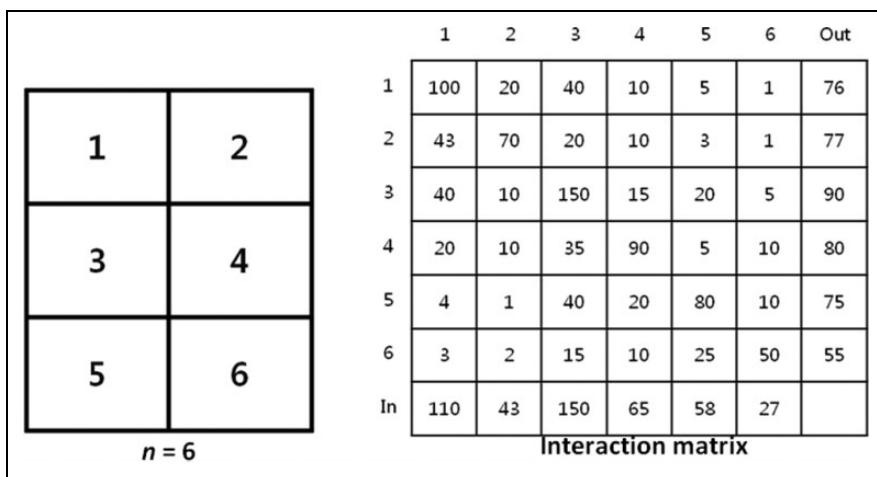
19 as a second-tier center and spatial units 16 and 25 as subregional centers of 19 and 23, respectively. Based on central place theory, these regionalization patterns indicate service areas or hinterlands for the  $p$ -functional centers, as well as the regional hierarchy.

## Application to a Larger Problem: ATR

### *Model Behavior of the PFRP*

Table 1 seems to suggest that the model can solve a set of more than twenty-five areal units. In general, given  $n$  areal units, computational time rapidly decreases as the number of  $p$  regions increases. Large computational times are only observed for lower levels of  $p$  with a higher level of complexity in terms of the numbers of constraints and variables. Thus, large computational times may be a problem if exact solutions are required at those levels. In terms of the model structure, note that constraints (6), (7), and (8) include the constant  $(n - p)$ , which dictates the maximum number of allowable areal units constituting a region. In other words, a decrease in  $p$  increases the value of the constant, implying that the possibility of unit flows among the assigned spatial units also increases. In addition, an increase in  $p$  increases the objective value because most areal units have intraflows that far exceed the maximum outflows. In contrast, a decrease in  $p$  allows the intraflows of many areal units to be substituted for the largest, second largest, or third largest outflows.

As noted earlier, an interesting observation from Table 1 is that the set of centers selected at  $p - 1$  is a subset of centers obtained for the range of  $p$ . This observation is



**Figure 7.** Sample data with six spatial units.

due to the behavior of the objective function of the model. Basically, the model's objective function searches for potential centers from either (1) areal units receiving large inflows from unselected areal units or (2) units with large intraflows. In theory, areal units with relatively small intraflows or large outflows can be considered potential centers. However, the model tends to designate areal units as potential centers if there are substantial differences between the magnitude of the intraflows and that of the largest (hereafter referred to as DIFF1) or second largest (hereafter referred to as DIFF2) outflows from a unit. These differences represent the impact on the objective value when a spatial unit is assigned to a center instead of being selected as a center. If only DIFF1 and DIFF2 are considered in the model, then the model has similar characteristics to the Nystuen–Dacey model in which centers are determined only by dominant flows from other nodes. However, note that the contiguity requirement of the model restricts the center to be within delimited regions. As a result, regions cannot be organized in a nonplanar type of hierarchy.

This behavior of the model rarely changes when the level of  $p$  decreases or increases. The results indicate that it is very unlikely that areal units that are not selected as potential centers at  $p$  are selected as centers at  $p - 1$ . Using hypothetical data with six spatial units, Figure 7 and Table 2 show conceptually and in detail how the PFRP functions in determining the locations of centers and assigning unselected areal units to the centers. As an initial step in understanding the behavior of the model, DIFF 1 and DIFF 2 are calculated by subtracting the largest and second largest outflows, respectively, of an areal unit from its intraflows (see Table 2). For  $p = 5$ , areal unit 6, which has the smallest effect (25) on the objective value in terms of DIFF1, maximizes the objective function when areal unit 6 is assigned to areal unit 5. Similarly, for  $p = 4$ , spatial unit 2, for which DIFF1 corresponds to the second

**Table 2.** Model Behavior with Analytical Target Reduction.

Unit	Intra flows (A)	Largest outflows (B)	Second largest outflows (C)	DIFFI <sup>a</sup> (A-B)	DIFF2 <sup>b</sup> (A-C)	Assignment	Objective	Illustration
(a) $p = 5$ , objective = 515	100 2 3 4 5 6	40 (3) 43 (1) 40 (1) 35 (3) 40 (3) 25 (5)	20 (2) 20 (3) 20 (5) 20 (1) 20 (4) 15 (3)	60 27 110 55 40 25	80 50 130 70 60 35	Center Center Center Center Center Center 5	100 70 150 90 80 25	n = 6
(b) $p = 4$ , objective = 488	100 70 150 90 80 50	40 (3) 43 (1) 40 (1) 35 (3) 40 (3) 25 (5)	20 (2) 20 (3) 20 (5) 20 (1) 20 (4) 15 (3)	60 27 110 55 40 25	80 50 130 70 60 35	Center Center Center Center Center Center 5	100 43 150 90 80 25	n = 6
(c) $p = 3$ , objective = 433	100 70 150 90 80 50	40 (3) 43 (1) 40 (1) 35 (3) 40 (3) 25 (5)	20 (2) 20 (3) 20 (5) 20 (1) 20 (4) 15 (3)	60 27 110 55 40 25	80 50 130 70 60 35	Center Center Center Center Center Center 5	100 43 150 90 80 25	n = 6
(d) $p = 2$ , objective = 383	100 70 150 90 80 50	40 (3) 43 (1) 40 (1) 35 (3) 40 (3) 25 (5)	20 (2) 20 (3) 20 (5) 20 (1) 20 (4) 15 (3)	60 27 110 55 40 25	80 50 130 70 60 35	Center Center Center Center Center Center 5	100 43 150 90 80 25	n = 6

Note: <sup>a</sup>IFFI: difference between the amount of intraflow and the largest outflows in terms of a spatial unit.

<sup>b</sup>DIFF2: difference between the amount of intraflow and the second largest outflows in terms of a spatial unit. The gray colored lines indicate the selected units as functional centers.

smallest effect, becomes a unit to be dropped from a set of candidate centers at  $p = 5$  and is assigned to center 1. Districting three regions is more complicated because the contiguity constraint must be considered. For  $p = 3$ , although areal unit 5 has the third smallest effect after units 2 and 6, areal unit 5 is already grouped with unit 6. If unit 5 is assigned to unit 3, then unit 6 is assigned to unit 3 or 4 based on contiguity. If unit 5 is assigned to 3 and unit 6 is assigned to 4, then the decrement in the objective value is 80 ( $5 \rightarrow 3: 40; 6 \rightarrow 4: 40$ ). Alternatively, if both units 5 and 6 are assigned to unit 3, then the objective value decreases by 75 ( $5 \rightarrow 3: 40; 6 \rightarrow 3: 35$ ), corresponding to a better grouping. However, the best case is found when assigning units 2, 4, and 6 to units 1, 3, and 5, respectively. This assignment results in the minimum decrement (55), thus maximizing the objective value. Finally, when a similar process is applied to  $p = 2$ , the maximum objective value is found once unit 2 is assigned to center 1 and units 4, 5, and 6 are grouped around center 3.

## ATR

Based on this hypothetical example and the empirical experiment, we propose the ATR method to solve  $p$ -regions problems for a larger number of areal units. At a given  $p$  level, the global optimal or near-optimal solutions are found by considering the solution set found at the  $p + q$  level, rather than for all combinations of areal units ( $n$ ) with functional centers. Here,  $q$  represents the ATR step size where  $1 \leq q \leq (p - 1)$ . ATR is based on the principle that the areal unit with the smallest contribution to the objective value during the delimitation of  $p$  regions is likely to have the smallest impact on the objective value when delineating  $(p - q)$  regions. Further, the centers that are selected at a level  $p$  are likely to be selected as centers at a lower level. This suggests that unselected areal units in the  $p$ -regions problem can be reasonably excluded from a set of potential centers when solving the  $(p - q)$ -regions problem.

ATR proceeds by the following steps. First, the  $p = (n - q)$  case is solved. In this step, the set of potential seeds is  $I (|I| = n)$ . The procedure can be started at an intermediate level of  $p$ , that is, with a high  $q$  value, but a relatively high  $q$  value may require significantly longer computation times. Second, unselected spatial units are eliminated from the set of potential centers and are added to the set of assigned spatial units  $U$ . Third, two additional constraints, (12) and (13), restrict the areal units in  $U$  to be centers and sinks for the next ATR case:

$$y_{ij} = 0 \quad \forall i, j \in U, \quad (12)$$

$$w_{ij} = 0 \quad \forall i, j \in U. \quad (13)$$

Fourth, the previous steps are repeated until  $p = 2$  or  $p < q$  is reached. This approach has considerable benefits because the complexity of the problem decreases significantly from  $_nC_p$  to  $_pC_{(p-q)}$  at the location level. The complexity of the

allocation level decreases because of the narrowed search space at a given location level, although the contiguity structure among the areal units may influence the solution efficiency for a particular case. It should be noted that although a lower  $q$  value requires less computational effort in finding the optimal solution, the probability of finding the global optimum increases, along with the solution time, as the value of  $q$  increases.

The performance of ATR can be improved by considering another important factor from Tables 1 and 2. Note that some areal units are consistently selected as functional centers regardless of the  $p$  level. In general, such units (denoted as the set  $S$ ) have substantial intraflows (self-sustaining spatial units) or inflows from other areal units. An areal unit with dominant inflows from other units can be regarded as a functional center. Thus, the solution space can be reduced and the computational efficiency improved by considering one or two areal units as centers regardless of the  $p$  level. Applying this concept to ATR requires the following constraints (14) and (15):

$$y_{ii} = 1 \quad \forall i \in S, \quad (14)$$

$$w_{ii} = 1 \quad \forall i \in S. \quad (15)$$

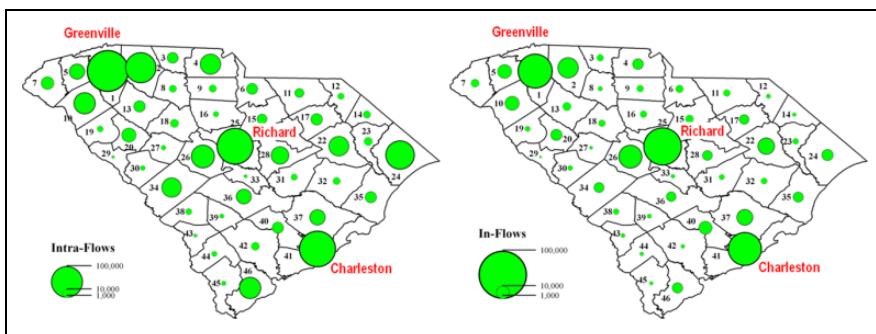
Finally, note that the exception that is observed at  $p = 15$  in Table 1 may occur when ATR is applied with  $q = 1$ . This exception occurs under three conditions: (1) the largest outflows of an areal unit are to a center that is not adjacent to the areal unit (which is defined as the first law of geography), (2) no unit flow route is ensured from the areal unit to the center, and (3) areal units that similarly impact the objective value are distributed around the center. This problem can be resolved using a higher  $q$  value.

Table 3 demonstrates how ATR models are more effective compared to the original model. An ATR model substantially reduces the computational time for all cases, without losing optimality. The longest computational time is less than five minutes, with most cases requiring only 1 percent of the solution time of the original model. In addition, the ATR model shows a 99 percent reduction in the number of iterations relative to the original model. Not surprisingly, these results are due to significant reductions in complexity at each location level, that is, the ATR model selects  $p - 1$  functional centers from  $p$  predetermined centers. Among the twenty-two cases, the solution at  $p = 14$  is the only case in which a solution set is not covered by a set at the next higher level,  $p = 15$  ( $q = 1$ ). Thus, the remaining twenty-one cases (95.5 percent) are globally optimal solutions. If a more conservative approach is required, then a higher  $q$  value can be used and computational times remain reasonable. The third column of Table 3 (an ATR model with  $q = 2$ ) shows that all solution sets are globally optimal and that the longest computational time is only 16.9 minutes, which is only 1.6 percent of the computational time for the original model. In addition, the ATR model reduces the number of iterations required by

**Table 3.** Comparison of Computational Results: Original versus ATR Models.

$p$	Objective	Original model			ATR model ( $q = 1$ )			ATR model ( $q = 2$ )			
		Time (seconds)	Iterations	Nodes	Objective (seconds)	Time (seconds)	Iterations	Nodes	Objective (seconds)	Time (seconds)	Iterations
2	396.518	64,079	277,273,010	1,128,647	396,518	279.2	31,369,018	1,047,757	396,518	1,013.7	75,206,773
3	455.851	96,416	358,923,930	1,835,554	455,851	51.2	5,369,887	286,474	455,851	122.5	12,354,844
4	488.844	59,878	258,330,152	1,365,187	488,844	34.9	3,984,860	144,726	488,844	664.3	61,189,028
5	520.310	5,121	24,597,894	152,677	520,310	15.5	1,604,440	59,226	520,310	86.1	6,337,854
6	547.029	11,520	54,995,215	504,360	547,029	3.5	320,373	12,730	547,029	255.6	19,905,918
7	571.863	5,506	78,369,039	574,614	571,863	2.2	152,495	6,590	571,863	157.2	11,593,149
8	596.464	3,037	46,693,117	378,875	596,464	1.3	71,136	2,980	596,464	179.1	12,572,290
9	620.456	800	12,749,791	125,034	620,456	1.7	109,222	6,536	620,456	55.8	3,041,191
10	644.104	867	13,569,523	137,718	644,104	1.7	105,221	6,638	644,104	46.1	2,193,551
11	666.833	369	5,551,374	63,123	666,833	0.9	25,767	1,622	666,833	45.9	1,889,692
12	688.582	504	10,509,338	141,060	688,582	0.8	12,031	923	688,582	14.2	411,107
13	707.236	109	1,494,960	26,149	707,236	0.9	16,020	1,058	707,236	6.4	218,008
14	723.520	89	762,827	12,489	721,677	0.8	1,2853	1,161	723,520	5.5	211,467
15	738.098	108	1,132,701	17,323	738,098	0.8	6,646	545	738,098	5.1	248,299
16	754.382	428	5,417,335	130,538	754,382	0.5	5,167	501	754,382	2.4	72,781
17	768.465	132	1,644,992	28,416	768,465	0.4	4,188	498	768,465	1.4	33,464
18	782.006	23	227,759	5,144	782,006	0.6	3,570	473	782,006	1.3	14,237
19	794.759	19	228,578	6,028	794,759	0.5	3,276	496	794,759	1.7	19,361
20	807.217	7	33,287	1,180	807,217	0.7	8,119	516	807,217	1.3	14,518
21	819,639	4	18,690	600	819,639	0.2	622	11	819,639	0.8	8,080
22	831,203	1	17,606	496	831,203	0.2	610	9	831,203	1.0	17,606
23	839,645	1	2,950	312	839,645	0.5	2,950	312	—	—	—

Note. ATR = analytical target reduction. All instances are solved with optimality in both models using CPLEX 12 executed on a Intel Xeon Quad Core 3.2 GHz, 3 GB RAM with Window 7 operating system.



**Figure 8.** Commuting flows in South Carolina (2000).

the original model by 72.9 percent. These results demonstrate that ATR is a promising approach for a large size problem when optimal solutions are desired. This approach is particular applicable to regionalization problems involving spatial interactions such as migration and commuting with a short-distance tendency (Ravenstein 1885; Tobler 1987) or interactions derived from the central place theory (Christaller 1966).

### *Application of ATR: South Carolina Case Study*

Here, we test the PFRP model by incorporating ATR for a large data set, namely forty-six counties in South Carolina using journey-to-work data from the 2000 US Census. As with the previous  $p$ -regions models, the MIP version of the PFRP is limited to a set of approximately twenty-five areal units for solving the problem with optimality. Thus, we apply ATR to assess its impacts on the solution capability of the original PFRP. To predetermine the initial seeds, Figure 8 provides the spatial pattern of interactions among the forty-six counties in terms of intraflows and inflows, which are used as a reference when adding constraints (14) and (15). Two counties (Greenville and Richard) show dominant intraflows (Figure 8a) as well as inflows. In contrast, there are several counties in the southwestern and northern areas that show minimal intraflows and/or inflows (Figure 8b). Intuitively, these small counties are clearly not likely to be functional centers, whereas the two counties with dominant intraflows are.

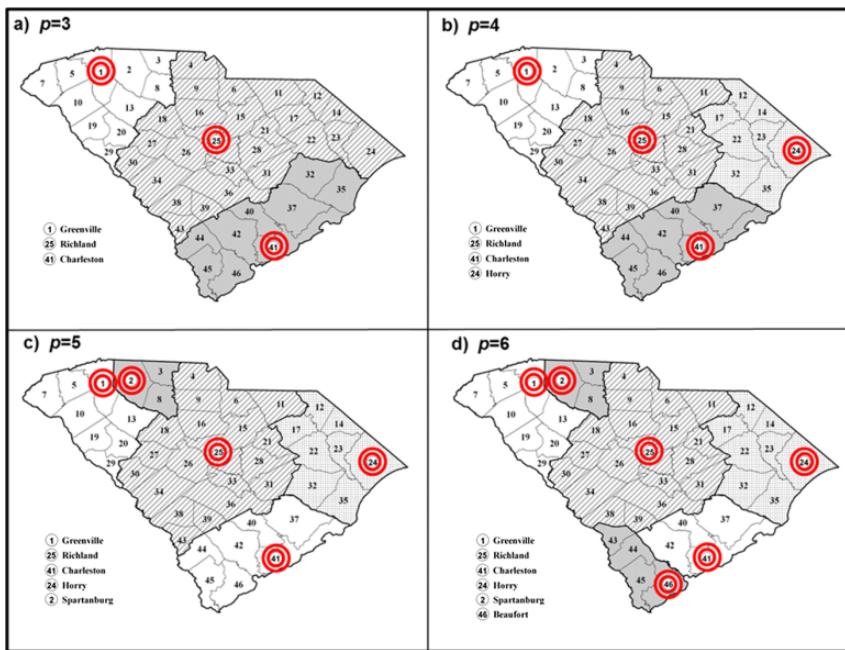
Table 4 shows the computational results obtained for the PFRP by applying ATR with  $q = 1$  sequentially from  $p = 44$  to  $p = 2$ . Because the optimal functional centers are identified at each  $p$ , those units that are not selected as functional centers at the  $p - 1$  level are excluded using constraints (11) and (12). As a result, optimal solutions for thirty-six cases (84 percent) are derived for  $p = 9$  to  $p = 44$ , for which the computational time ranges from several seconds to a few hours. Note that seven cases are terminated due to memory problems, with a gap ranging from 3.78 percent to 41.7 percent. However, despite the gap, the objective values are global optima,

**Table 4.** Computational Results of the  $p$ -Functional Regions Problem Using the Analytical Target Reduction.

$p$	Objective	Solution time (sec)	Iterations	Nodes	Gap (%)	Spatial units <sup>b</sup>
44	1,322,352	0.8	2,278	0	0.0	40, 45
43	1,322,044	0.9	1,908	0	0.0	33
42	1,321,678	0.7	1,856	0	0.0	37
41	1,320,796	0.8	1,746	2	0.0	29
40	1,319,628	0.8	1,706	0	0.0	30
39	1,317,876	1.8	4,490	455	0.0	21
38	1,316,111	0.9	2,177	219	0.0	43
37	1,313,973	1.2	4,267	523	0.0	16
36	1,311,759	1.5	5,075	498	0.0	27
35	1,309,286	1.3	4,307	524	0.01	39
34	1,305,757	3.0	5,521	529	0.01	44
33	1,302,130	2.7	5,222	493	0.01	19
32	1,298,398	3.2	31,664	2,425	0.01	31
31	1,293,539	9.6	232,240	9,463	0.01	38
30	1,288,513	3.7	64,551	3,736	0.01	8
29	1,283,462	19.3	361,931	16,051	0.01	9
28	1,278,306	16.4	247,702	11,796	0.01	12
27	1,272,019	21.0	666,641	34,639	0.01	32
26	1,265,091	9.0	409,807	17,732	0.01	15
25	1,257,987	6.3	177,971	9,395	0.01	14
24	1,249,928	9.4	423,262	15,841	0.01	42
23	1,241,708	34.5	783,609	27,587	0.01	18
22	1,233,349	20.4	958,729	37,824	0.01	23
21	1,224,154	72.2	3,346,838	68,580	0.01	13
20	1,214,398	44.6	2,046,219	54,677	0.01	17
19	1,202,283	382.4	16,710,989	452,645	0.01	3
18	1,189,973	521.8	20,872,749	592,398	0.01	11
17	1,177,028	940.3	39,820,024	681,754	0.01	35
16	1,163,258	7,129.9	249,147,517	3,776,979	0.01	6
15	1,149,402	14,193.3	493,651,814	538,534	0.01	5
14	1,133,758	24,265.2	804,460,023	7,750,666	0.01	26
13	1,114,914	12,064.7	429,158,698	3,403,243	0.01	7
12	1,088,994	5,186.7	151,532,959	1,511,818	0.01	20
11	1,061,972	749.8	20,494,435	161,889	0.01	36
10	1,023,965	15,372.7	324,999,231	1,291,539	0.01	28
9	984,035	38,443.7	100,011,578	2,406,117	0.01	10
8 <sup>a</sup>	936,770	28,946.2	535,571,031	1,842,094	3.71	34
7 <sup>a</sup>	883,587	26,465.8	452,195,119	1,187,111	5.3	4
6 <sup>a</sup>	827,928	19,747.0	393,610,468	1,861,842	6.1	22
5 <sup>a</sup>	770,606	10,623.0	214,255,298	857,179	6.9	46
4 <sup>a</sup>	683,596	7,284.6	216,684,425	1,187,180	12.2	2
3 <sup>a</sup>	590,381	15,531.6	427,453,529	2,606,256	14.6	24
2 <sup>a</sup>	411,377	21,308.2	595,780,096	2,914,351	41.7	41

Note. <sup>a</sup>The solutions are found with a gap after the termination of out-of-memory situations in current computational environment but are proven to be as optimal by enumeration-based results of all possible cases for  $p = 2\text{--}8$ .

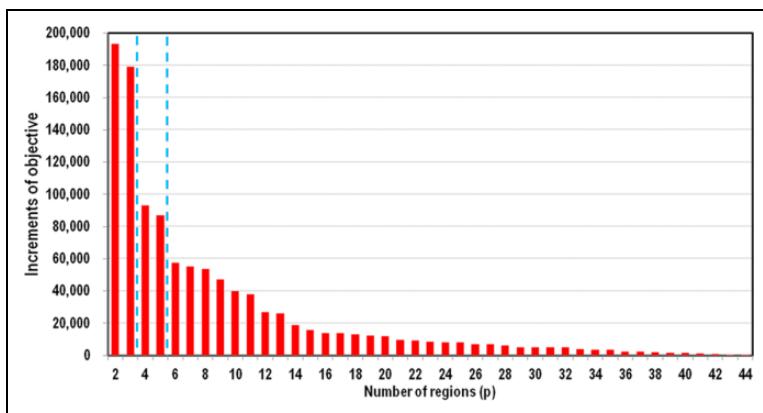
<sup>b</sup>Spatial units which are excluded cumulatively for the solving process of the next lower  $p$ .



**Figure 9.** Optimal solutions of South Carolina ( $p = 3\text{--}6$ ).

which are verified through enumerations for  $p = 2\text{--}8$ . Note also that most of the excluded spatial units are counties with minimal intraflows or inflows from other counties, whereas the top ten counties (10, 34, 4, 22, 46, 2, 24, 41, 25, and 1) satisfying these criteria remain in the set of potential centers until the sequential process reaches  $p = 9$ .

Figure 9 demonstrates that ATR is appropriate, not only for the PFRP but also for other districting model variants. As expected, at  $p = 3$ , South Carolina is simply districted with three regions centered at Greenville, Richland, and Charleston counties, which have the three largest populations in South Carolina. At  $p = 4$ , Horry county, which ranks fourth in terms of intraflows, is added as a functional center by annexing some neighboring counties from regions covered by Richland and Charleston. At  $p = 5$  and 6, Spartanburg and Beaufort counties are added as regional centers. It is noteworthy that these three counties show dominant intraflows but small inflows and are grouped with only a few neighboring counties. This finding may indicate that these counties are more self-sustaining regions than regions with functional centers in Greenville, Richland, and Charleston counties. The regional delimitation of South Carolina can be explained using a combination of central location theory and labor-shed regions. Not surprisingly, Richland County, which includes Columbia,



**Figure 10.** Increments of the objective value with  $p$  (South Carolina).

the capital of South Carolina, is a major functional center for administrative services. The most populous counties, Greenville and Charleston, have experienced a rapid population growth in conjunction with a corresponding increase in job opportunities. The most appropriate level of  $p$  for regionalization can be identified from the variation of the objective function values with  $p$ . In Figure 10, the most significant increment in the objective function is clearly observed as  $p$  changes from 3 to 4, indicating that the optimal number of regions is three. If more regions are considered, then five regions are reasonable because of an alternative natural break found in the variation in the objective function values from  $p = 5$  to  $p = 6$ . In terms of an urban hierarchy, the first-tier consists of three centers (Greenville, Richard, and Charleston counties), and the second-tier comprises Horry and Spartanburg counties as subfunctional centers.

## Concluding Remarks

This article considers a flow-based regionalization problem (PFRP) in terms of an MIP formulation that determines  $p$  functional centers and aggregates  $n$  areal units into  $p$  contiguous regions based on spatial interactions. This PFRP model is based on classical location theory, which assumes that a region consists of functional centers that are often referred to as central places, urban centers, or CBDs, depending on geographic scales, surrounded by a hinterland that is linked through spatial interaction with the centers. Spatial interactions, including inflows, outflows, and intra-flows, are key to transforming areal units into  $p$ -functional regions, where the pattern of spatial interactions may represent an area of commuting or similar economic flows. In our results, the PFRP only groups commuting flows among regions with a set of functional centers. However, note the empirical observation that the

functional regions formed by a pattern of spatial interactions among regions, such as commuting flows, are also governed by other political or economic activities.

There are three main differences between the PFRP presented here and other  $p$ -regions models. First, the PFRP fundamentally belongs to a class of location-allocation models with a contiguity requirement. The models in Duque, Church, and Middleton (2011) and Shirabe (2009) are basically allocation models with explicit contiguity requirements, but these models do not identify the significant areal unit within each region. In contrast, the PFRP determines the locations of functional centers by considering the magnitude of spatial interactions among areal units, while also ensuring contiguity. Second, the PFRP is a maximization model for districting regions. Most other  $p$ -regions models focus on minimizing *dissimilarity* as an objective function rather than maximizing *similarity* among spatial units. The PFRP generally maintains compact shapes because the mot flows, such as commuting, are inherently dictated by physical distance, such that spatial contiguity is satisfied in capturing the spatial pattern of commuting flows. Third, in terms of model performance, the PFRP provides more efficient solvability than previous models. The PFRP model can be applied to problems with up to thirty areal units with no additional constraints, which is sufficient to validate many regionalization problems with optimality. Finally, the PFRP combined with ATR can be applied to relatively large problems (e.g., forty-six counties in South Carolina).

From a planning perspective, ATR is useful for situations in which the decision-making process requires that an appropriate level of  $p$  be determined. Although this approach may not guarantee optimality in extremely complex cases, ATR is potentially applicable to any  $p$ -regions problem with criteria involving geographical movements among the regions, which is exemplified here in the cases of Seoul and South Carolina. In particular, ATR is effective in reinforcing optimality for small or moderate size problems if an exact solution is required. Obviously, ATR may have limited applicability for a large number of areal units, which is still the main issue in using a mathematical programming approach to solve regionalization problems. In this regard, future research should be two-pronged. First, the suitability and effectiveness of ATR should be evaluated for use with complex data sets, followed by the development of heuristic techniques for practical applications (Duque, Romas, and Suriñach 2007).

### **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### **Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Research Foundation of Korea Grant, funded by the Korean Government (NRF-2011-327-B00856).

## References

- Alvanides, S., S. Openshaw, and O. Duke-Williams. 2000. "Designing Zoning Systems for Flow Data." In *GIS and Geocomputation*, edited by D. Martin and P. M. Atkinson, 115–34. London, UK: Taylor & Francis.
- Berry, B. J. L., and W. L. Garrison. 1958. "The Functional Bases of the Central Place Hierarchy." *Economic Geography* 34:145–54.
- Bozkaya, B., E. Erkut, D. Haight, and G. Laporte. 2011. "Designing New Electoral Districts for the City of Edmonton." *Interfaces* 41:534–47.
- Bozkaya, B., E. Erkut, and G. Laporte. 2003. "A Tabu Search Heuristic and Adaptive Memory Procedure for Political Districting." *European Journal of Operational Research* 144: 12–26.
- Brown, L. A., and J. H. Holmes. 1971. "The Delimitation of Functional Regions, Nodal Regions and Hierarchies by Functional Distance Approaches." *Journal of Regional Science* 11:57–72.
- Brown, L. A., and F. E. Horton. 1970. "Social Area Change: An Empirical Analysis." *Urban Studies* 7:271–88.
- Christaller, W. 1966. *Central Places in Southern Germany*, Translated by C. W. Baskin. Englewood Cliffs, NJ: Prentice Hall.
- Coombes, M. G., J. S. Dixon, J. B. Goddard, S. Openshaw, and P. J. Taylor. 1979. "Daily Urban Systems in Britain: From Theory to Practice." *Environment and Planning A* 11: 565–74.
- Cova, T. J., and R. L. Church. 2000. "Contiguity Constraints for Single-Region Site Search Problems." *Geographical Analysis* 32:306–29.
- Duque, J. C. 2004. "Design of Homogeneous Territorial Units: A Methodological Pproposal and Applications." PhD dissertation, University of Barcelona, Barcelona, Spain.
- Duque, J. C., L. Anselin, and S. J. Rey. 2012. "The Max-P-Regions Problem." *Journal of Regional Science* 52:397–419.
- Duque, J. C., R. L. Church, and R. S. Middleton. 2011. "The  $p$ -Regions Problem." *Geographical Analysis* 43:104–26.
- Duque, J. C., R. Romas, and J. Suriñach. 2007. "Supervised Regionalization Methods: A Survey." *International Regional Science Review* 30:195–220.
- Fischer, M. M. 1980. "A Comparison of Some Hierarchic and Non-Hierarchic Strategies." *Regional Science and Urban Economics* 10:503–37.
- Garfinkel, R. S., and G. L. Nemhauser. 1970. "Optimal Political Districting by Implicit Enumeration Techniques." *Management Science* 16:B495–508.
- George, J. A., B. W. Lamar, and C. A. Wallace. 1997. "Political District Determination using Large-Scale Network Optimization." *Socio-Economic Planning Sciences* 31:11–28.
- Gordon, A. D. 1999. *Classification*. 2nd ed. London, UK: Chapman & Hall/CRC.
- Green, F. H. W. 1958. "Community of Interest Areas: Notes on the Hierarchy of Central Places and their Hinterlands." *Economic Geography* 34:210–26.
- Hess, S. W., and S. A. Samuels. 1971. "Experiences with a Sales Districting Model: Criteria and Implementation." *Management Science* 18:41–54.

- Hess, S., J. Weaver, H. Siegfeldt, J. Whelan, and P. Zitzlau. 1965. "Nonpartisan Political Redistricting by Computer." *Operations Research* 13:998–1006.
- Kim, C., S. Sang, Y. Chun, and W. Lee. 2012. "Exploring Urban Commuting Imbalance by Jobs and Gender." *Applied Geography* 32:532–45.
- Lösch, A. 1938. "The Nature of Economic Regions." *Southern Economic Journal* 5:71–8.
- Masser, I., and P. J. B. Brown. 1975. "Hierarchical Aggregation Procedures for Interaction Data." *Environment and Planning A* 7:509–23.
- Masser, I., and J. Scheurwater. 1980. "Functional Regionalisation of Spatial Interaction Data: An Evaluation of Some Suggested Strategies." *Environment and Planning A* 12:1357–82.
- Mehrotra, A., E. L. Johnson, and G. L. Nemhauser. 1998. "An Optimization Based Heuristic for Political Districting." *Management Science* 44:1100–14.
- Morrill, R. 1981. *Political Redistricting and Geographic Theory*. New York: Association of American Geographers.
- Murray, A., and V. Estivill-Castro. 1998. "Cluster Discovery Techniques for Exploratory Spatial Data Analysis." *International Journal of Geographical Information Science* 12:431–43.
- Noronha, V. T., and M. F. Goodchild. 1992. "Modeling Interregional Interaction: Implications for Defining Functional Regions." *Annals of the Association of American Geographers* 82:86–102.
- Nystuen, J. D., and M. F. Dacey. 1961. "A Graph Theory Interpretation of Nodal Regions." *Papers and Proceedings of the Regional Science Association* 7:24–42.
- Openshaw, S. 1973. "A Regionalization Program for Large Data Sets." *Computer Applications* 3/4:136–47.
- Openshaw, S. 1977. "A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning, and Spatial Modelling." *Transactions of the Institute of British Geographers* 2:459–72.
- Openshaw, S., and L. Rao. 1995. "Algorithms for Reengineering 1991 Census Geography." *Environment and Planning A* 27:425–46.
- Openshaw, S., and C. Wymer. 1995. "Classifying and Regionalizing Census Data." In *Census Users Handbook*, edited by S. Openshaw, 239–70. Cambridge, UK: GeoInformation International.
- Plane, D. A. 1982. "Redistricting Reformulated: A Maximum Interaction/Minimum Separation Objective." *Socio-economic Planning Sciences* 16:241–44.
- Ravenstein, E. G. 1885. "The Laws of Migration." *Journal of the Statistical Society of London* 48:167–235.
- Rosing, K. E., and C. ReVelle. 1986. "Optimal Clustering." *Environment and Planning A* 18: 1463–76.
- Salazar-Aguilar, M. A., R. Z. Ríos-Mercado, and M. Cabrera-Ríos. 2011. "New Models for Commercial Territory Design." *Networks and Spatial Economics* 11:487–507.
- Schwanen, T., F. M. Deileman, and M. Dijst. 2003. "Car Use in Netherlands Daily Urban Systems: Does Polycentrism Result in Lower Commute Times?" *Urban Geography* 24: 410–30.
- Shirabe, T. 2005. "A Model of Contiguity for Spatial Unit Allocation." *Geographical Analysis* 37:2–16.

- Shirabe, T. 2009. "Districting Modeling with Exact Contiguity Constraints." *Environment and Planning B: Planning and Design* 36:1053–66.
- Song, Y., and H. Kim. 2015. "Evolution of Subway Network Systems, Subway Accessibility and Change of Urban Landscape: A Longitudinal Approach to Seoul Metro-politan Area." *International Journal of Applied Geospatial Research* 6(2):53–76.
- Song, Y., K. Lee, W. P. Anderson, and T. R. Lakshmanan. 2012. "Industrial Agglomeration and Transport Accessibility in Metropolitan Seoul." *Journal of Geographical Systems* 14: 299–318.
- Taaffe, E. J., H. L. Gauthier, and M. E. O'Kelly. 1996. *Geography of Transportation*. 2nd ed. Upper Saddle River, NJ: Prentice Hall.
- Tobler, W. 1987. "An Experiment in Migration Mapping by Computer." *The American Cartographer* 14:155–63.
- van der Laan, L. 1998. "Changing Urban Systems: An Empirical Analysis at Two Spatial Levels." *Regional Studies* 32:235–47.
- Wheeler, J. O., and R. L. Mitchelson. 1989. "Information Flows Among Major Metropolitan Areas in the United States." *Annals of the Association of American Geographers* 79: 523–43.
- White House. 2009. "Update of Statistical Area Definitions and Guidance on Their Use." OMB Bulletin No. 10-02. Accessed December 1, 2009. <http://www.whitehouse.gov/sites/default/files/omb/assets/bulletins/b10-02.pdf>
- Williams, J. C. 1995. "Political Redistricting: A Review." *Papers in Regional Science* 74: 13–40.
- Wise, S. M., R. P. Haining, and J. Ma. 1997. "Regionalisation Tools for Exploratory Spatial Analysis of Health Data in Recent Development." In *Spatial Analysis: Spatial Statistics, Behavioural Modelling, and Computational Intelligence*, edited by M. M. Fischer and A. Getis, 83–100. New York: Springer.
- Yeates, M. 1963. "Hinterland Delimitation: A Distance Minimizing Approach." *The Professional Geographer* 15:7–10.
- Zoltners, A. A., and P. Sinha. 1983. "Sales Territory Alignment: A Review and Model." *Management Science* 29:1237–56.