Computers, Environment and Urban Systems 36 (2012) 456–469

Contents lists available at SciVerse ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

Prescriptive modeling with map algebra for multi-zone allocation with size constraints

Takeshi Shirabe *

School of Architecture and the Built Environment, Royal Institute of Technology (KTH), Drottning Kristinas väg 30, 100 44 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 11 August 2010
Received in revised form 9 December 2011
Accepted 9 December 2011
Available online 4 January 2012

Keywords:
Map algebra
Cartographic modeling
Raster-based geographic information systems
Districting
Land use allocation

ABSTRACT

Map algebra is a methodology for organizing and processing digital cartographic data in a geographic information system (GIS). While its capabilities to describe patterns existing (or hidden) in data have been well studied, its capabilities to prescribe new patterns in response to specific requirements have not been much explored. The latter, prescriptive capabilities help planners address a class of spatial problems called "cartographic allocation" (Tomlin, 1990), which concerns allocation of subsets (or zones) of a cartographic space to certain uses according to one or more criteria. Taking a school districting problem as an example, this paper introduces a systematic approach to designing a map algebraic procedure for a cartographic allocation problem with capacity constraints. It is found that a classical trial-and-error heuristic can be refined to a more formal approximation method and serve as a good alternative to other solution methods when the problem involves a large number of spatial units as is often the case with a raster-based GIS.

1. Introduction

Map algebra (Tomlin, 1994) generally refers to a methodology—and sometimes a language—for organizing and processing data represented in cartographic form. It organizes cartographic data as a set of single-attribute maps called (map) layers and processes them on a layer-by-layer basis, i.e. accepts one or more layers as input and returns a single layer as output. A specific embodiment of organizing and processing cartographic data according to the conventions of this algebra is called a cartographic model.

Map algebra is not just a conceptual framework but a practical tool thanks to the advance of modern computing technology. Nowadays many existing geographic information systems (GISs) adopt map algebra in one form or another and help planners develop a variety of cartographic models. These may be designed to compare vegetation and soil types, measure distance from sources of pollution, characterize land form, summarize the timber volumes of forest stands, and so on.

The models mentioned above were deliberately chosen so as to exemplify *descriptive* models, which relate to *what it is*. *Prescriptive* models are another type, concerning *what should be*—e.g. what arrangement of land use zones should best satisfy given environmental criteria. In contrasting the two kinds of models, Tomlin stated: "To move from description to prescription, however, a new set of techniques is required. These are techniques that broaden the role of cartographic modeling from relatively passive inquiry to much more active intent. Descriptive models answers questions; prescriptive models solve problems" (Tomlin, 1990, p. 198).

"Cartographic allocation" (Tomlin, 1990) is a major problem class that map algebra addresses. It is, in general, to select one or more zones from a given cartographic space for certain land uses according to specific criteria. Some of these criteria are called "atomistic" (Tomlin, 1990) as they are satisfied by each individual location. For example, an atomistic criterion prohibits a planned residential zone from being located on slopes steeper than 5%. On the contrast, "holistic" (Tomlin, 1990) criteria are satisfied by a collection of locations as a whole. An example is such that the residential zone must be smaller than 10 km^2 in size and circular in shape.

An atomistic criterion may be mechanically resolved by a simple screening of locations based on their values stored on a layer of interest. To address multiple atomistic criteria, as has been done manually by landscape architects like McHarg (1969), it is convenient to combine their associated layers into a single layer—often referred to as a *suitability* map. GIS users may conduct this kind of information synthesis, for example, using the weighted addition or linear combination method, which gives each relevant layer a weight according to its relative importance (possibly determined by experts or decision makers) and adds all the weighted layers on a location-by-location basis. GIS-based suitability analysis as such has been extensively studied in the series of work by Eastman and his team (Eastman, Jiang, & Toledano, 1998; Eastman, Jin, Kyem, & Toledano, 1995; Eastman, Kyem, & Toledano, 1993).

---

* Tel.: +46 8 790 6363; fax: +46 8 790 8580.
E-mail address: shirabe@kth.se

0198-9715/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2011.12.003

Once a holistic criterion is introduced, an otherwise trivial allocation problem may suddenly become (at least computationally) intractable. This is more so for those involving multiple zones and multiple holistic criteria. Solution techniques for such problems "rapidly become quite a bit more sophisticated" (Tomlin, 1990, p. 220). Tomlin & Johnston, 1990 demonstrated one such technique, which takes a cyclic process of (1) allocating zones to land uses according to specified criteria, (2) evaluating the quality of the current allocation, and (3) modifying the criteria for better allocation. In the last step of this technique, each criterion is modified by adjusting the importance of the corresponding layer. For example, if a layer did not influence the last allocation as much as it should have, its importance will be raised in the next round. A similar heuristic known as "the Multi-Objective Land Allocation (MOLA) procedure" (Eastman et al., 1993) has been incorporated in a commercial GIS (IDRISI) from the Clark Labs.

Despite its rather simplistic mechanism, the trial-and-error technique has been found of practical use in reconciling multiple, competing interests in limited space. Unfortunately, however, the development of a theory behind this classic technique has been slow. It is true that map algebra has drawn a great deal of interest from the GIS research community over the last few decades, but more focus has been placed on its conventions than on its use as a prescriptive, problem-solving tool. For example, Takeyama and Couclelis (1997) introduced a new data type called meta-relational map and showed its utility in linking map algebra and cellular automata. Cubic map algebra (Mennis, Viger, & Tomlin, 2005) is an extension of map algebra to handle three dimensional data (with one dimension possibly used for time). Frank (2005) discussed a formalization of map algebra for spatio-temporal data. Cordeiro, Câmara, Moura de Freitas, and Almeida (2009) proposed "yet another algebra" for efficient implementation of cellar automata-like dynamic models. Most recently, Li and Hodgson (2004) and French and Li (2010) designed a map algebra for vector data.

While successful refinement of map algebra contribute to our ability to express data-processing procedures, including one for solving a cartographic allocation problem, it is important to note that map algebra itself does not guarantee their correctness. It is not unusual that one's intuition and imagination play a critical role in cartographic modeling but they are not rigorous or specific enough to address technical or computational details like "How should the importance of a layer be raised?", "By adding some value to it? Why not multiplying?", "How big should that value be?" and "May that value be negative?" These are the kinds of questions that need to be answered in order to enhance the accuracy of the solutions generated through (any variant of) map algebra.

Therefore, the purpose of this paper is to explore map algebra's prescriptive, problem-solving capabilities while respecting the original conventions of map algebra. More specifically, it will be shown how a map algebraic heuristic is developed for a multi-zone cartographic allocation problem that seeks a least-cost allocation of two or more zones to specified uses, each of which has a fixed size limit or requirement.

The remainder of the paper is organized as follows. Section 2 reviews the conventions of map algebra. Section 3 introduces a school districting problem that serves as an example of the multi-zone cartographic allocation problem. Section 4 develops a map algebraic procedure for its approximate solution, and extends the procedure to address a more general case. Section 5 reports computational experiments and evaluates the accuracy and efficiency of the map algebraic heuristic. Section 6 concludes the paper.

2. Map algebra

Map algebra is a set of conventions of data, data processing, and data-processing control for analyzing and synthesizing digital cartographic data. The spirit of map algebra is that all these components are decomposed into fundamental units so that (re)compositions of such units generates a wide range of flexibility and applicability.

Cartographic data (i.e. recorded facts in map form) are decomposed to several layers. Unlike conventional maps, map algebraic layers are represented by numeric values rather than by colors or graphic symbols. This makes it possible to design unambiguous mathematical operations on layers.

Data-processing methods are decomposed to a set of basic operations. Their counterparts in elementary algebra include addition, subtraction, division, and multiplication. Tomlin (1990) classifies map algebraic operations into three categories: local, focal, and zonal. Local operations assign each *location* a new value as a function of the location's value(s). A *location* occupies the smallest portion of a layer that is uniquely identified by an ordered pair of planar coordinates, and the same set of locations are shared by all layers in a single model. Focal operations assign each location a new value as a function of the values of all locations within the *neighborhood* to which that location belongs. A neighborhood of a location is usually specified by distance and/or directional relationship to that location. Zonal operations assign each location a new value as a function of the values of all locations within the *zone* to which that location belongs. A zone is defined by a set of locations having the same value on a specific layer. While each of these operations may be useful on its own, one can combine two or more of them into a procedure that performs a more complicated task.

Data-processing controls, i.e. means to specify and arrange data-processing methods, are decomposed to strings of symbols resembling those of elementary algebra and English. Strings of symbols can be combined to express a statement which represents an operation. A set of statements can be further combined to describe a program that represents a procedure. As an example, the following map algebraic program represents a map algebraic procedure for averaging four different cost layers on a location-by-location basis (excerpted from Tomlin (1990), p. 60).

TotalCost = LocalSum of YourCost and MyCost
TotalCost = LocalSum of TotalCost and HerCost
TotalCost = LocalSum of TotalCost and HisCost
AverageCost = LocalRatio of TotalCost and 4

3. Districting as multi-zone cartographic allocation

This section formally describes the multi-zone cartographic allocation problem we investigate in this paper. The problem takes the form of a classic planning problem, the districting (or regionalization) problem. It in general involves partitioning of a given area into a specified number of districts, and is often cast in terms of grouping of discrete areal units into larger clusters. In the context of map algebra, such units are normally grid cells, because raster-based GIS are, at least for now, the most common platform for cartographic modeling with map algebra.

3.1. Problem statement

We begin with a simplest districting problem we could pose to raster-based GIS. It aims to allocate each cell to one and only one zone, which may not exceed a specified size limit, such that the sum of the costs of all zones is minimized. The size of a zone is defined as the sum of the values (of a chosen attribute) associated with all the cells comprising that zone. The cost of a zone is defined in the same manner. For ease of illustration, the problem is considered in the school districting context taken from Yeates (1963) where student population is regarded as size attribute and travel cost as cost attribute.

> Given a study area comprising *n* grid cells, each having a fixed number of students, and containing m schools, each having a fixed upper limit (or capacity) on its enrollments, partition the study area into *m* school zones subject to the following constraints. The total travel time spent by all students must be minimized, no school can accommodate more students than its capacity, and each grid cell must belong to one and only one school district.

It should be noted that the purpose of studying this particular problem is to learn and demonstrate an important theory applicable to similarly formulated problems that take place in raster space. We acknowledge that there are many cases in which irregularly-shaped polygons (rather than grid cells) such as ZIP code areas, census tracts, and city blocks are preferred spatial units, and also, in addition to capacity and travel time (or distance), other criteria such as contiguity (Caro, Shirabe, Guignard, & Weintraub, 2004), racial balance (Belford & Ratliff, 1972; Franklin & Koenigsberg, 1973; Diamond & Wright, 1987; Caro et al., 2004; Church & Murray, 1993; November, Cromley, & Cromley, 1996; Schoepfle & Church, 1991), socio-economic distribution (desJardins et al., 2007), student-teacher ratios (November et al., 1996), changes in school over time (Caro et al., 2004; Holloway, Wehrung, & Zeitlin, 1975; Lemberga & Church, 2000), school closing (Diamond & Wright, 1987; Church & Murray, 1993), and population change (Armstrong, Lolonis, & Honey, 1993; Ploughman, Darnton, & Heuser, 1968) may be equally important in practice.

3.2. Problem formulation

The raster school districting problem stated above can be formulated as a set of equations involving 0–1 decision variables commonly referred to as an integer programming model.

$$\min z = \sum_{i \in I} \sum_{j \in J} d_{ij} a_i x_{ij} \tag{1}$$

subject to

$$\sum_{i \in I} a_i x_{ij} \leqslant b_j \quad \forall j \in J \tag{2}$$

$$\sum_{j \in J} x_{ij} = 1 \quad \forall i \in I \tag{3}$$

$$x_{ij} \in (0, 1) \quad \forall i \in I \quad \forall j \in J \tag{4}$$

where I is a set of cells, J is a set of schools, d_{ij} is a constant indicating the travel time from cell i to school j, a_i is a constant indicating the number of students residing in cell i, b_j is a constant indicating the capacity of school j, x_{ij} is a decision variable that equals to 1 if cell i is allocated to school (zone) j, 0 otherwise.

Eq. (1) defines the objective of the problem, that is, to minimize the total commuting time of all the students. Eq. (2) imposes an upper bound on the number of students enrolled in each school. Eq. (3) ensures that all students in each cell are assigned to schools. Eq. (4) prevents fractional assignments of cells to multiple schools.

While there are a variety of solution techniques (e.g. branch-and-bound algorithms) for the problem as formulated above, exploration of these techniques is not the purpose here. Still it is worth noting that the problem belongs to a computationally difficult problem class commonly known as NP-hard, which implies that it becomes exceedingly more intractable as the numbers of grid cells and of zones increase, which, in turn, justify use of a heuristic.

3.3. Related problems

The districting problem appears in a wider range of contexts than considered in the present paper. Included are political districting (e.g. Hess et al., 1965, Garfinkel & Nemhauser, 1970, Forman & Yue, 2003; Gutiérrez-Ándrade & García, 2009; Mehrotra, Johnson, & Nemhauser, 1998; Williams, 1995), sales territory design (e.g. Drexl & Haase, 1999; Fleischmann & Paraschis, 1988; Hess & Samuels, 1971; Marlin, 1981; Salazar-Aguilar, Ríos-Mercado, & Cabrera-Ríos, 2011; Segal & Weinberger, 1977; Shanker, Turner, & Zoltners, 1975; Vidale, 1956; Zoltners & Sinha, 1983, 2005), and land use allocation (e.g. Aerts, Eisinger, Heuvelink, & Stewart, 2003; Aerts & Heuvelink, 2002; Crema, 1996; Eastman et al., 1993; Gilbert, Holmes, & Rosenthal, 1985; Wright, Revelle, & Cohon, 1983). Also, reserve design problems (e.g. Cerdeira, Gaston, & Pinto, 2005; Church, Gerrard, Gilpin, & Sine, 2003; Mcdonnell, Possingham, Ball, & Cousins, 2002; Nalle, Arthur, & Sessions, 2003; Önal & Briers, 2005; Önal & Wang, 2008) may be seen as a special case in which a single district is sought.

The school districting problems formulated as Eqs. (1)–(4) has a feature that makes it distinct from other districting problems: schools are designated as district centers. If no such locations are given, then the problem becomes one that simultaneously locates district centers and allocates spatial units to districts so as to minimize the sum of the population-weighted travel cost between each spatial unit and its district center, without violating the capacity of any district. This is another NP-hard location–allocation problem, known as the capacitated *p*-median problem (CPMP). The formulation of a districting problem as CPMP dates at least back to Hess, Weaver, Siegfeldt, Whelan, and Zitlau (1965). They developed an iterative solution procedure called "REDIST" and later evolved to "GEOLINE" (Hess & Samuels, 1971), which starts with promising district centers, allocates spatial units to them by solving a transportation problem, and relocates the temporary center of each district to a 1-median of that district. More sophisticated heuristic methods for CPMP have been proposed by, among others, Lorena and Senne (1994, 2003), Maniezzo, Mingozzi, and Baldacci (1998), Ceselli and Righini (2005), Diaz and Fernandez (2006), and Santos-Correa, Steiner, Freitas, and Carnieri (2004).

The use of a computer to solve the districting problem has been extensively investigated. One of the earliest attempts was, as described above, based on the transportation model, and has been employed by many researcher including Hess et al. (1965), Segal and Weinberger (1977), Marlin (1981), and Fleischmann and Paraschis (1988). A multi-objective mixed integer programming model was developed by Diamond and Wright (1987) and later revised by Church and Murray (1993) to simultaneously address school closing and consolidation problems with a primary focus on the resulting school capacity utilization after school closing and consolidation. Schoepfle and Church (1991) transformed a school districting problem with racial balance considerations into a network flow problem, which can be solved by a commercial solver. November et al. (1996) used goal programming models to minimize disparities in minority enrollments, dollars spent per student, student–teacher ratios, and overall enrollment as well as total transportation cost. Mehrotra et al. (1998) developed a heuristic, based on branch-and-price and column generation techniques, to solve a political districting problem for the state of South Carolina.

In addition to those deterministic optimization algorithms and heuristics, stochastic approaches have recently gained increasing attention. For example, Macmillan (2001), Aerts and Heuvelink (2002), Gutiérrez-Ándrade and García (2009) implemented a simulated annealing method for a generic districting problem, for a land use allocation problem, and for an electoral districting problem, respectively. As an alternative to simulated annealing, Bação, Lobo, and Painho (2005) proposed genetic algorithms in the context of political redistricting. Also, Forman and Yue (2003) modified a genetic algorithm originally designed to solve the traveling salesman problem for a congressional districting problem. Another unique heuristic called "geographical distillation" was proposed by Middleton (2006). It addresses not only districting problems but a variety of other spatial optimization problems including the p-median problem and the traveling salesman problem. It does so by first identifying structures that comprise locally good solutions and then combining them into globally good solutions. For more thorough reviews of districting models and algorithms, see, e.g., Kalcsics, Nickel, and Schröder (2005) and Duque, Ramos, and Surinach (2007).

4. Map algebraic heuristic

This section presents a map algebraic heuristic for the school districting problem formulated above. For facilitate our discussion, consider that two raster layers are initially given: one representing the location of the 17 schools (Fig. 1) whose capacities are recorded in Table 1 and the other on which each cell has a value indicating the number of students residing in it.

Traditional but experienced planners may manage to draw reasonable school zone boundaries by visual inspection of these maps, perhaps assuming that travel time is proportional to straight-line distance. Such an approach is usually regarded as a rough-and-ready art, but can be refined to a systematic approximation heuristic like a "graphical solution to the transportation problem" by Vidale (1956), which is reviewed below.

4.1. Vidale's graphical solution

Vidale's problem is restated here: "Given sources having fixed production capacities of certain goods and consuming centers having fixed requirements of the goods, which sources serve which centers to minimize the total transportation costs?" This is known as the simple transportation problem or Hitchcock (1941) transportation problem and takes a similar formulation as Problem (1)–(4) except that x_{ij} is not required to be integral. Therefore, Vidale's problem is formulated as Eqs. (1)–(3) and

$$x_{ij} \geqslant 0 \quad \forall i \in I, \forall j \in J \tag{5}$$

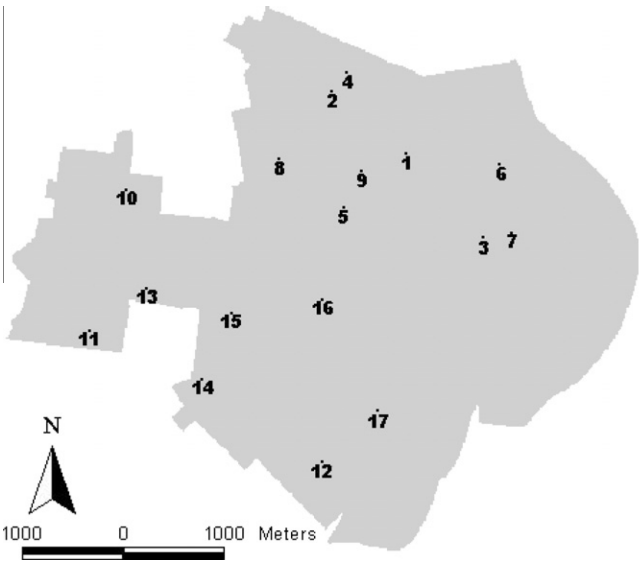

Fig. 1. Location of schools (annotated with their ID's) in a study area (shaded). Note that this layer is referenced to a cartographic frame whose resolution (or cell size) is 15.24 m by 15.24 m, and so are all other layers that follow.

Note that I and J are here regarded as a set of consuming centers and a set of production sources, respectively.

To solve the problem, Vidale began with an idea that if every source has an unlimited capacity, an optimal solution would be easily obtained by having all consuming centers served by their least-cost sources. The resulting coverage of each source is called its natural market. In actuality, however, while some sources cannot serve all consuming centers in their own natural markets, others have to serve more consuming centers beyond their own natural markets. Vidale resolved this unbalance by reasoning that each source implicitly has its own "relative location advantage" (Vidale, 1956)—which is by definition nonnegative—and that if this hidden value is added to the shipping cost of each unit of the goods, the consequent natural markets are optimal markets.

By constructing the dual of Problem (1)–(3) and (5), one can prove the correctness of Vidale's proposition. The dual problem is formulated below.

$$\max \quad \sum_{i \in I} v_i - \sum_{j \in J} b_j u_j \tag{6}$$

subject to

$$d_{ij}a_i + a_i u_j - v_i \geqslant 0 \quad \forall i \in I, \forall j \in J \tag{7}$$

$$u_j \geqslant 0 \quad \forall j \in J \tag{8}$$

$$v_i \geqslant 0 \quad \forall i \in I \tag{9}$$

where u_j and v_i are decision variables.

According to the dual complementary slackness theorem (see, e.g. Rardin, 1997), the optimality conditions for the two problems are:

$$(d_{ij}a_i + a_i u_j - v_i)x_{ij} = 0 \quad \forall i \in I, \forall j \in J \tag{10}$$

This implies that if $x_{ij} > 0$, then $d_{ij}a_i + a_iu_j - v_i = 0$ and if $d_{ij}a_i + a_iu_j - v_i > 0$, then $x_{ij} = 0$.

In other words, assuming $a_i > 0$, if center i is served by source j, then $d_{ij} + u_j = v_i/a_i$ and if $d_{ij} + u_j > v_i/a_i$, then center i is not served by source j. It follows that in an optimal market division, each center i is served by a source j that has the smallest $d_{ij} + u_j$ (which equals to v_i/a_i). By regarding u_j as the relative location advantage of source j, Vidale's proposition has been proved.

Now suppose that center i lies on the boundary between sources j and k in an optimal solution. According to Vidale's proposition, the following relation holds.

$$d_{ij} + u_j = v_i/a_i = d_{ik} + u_k \tag{11}$$

Equivalently

$$d_{ij} - d_{ik} = u_k - u_j \tag{12}$$

This implies that the boundary between two sources is a line of constant cost differential between them.

Suppose further that center i lies on the boundary between sources j and l (not k) in the same optimal solution. Similarly, the following must be true.

$$d_{ij} + u_j = d_{il} + u_l \tag{13}$$

From Eqs. (11) and (13), we have

$$d_{ik} + u_k = d_{il} + u_l \tag{14}$$

This implies that center i is also on the boundary between sources k and l. The same property is found in Thiessen polygons (see Okabe, Boots, Sugihara, & Chiu, 2000).

These results lead to Vidale's other proposition that optimal markets are defined by boundaries satisfying the following conditions:

Table 1
Capacities of schools.

School ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Capacity	504	450	440	400	564	533	270	755	474	895	900	601	665	720	960	782	350

1. The boundary between two sources comprises points the difference of whose transportation costs from the sources is a constant.
2. At a point where two boundaries meet, at least one other boundary must meet.

Vidale designed a procedure for finding boundaries that satisfy the two conditions. It superimposes a layer representing lines of constant transportation cost difference for each pair of sources on top of a layer representing the location of consuming centers (see Fig. 2), and repeats the following. First, guess a correct topological configuration of optimal markets. Then, select a cost difference line for a pair of adjacent sources from the corresponding cost difference layer and perform this routine sequentially for all other pairs. Finally, if the capacity constraint is violated by any one of the resulting markets, select another cost difference line in the second step. If no set of boundaries are found satisfactory, start with a different topological configuration. Vidale stated that a problem with five to seven sources and over a thousand consuming centers could be solved in a few hours (Garrison, 1959).

Vidale's graphical solution seems applicable to our school districting problem as schools can be regarded as production sources, (centers of) cells as consuming centers, and travel time as transportation cost. However, the manual technique will not be used here for at least three reasons. First, an exceedingly large number of layers may need to be inspected. Vidale stated that although many could be counted out beforehand (since only a limited number of topological arrangements apply to a given set of sources), theoretically m sources generate $m(m-1)/2$ cost difference layers. Second, the topological configuration of optimal school zones must be guessed in advance. This is a difficult task when a large number of schools are involved. Third, it is assumed that transportation cost increases monotonically with distance, which severely limits the applicability of the method. For example, if different degrees of mobility are assumed on and off streets, travel time will no longer be linear to straight-line distance (Upchurch, Kuby, Zoldak, & Barranda, 2004). Moreover, in a more general cartographic allocation problem, allocation costs may be distributed over the study area with no apparent regularity (See Section 5). This will make it difficult (if not impossible) to create cost difference layers manually, and even if it were possible, one might need to inspect all $m(m-1)/2$ cost difference layers since one cannot tell a priori which pairs of zones are likely to be adjacent. However, these difficulties are ascribed mostly to human visual capacity and, as shown next, can be easily resolved by GIS and map algebra.

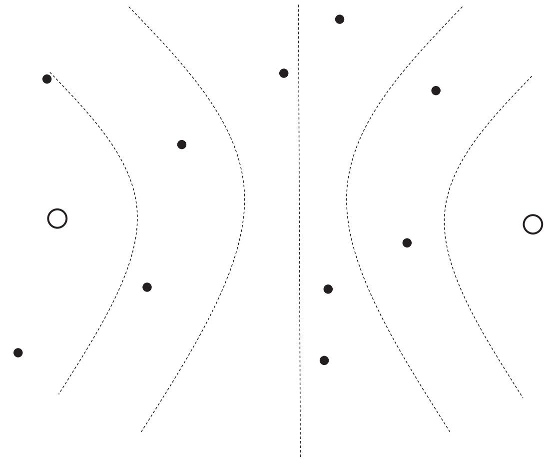

Fig. 2. Location of consuming centers (black dots) and lines of constant transportation cost difference (dashed lines) for two production centers (circles).

4.2. Cartographic modeling solution

Although we have decided not to take Vidale's method as it is, we have learned an important implication from it. That is, if we know the *true* cost (i.e. travel time d_{ij} plus relative location advantage u_j) of assigning each cell to each school in advance, we can solve the problem simply by creating natural markets (school zones) in terms of this cost. Here remember that our school districting problem requires all students in a cell be assigned to the same school (see Eq. (4)).

The travel time portion of the true cost of assigning cell i to school j is the product of the travel time d_{ij} between them and the number of students a_i residing in cell i. We refer to the resulting value $d_{ij}a_i$ as cell-based travel time. A layer representing cell-based travel time to a school can be easily generated by map algebra's *LocalProduct* of the travel time layer for that school and the student population layer.

Similarly, since a school's relative location advantage is imposed uniformly on each student (cf. a source's relative location advantage in Vidale's problem, which is imposed on each unit of goods), the relative location advantage portion of the true cost of assigning cell i to school j is the product of the relative location advantage u_j of school j and the number of students a_i in cell i. We refer to the resulting value u_ja_i as cell-based relative location advantage. Again, this calculation can be performed by *LocalProduct* of the student population layer and a layer on which all cells have the same value equal to the relative location advantage of a school.

From a map algebraic point of view, Vidale's relative location advantage can be seen as a weight given to the population layer. This observation may not seem to bring any benefit to the present problem, but it does to a problem involving two or more different kinds of sizes (not just population size) because all the corresponding layers can be handled mechanically. For example, if a school offers multiple grades each of which has a fixed capacity, its true cost layer is created by weighting each graders' population layer and adding the resulting layer to the cell-based travel time layer (see Fig. 3).

The mechanism of weighted addition described above is easy to understand, but not so to perform because weights are not given but to be found. In the school districting problem, it is intuitively true that a school that receives more students than its capacity allows should give the student population layer a higher weight in order to make attendance more costly, and a school whose capacity is not filled can do otherwise. No such judgment, however, can be made before all school zones are delineated. This is a typical chicken-and-egg dilemma. Nevertheless, this intuition suggests the following heuristic, which iteratively adjusts weights until all capacity violations are resolved.

Step 1: Construct zones by assigning each cell to its least cost school with respect to the current true cost. Initially, every school gives the student population layer a weight of 0 so that the true cost layers is identical to the cell-based travel time layer.
Step 2: For each zone, calculate the number of students.

Step 3: For each zone, calculate the population deviation, which is the number of students within that zone minus the capacity of the corresponding school. Stop if no zone has a positive deviation.
Step 4: For each zone, update the weight for the student population layer by adding the product of that zone's population deviation and a certain step size. If the resulting weight becomes smaller than 0, set it equal to 0.
Step 5: For each zone, update the weighted student population layer by multiplying the student population layer by the updated weight.
Step 6: For each zone, update the true cost layer by adding the cell-based travel time layer and weighted student population layer. Return to Step 1.

As shown below, the procedure can be easily translated into map algebraic terms (see also Fig. 4). Details of the map algebraic operations it refers to are found in Tomlin (1990). The map algebraic procedure takes the following layers as input (assuming that each school is indexed with a unique integer ranging from 1 to n).

- *Population* layer on which each cell is assigned a value indicating the number of students in that cell.
- *StudyArea* layer on which all cells are assigned the same arbitrary value.
- *Capacity*(i) layer on which each cell is assigned the same value indicating the capacity of school i.
- *StepSize*(i) layer on which all cells are assigned the same value indicating the step size chosen for School i.
- *TravelTime*(i) layer on which each cell is assigned a value indicating the sum of the time it takes all students in that cell to travel to school i.

and returns a layer called *EachZone* on which each cell is assigned a value indicating the zone to which that cell belongs.

Map algebraic heuristic for school districting with capacity constraints

```
// Initialization
for $i$ = 1 to $n$
  TrueCost($i$) = LocalRating of TravelTime($i$)
  Weight($i$) = LocalRating of 0
endfor
do
  // Step 1
  MinCost = LocalRating of $\infty$
  for $i$ = 1 to $n$
    MinCost = LocalMinimum of MinCost and TrueCost($i$)
  endfor
  EachZone = LocalRating of MinCost
  for $i$ = 1 to $n$
    EachZone = LocalRating of EachZone with i for TrueCost($i$)
  endfor
  // Step 2
  EachPop = ZonalSum of Population within EachZone
  for $i$ = 1 to $n$
    Pop($i$) = LocalRating of EachZone with EachPop for i with $-\infty$ for ... $(i - 1)$ with $-\infty$ for $(i + 1)$ ...
    Pop($i$) = ZonalMaximum of Pop($i$) within StudyArea
  endfor
  // Step 3
  for $i$ = 1 to $n$
    Deviation($i$) = LocalDifference of Pop($i$) and Capacity($i$)
  endfor
```

* (*continued*)

Map algebraic heuristic for school districting with capacity constraints

```
  // Step 4
  for $i$ = 1 to $n$
    WeightChange($i$) = LocalProduct of Deviation($i$) and StepSize($i$)
    Weight($i$) = LocalSum of Weight($i$) and WeightChange($i$)
    Weight($i$) = LocalMaximum of Weight($i$) and 0
  endfor
  // Step 5
  for $i$ = 1 to $n$
    WeightedPop($i$) = LocalProduct of Population and Weight($i$)
  endfor
  // Step 6
  for $i$ = 1 to $n$
    TrueCost($i$) = LocalSum of TotalTime($i$) and WeightedPop($i$)
  endfor
until all Deviation($i$)'s become a layer on which no cell has a positive value.
```

It should be noted that the procedure presented above is expected but not guaranteed to converge to a single solution. The chance (and speed) of convergence depends on the step size chosen in Step 4. From our experience, the step size for each j, here denoted by s_j, may be initially set to:

$$s_j = \frac{\sum_{i \in I} d_{ij} a_i}{\delta \cdot b_j \sum_{i \in I} a_i} \tag{15}$$

where d_{ij}, a_i, and b_j are constants in Eqs. (1) and (2) and δ is a constant ($\approx$100) chosen to make s_j sufficiently small. If the step size is found too small or too large after several iterations, then try the following modification before starting over: if intermediate solutions are almost identical, choose a smaller δ; if they diverge or alternate between two or more solutions, choose a larger δ. An alternative rule of thumb is a mathematical technique called the subgradient method for Lagrangian relaxation (Fisher, 1981, 1985), which is briefly explained in Appendix.

One additional trick we have found effective to improve the chance of convergence is to increase/decrease each of the weights updated in Step 4 by a small rate (between −1% and 1%) randomly chosen. This seemingly redundant routine helps avoid getting stuck in an infinite loop of infeasible solutions.

The proposed heuristic can be implemented in a typical office computer. To show this, it was programmed in Java and tested on an Intel Core i5 2.35 GHz CPU with 3.42 GB of RAM. Although no feasible solution was found after several hours of computation, it took only 22 s (or 665 iterations) to encounter a practically satisfactory solution (Fig. 5) in that no school violates its capacity by more than 1% (Table 2). The table also shows that some schools could accommodate more students, that is, they are currently not subject to the capacity constraint. This implies that those schools have natural zones, which, in turn, explains their zero weighting of the student population layer. On the other hand, the schools that used up their capacity had to give the student population layer a positive weight. Finally it is worth noting that the solution illustrated in Fig. 5 contains a fragmented school zone as the map algebraic heuristic has taken into account only capacity constraints. Therefore, the present example should be considered illustrative rather than practical.

4.3. Extension

So far it has been assumed that each school has a finite capacity; that is, the population size of each zone is bounded from above. A natural extension of it addresses lower bound constraints.

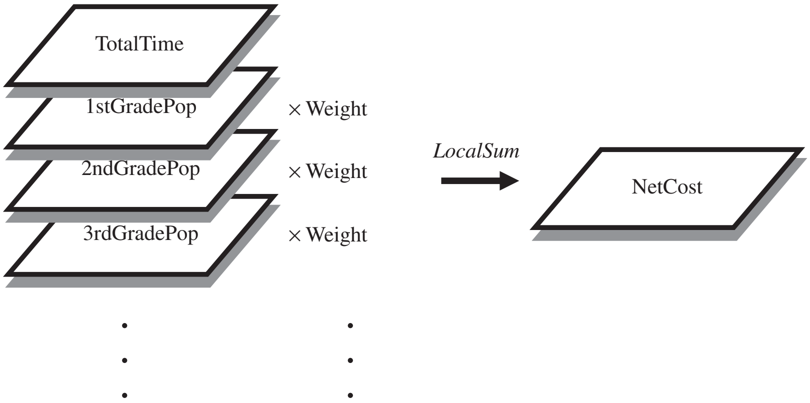

Fig. 3. "True cost" layer as *LocalSum* of "Total Time" layer and multiple weighted population layers.

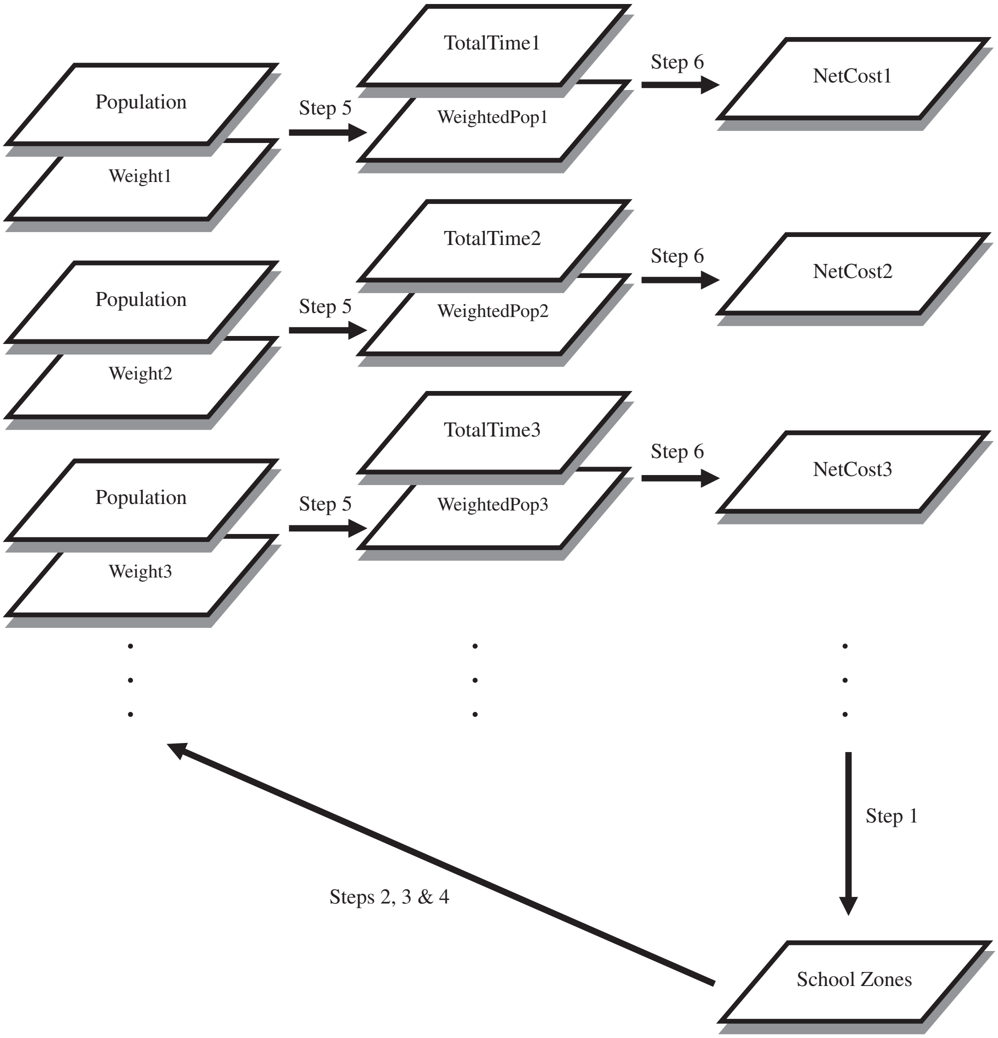

Fig. 4. A schematic diagram of a cartographic model for school districting.

Suppose that each school requires a minimum number of students in order to remain in operation. This is a lower bound constraint, and formally expressed as:

$$\sum_{i \in I} a_i x_{ij} \geqslant l_j \quad \forall j \in J \tag{16}$$

where l_j is the lower bound on the enrollment of school j.

Equivalently,

$$\sum_{i \in I} (-a_i) x_{ij} \leqslant -l_j \quad \forall j \in J \tag{17}$$

The latter form is identical to the upper-bound constraint expressed by Eq. (2) and thus can be interpreted such that the size of each zone with respect to the *negative* of population may not

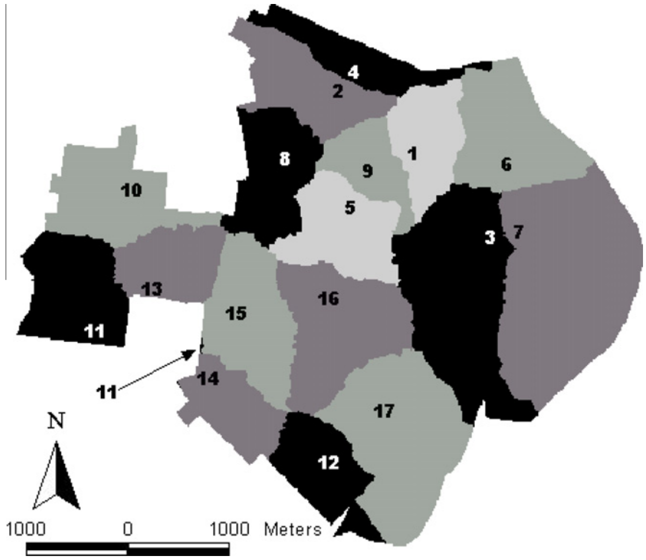

Fig. 5. An approximate solution to Problem (1)–(4).

exceed the *negative* of a lower bound. Accordingly, the procedure remains valid for a problem with lower-bound constraints, if Step 3 is modified in a way that replaces the student population layer with its inverse and the upper bound with the negative of the lower bound. This leads to the following variation of Step 3, followed by Step 4 (unchanged).

> Step 3: For each zone, calculate the population deviation, which is the *negative* of the number of students within the zone minus the *negative* of the lower bound of the corresponding school. Stop if no zone has a positive deviation.
> Step 4:For each zone, update the weight of the student population by adding the product of that zone's population deviation and a certain step size. If the resulting weight becomes smaller than 0, set it equal to 0.

These steps can be mechanically simplified to:

> Step 3: For each zone, calculate the population deviation, which is the number of students within the zone minus the lower bound of the corresponding school. Stop if no zone has a *negative* deviation.
> Step 4:For each zone, update the weight of the student population layer by adding the product of that zone's population deviation and a certain step size. If the resulting weight becomes *larger* than 0, set it equal to 0.

The modified steps can be intuitively understood as follows. Schools that receive less than a sufficient number of students should decrease the weight of the student population layer in order to make attendance less costly and that schools that receive more than a sufficient number of students can do the opposite. What may not be intuitive is that the resulting weight is nonpositive.

If a school has both a capacity and a minimum enrollment requirement, it will need to give the student population layer two different weights—a nonnegative weight for the former constraint and a nonpositive weight for the latter constraint—and process them independently. Upon the completion of the procedure, one may elect to add the two weights and designate the result as a single weight, which is obviously not restricted in sign.

With the means to handle both lower and upper bound constraints, we are in position to depart the school districting scenario and consider a more general cartographic allocation problem. Suppose that zones need to be allocated to certain uses within a study area, and a suitability analysis has generated a suitability layer for each use. In addition, each use requires its zones to be of a certain size in terms of each relevant attribute represented by a layer. The school districting procedure is easy to be applied to this general allocation problem. To see this, we only need to regard the suitability layer for each use as the cell-based travel time layer for each school and a layer in terms of which a size of each zone is measured as the population layer, and then choose an appropriate weight sign for each size-constraining layer. See Fig. 6 for a schematic diagram.

5. Computational experiments

To evaluate the accuracy and complexity of the proposed heuristic, we conducted computational experiments with sample districting problems populated with hypothetical data. All computation was performed on an Intel Core i5 2.35 GHz CPU with 3.42 GB of RAM.

5.1. Accuracy

For accuracy testing, we created five 100-by-100 random raster layers on which each cell was assigned an integer ranging from 1 to 10 with an equal probability. These are referred to as R10 layers. Similarly, we created five R100 layers, five R1000 layers, and five R10000 layers. Then we smoothed all those random layers by applying the *FocalMean* operation with a 3-by-3 neighborhood three times successively. The resulting layers are autocorrelated as is the case with many geographic phenomena, and referred to as A10, A100, A1000, and A10000 layers.

A problem that we call R10&100 seeks a minimum-cost division of the study area (i.e. a 100-by-100 raster space) into five zones in a way that:

- a R10 layer serves as the layer in terms of which the size of each zone is measured,
- each R100 layer serves as the cost layer for a zone,
- the size of zone 1 is approximately 50% of the size of the study area, and those of zones 2, 3, 4, and 5 are approximately 25%, 12.5%, 6.25%, and 6.25% of it, respectively. "Approximately" here means that the R10 size of each zone may not deviate from its specified target size by more than 10%, and
- the sum of the costs of all zones must be minimized without any of the above size constraints being violated.

Table 2
Enrollments and weights of schools in a heuristic solution. Note: all values have been rounded.

School ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Capacity	504	450	440	400	564	533	270	755	474	895	900	601	665	720	960	782	350
Resulting enrollment	496	392	442	349	569	535	228	746	418	903	727	607	669	716	969	779	351
Weight	122	0	0.17	0	77.9	76.1	0	226	0	175	0	1298	213	880	471	520	978

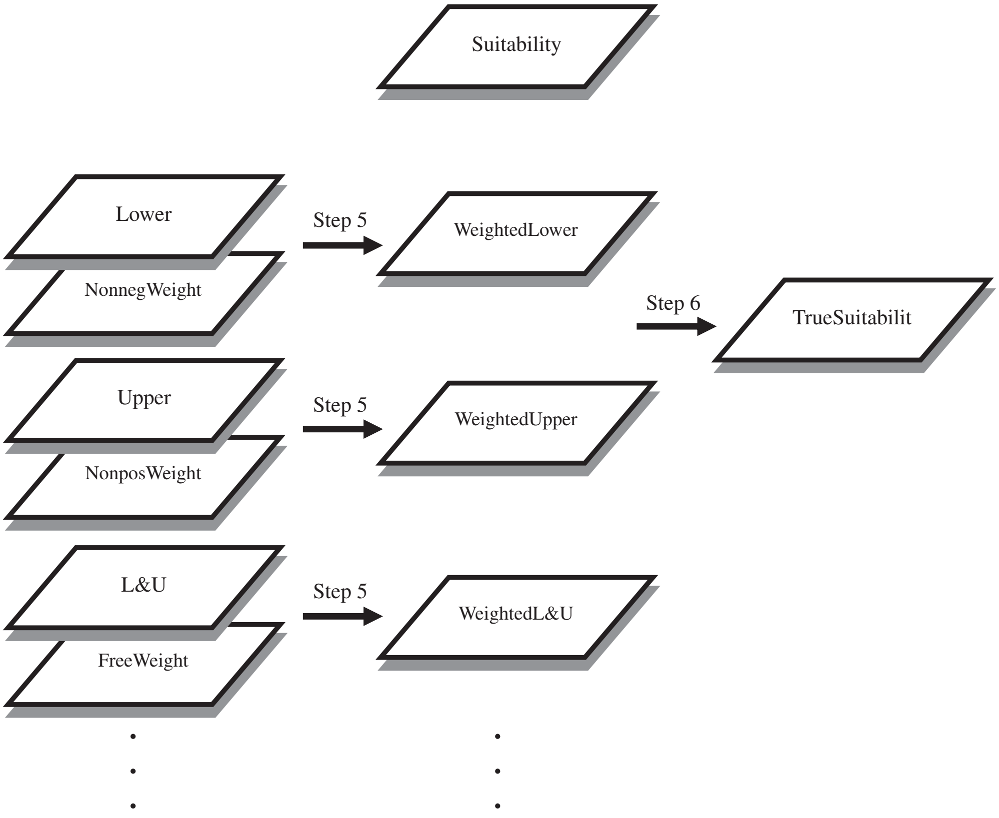

Fig. 6. Layer weighting routine. The layer annotated with "Lower" is a layer in terms of which the size of a zone is measured and bounded from below, and is accompanied by a layer annotated with "NonnegWeight" representing its nonnegative weight. The layer annotated with "Upper" is a layer in terms of which the size of the zone is measured and bounded from above, and is accompanied by a layer annotated with "NonposWeight" representing its nonpositive weight. The layer annotated with "L&U" is a layer in terms of which the size of a zone is measured and bounded both from below and from above, and is accompanied by a layer annotated with "FreeWeight" representing its "free" (i.e., not restricted in sign) weight. All of them are transformed through LocalProduct operation into their respective weighted layers (i.e. "WeightedLower," "WeightedUpper," or "WeightedL&U.") and then added to the "Suitability" layer with respect to which the size of the zone is maximized. The resulting layer indicates the "true suitability" of each cell for the land use under consideration.

Table 3
(a) Heuristic and exact solutions to problems with 100-by-100 random layers. (b) Heuristic and exact solutions to problems with 100-by-100 autocorrelated layers. Note that: # of iteration is the number of iterations performed by the heuristic, time is the running time (in seconds) of the heuristic, Obj_Heu is the objective value of the heuristic solution, Obj_Exa is the objective value of the exact solution (if the problem is optimally solved) or a lower bound on it (otherwise), and gap is the relative optimality gap between the heuristic solution and the exact or best possible solution computed as (Obj_Heu–Obj_Exa)/Obj_Exa.

Problem	# of Iterations	Time (s)	Obj_Heu	Obj_Exa	Gap
(a)					
R10&100	1997	1.984	212598	211976	0.002934294
R10&1000	2989	2.828	2058058	2052271	0.002819803
R10&10000	3762	3.609	20679972	20499321	0.008812536
R100&10	4931	4.703	25832	25572	0.010167371
R100&1000	599	0.64	2034599	2027470	0.003516205
R100&10000	327	0.375	20518944	20431557	0.00427706
R1000&10	972	1.016	25635	25472	0.006399183
R1000&100	336	0.359	208421	207798	0.002998104
R1000&10000	199	0.25	20444748	20286944	0.007778587
R10000&10	3440	3.344	25803	25548	0.009981212
R10000&100	609	0.656	210792	209051	0.008328111
R10000&1000	179	0.235	2052743	2035418	0.008512
(b)					
A10&100	112	0.157	455278	454297	0.002159586
A10&1000	824	0.765	4433180	4431910	0.000286558
A10&10000	907	0.813	44475788	44450610	0.000566426
A100&10	242	0.266	53898	53785	0.002100958
A100&1000	295	0.297	4436336	4434288	0.000461855
A100&10000	264	0.266	44527012	44510155	0.000378723
A1000&10	95912	81.438	53883	53776	0.001989735
A1000&100	37	0.079	454214	454018	0.000431701
A1000&10000	52	0.078	44440568	44425589	0.000337174
A10000&10	102311	86.703	53801	53774	0.000502101
A10000&100	386	0.437	454343	454274	0.000151891
A10000&1000	72	0.11	4432151	4430741	0.000318231

Table 4
Heuristic solutions to problems with 100-by-100, 500-by-500, and 1000-by-1000 random layers. Note that: layer size is the product of the number of rows and the number of column of the input raster, δ is a constant used in Eq. (15), # of iteration is the number of iterations performed by the heuristic, time is the running time (in seconds) of the heuristic, Obj_Heu is the objective value of the heuristic solution, L_Bound is a lower bound (obtained by the heuristic) on the objective value of the exact solution, and gap is the relative optimality gap between the heuristic solution and the best possible solution computed as (Obj_Heu–L_Bound)/L_Bound.

Problem	Layer size	δ	# of Iterations	Time (s)	Obj_Heu	L_Bound	Gap
R10&100	100 by 100	100	1997	1.984	212598	211745	0.004030505
		500	1658	1.61	216494	211264	0.024755225
		1000	5486	5.187	213085	211495	0.007518727
	500 by 500	100	158	3.953	5296885	5241703	0.010527495
		500	753	18.344	5275466	5242401	0.00630732
		1000	5872	139.907	5301704	5241029	0.011576925
	1000 by 1000	100	245	24.156	21125652	20977584	0.007058391
		500	1302	124.953	21215896	20975156	0.011477388
		1000	5748	560.86	21046876	20979424	0.00321515
R10&1000	100 by 100	100	2989	2.828	2058058	2051982	0.002961284
		500	3027	2.859	2063292	2047816	0.007557258
		1000	12324	11.468	2064277	2051705	0.006127648
	500 by 500	100	654	15.688	51622784	51345948	0.005391584
		500	2704	65.797	51539200	51347472	0.003733933
		1000	5601	132.984	51530104	51347868	0.003549047
	1000 by 1000	100	983	95.969	205948352	204746592	0.005869499
		500	3750	382.156	208803008	204521248	0.020935526
		1000	3092	296.875	206817920	204718256	0.010256359
R10&10000	100 by 100	100	3762	3.609	20679972	20495818	0.008984955
		500	1084	1.078	20947080	20423042	0.025659155
		1000	27123	25.563	20630444	20480756	0.007308715
	500 by 500	100	631	15.266	514117760	510787488	0.006519878
		500	3235	77.891	513098560	510815616	0.004469213
		1000	2222	54.844	517552896	510518848	0.013778234
	1000 by 1000	100	247	24.547	2055803008	2043563776	0.005989161
		500	725	71	2062268800	2043183232	0.009341095
		1000	2024	195.625	2097410432	2040447488	0.027916888
R100&10	100 by 100	100	4931	4.703	25832	25554	0.010868956
		500	3372	3.281	25793	25557	0.009224697
		1000	5324	5.094	26161	25485	0.02652525
	500 by 500	100	No solution found				
		500	18148	464.094	646863	636214	0.016738678
		1000	7602	180.954	643858	636342	0.011811954
	1000 by 1000	100	No solution found				
		500	7082	719.703	2578129	2544016	0.013409114
		1000	19227	1986.203	2583743	2542369	0.016273798
R100&1000	100 by 100	100	599	0.64	2034599	2026391	0.004050489
		500	2513	2.453	2055747	2025780	0.014792883
		1000	1161	1.203	2038601	2026116	0.006162222
	500 by 500	100	239	5.984	51351072	51089224	0.005125308
		500	1890	46.047	51468032	51085800	0.007482157
		1000	4159	100.079	51626880	51078156	0.010742831
	1000 by 1000	100	653	65.172	203929968	203645504	0.001396859
		500	1684	165.187	204953424	203623504	0.00653127
		1000	1542	147.562	207757808	203502864	0.020908521
R100&10000	100 by 100	100	327	0.375	20518944	20419806	0.004854992
		500	9129	8.641	20651124	20415284	0.011552129
		1000	2135	2.125	21028464	20385938	0.031518098
	500 by 500	100	338	8.421	511141120	507134816	0.00789988
		500	2530	62.172	519620352	506610176	0.025680842
		1000	4879	118.016	516204672	506764288	0.018628748
	1000 by 1000	100	548	53.016	2044883968	2028610560	0.008021948
		500	2106	206.406	2045396352	2028548736	0.008305256
		1000	1706	168.859	2047116288	2028514176	0.009170314
R1000&10	100 by 100	100	972	1.016	25635	25469	0.006504452
		500	2028	1.968	25791	25463	0.012872579
		1000	2013	2.016	26219	25408	0.031927489
	500 by 500	100	No solution found				
		500	6441	154.547	645388	635811	0.015063351
		1000	1891	45.703	644539	635946	0.013512751
	1000 by 1000	100	No solution found				
		500	5059	473.859	2579751	2541020	0.015242504
		1000	5560	524.984	2588844	2540298	0.019110457
R1000&100	100 by 100	100	336	0.359	208421	207788	0.00304562
		500	1289	1.265	209110	207761	0.00649107
		1000	1823	1.813	209511	207714	0.008649194
	500 by 500	100	545	13.438	5215290	5190803	0.004717382
		500	1019	25.047	5240548	5189998	0.009739888

(*continued on next page*)

Table 4 (*continued*)

Problem	Layer size	δ	# of Iterations	Time (s)	Obj_Heu	L_Bound	Gap
		1000	3673	88.984	5226785	5190345	0.007020825
	1000 by 1000	100	278	28.375	20823980	20744798	0.003816957
		500	817	81.438	21132332	20724534	0.019677065
		1000	5102	499.359	20870828	20743006	0.006162173
R1000&10000	100 by 100	100	199	0.25	20444748	20284480	0.007901016
		500	1032	1.031	20410928	20275350	0.006686839
		1000	2979	2.922	20442584	20273788	0.008325824
	500 by 500	100	650	15.875	510316864	508315584	0.003937082
		500	1146	28.375	511372992	508291328	0.006062791
		1000	3170	76.734	512030336	508268416	0.007401444
	1000 by 1000	100	189	19.422	2046162048	2031442432	0.007245894
		500	768	75.969	2052350848	2031226880	0.01039961
		1000	888	89.219	2045102208	2031485056	0.006703053
R10000&10	100 by 100	100	3440	3.344	25803	25538	0.010383185
		500	2250	2.218	25821	25536	0.011159786
		1000	3103	3	25845	25538	0.012002417
	500 by 500	100	No solution found				
		500	2993	72.266	643561	636580	0.010966216
		1000	1723	41.907	643958	636524	0.011678858
	1000 by 1000	100	No solution found				
		500	3361	316.875	2588661	2541340	0.018620492
		1000	6547	624.625	2602723	2537837	0.025567442
R10000&100	100 by 100	100	609	0.656	210792	208993	0.008608019
		500	1243	1.266	211314	208975	0.011192348
		1000	6339	6.125	211445	208957	0.011908572
	500 by 500	100	379	9.297	5219814	5187161	0.006294966
		500	786	19.391	5210964	5187518	0.004519695
		1000	2757	67.656	5205281	5187574	0.003413446
	1000 by 1000	100	275	28.109	20867572	20751700	0.005583735
		500	2940	286.313	20964672	20747700	0.010457641
		1000	5946	610.235	21259480	20726118	0.025733811
R10000&1000	100 by 100	100	179	0.235	2052743	2034921	0.00875808
		500	745	0.765	2039068	2035135	0.001932735
		1000	1910	1.89	2058570	2034662	0.011750541
	500 by 500	100	216	5.328	51054920	50870604	0.003623232
		500	763	18.781	51175816	50867464	0.006061871
		1000	4157	100.844	51048296	50870396	0.003497122
	1000 by 1000	100	358	35.687	204798832	202862592	0.009544589
		500	1520	148.016	203408720	202894368	0.002535073
		1000	1623	159.032	205280944	202808512	0.012190968

Problems R10&1000, R10&10000, R100&10, R100&1000, R100&10000, R1000&10, R1000&100, R1000&10000, R10000&10, R10000&100, R10000&1000, A10&100, A10&1000, A10&10000, A100&10, A100&1000, A100&10000, A1000&10, A1000&100, A1000&10000, A10000&10, A10000&100, A10000&1000 were similarly defined.

An exact solution (or a feasible solution with a known optimality gap) was found to all the problems listed above by the IP solver CPLEX 9.0, and a feasible solution by a stochastic version of our heuristic with δ set to 100 in Eq. (15). For each problem, we compared the objective value of the heuristic solution, OBJ_HEU, and that of the optimal solution (or a theoretical lower bound on it, if the problem cannot be optimally solved within two hours of computation), OBJ_EXA, and computed the optimality gap as (OBJ_HEU–OBJ_EXA)/OBJ_EXA. As shown in Table 3, all heuristic solutions were approximately within 1% of optimality. Note, however, that the exact solution method should be considered as a better option for problems of this size (5 zones and $100 \cdot 100$ cells) since it will likely find a proven optimal solution in a reasonable time (several seconds to minutes for most of our sample problems). As for the effect of autocorrelation, the heuristic achieved almost the same level of accuracy whether or not input layers were autocorrelated; however, this is not the case with running time (see the running time for A1000_A10 and A10000_A10 in Table 3).

5.2. Complexity

As stated earlier, the present heuristic is not guaranteed to find a solution in a finite number of iterations. To see how the running time increases with the size (i.e. the number of cells) of the input raster, we revisited Problems R10&100 to R10000&1000 (which had been considered on a 100-by-100 raster), reconstructed these 12 problems both on a 500-by-500 and 1000-by-1000 raster, and solved them using the map algebraic heuristic. Eq. (15) continued to be used for the determination of step sizes, but three different δ values, 100, 500, and 1000, were tried.

The results are summarized in Table 4. While the running time generally increases with the size of the input raster, the number of iterations does not seem to depend on it. Since the amount of computation per iteration is in theory linear to the number of cells (as well as the number of zones), the heuristic seems able to handle large-scale problems in practical sense. In fact, all the problems involving 1000-by-1000 layers were solved in tens of seconds to hundreds of seconds, except for one case (R100&10).

It is important to note that the problems involving 500-by-500 and 1000-by-1000 layers were too large to be solved by the exact solution algorithm, which prevented us from comparing heuristic solutions with exact solutions. As shown in Appendix, however, the present heuristic can provides not only a heuristic solution but a lower bound on the optimal objective value as a by-product.

Table 4 shows that each of the heuristic solutions obtained is good relative to this bound.

Another observation is that the choice of $\delta = 100$ led to a better solution and faster computation in most of our sample problems. However, when the size of the input raster was large (i.e. 500-by-500 or 1000-by-1000), this was not the case. As seen in Table 4, for the problems in which R10 layers represent cost (i.e. R100&10, R1000&10, and R10000&10), δ needed to be set larger for convergence. This may imply that the number of cells should have been factored into Eq. (15).

6. Conclusion and discussion

Map algebra is a widely used methodology for cartographic data analysis and synthesis with geographic information systems. Although it has been expected to be effective in prescriptive (not just descriptive) modeling, little has been known about how such a model can be systematically built in accordance with the conventions of map algebra.

Taking a school districting problem as a mean of illustration, this paper has presented a map algebraic procedure for least-cost allocation of multiple zones subject to capacity constraints. It approximately solves the problem through a cyclic process of zone allocation, zone size measurement, and layer weight adjustment, which are all specified by map algebra. This is an extension of a half-century-old manual technique for a large-scale transportation problem and general enough to be applicable in different contexts including districting and land allocation. Like its predecessor, this heuristic is intended to serve as an alternative to exact optimization methods such as integer programming when the NP-hard problem is too large to be solved exactly.

The map algebraic heuristic is not without shortcomings. It is designed to address only one kind of holistic allocation criterion, i.e. size. It is anticipated that planners encounter other holistic criteria, particularly shape-related criteria (see, e.g., Benabdallah & Wright, 1991; Ehler, Cowen, & Mackey, 1996). For instance, a fragmented, narrow or winding zone may be of little use in practice. Therefore, for more realistic and presumably less tractable problems, one may still need to rely on other heuristics outside the (currently recognized) scope of map algebra such as simulated annealing. Nevertheless, we hope that the cartographic modeling technique presented in this paper will help attract more interest in the prescriptive capabilities of GIS and map algebra, and expect that more sophisticated techniques exist and are worth pursuing in future research.

Finally, the present procedure may contribute to suitability analysis, too. While it is expected that the conventional weight-and-sum approach will remain useful, it is not without concerns. One of them relates to the validity of arithmetic operations on two values measured in different scales. For example, it makes little sense to weight and sum slope and distance values. However, this is the kind of routine that many GIS users do based on such an assumption that the less steep and the more distant to water body a location is, the more suitable for development it is. Then they somehow determine the weights of the two factors, say 2 for slope and 3 for distance-to-water, and calculate the suitability of each location accordingly. But on what basis can the correctness of the chosen weights be defended? One possible answer may be offered by solving a problem of finding a certain size of area whose total distance value is minimized and average slope value is bounded. A by-product of its solution is the weight of the slope layer relative to that of the distance layer. The roles of the two layers can be swapped in order to obtain the weight of the distance layer relative to that of the slope layer. We could even introduce a third layer (e.g. an acquisition cost layer) as a factor to be minimized/maximized while keeping the slope and distance layers for size measurement. A solution to the resulting problem will determine the weights of the slope and distance layers relative to that of the cost layer. In this way, one needs not to solicit for the weight of a layer per se but for the lower/upper bound on the size of a zone measured in terms of that layer. This implies that in suitability analysis there are some factors whose weights are not intrinsically fixed but vary depending on how target zones are constrained in size.

Acknowledgements

The author would like to thank the editor and anonymous reviewers for their valuable comments and suggestions, which have enhanced the structure and contents of this article. Any errors are, of course, the author's responsibility.

Appendix A

The map algebraic procedure presented in this paper involves a simple iterative process, which assigns each cell to its least-true-cost zone and adjusts the weight of the student population layer for each zone according to that zone's current (population) size. This supplementary section relates this intuitive yet rather informal approach to a mathematical technique called *Lagrangian relaxation*, and presents another heuristic for adjusting these weights.

In this appendix, let us consider the following problem that generalizes the districting problems considered in the paper.

$$\min z = \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} \tag{A1}$$

subject to

$$\sum_{i\in I} a_{ij}x_{ij} \leqslant b_j \quad \forall j \in J \tag{A2}$$

$$\sum_{j\in J} x_{ij} = 1 \quad \forall i \in I \tag{A3}$$

$$x_{ij} \in (0,1) \quad \forall i \in I, \forall j \in J \tag{A4}$$

Note that School Districting Problem (1)–(4) can be seen as a special case of this by letting $c_{ij} = d_{ij}a_i$ and $a_{ij} = a_i$ for each $i \in I$ and $j \in J$.

Lagrangian relaxation is a technique to transform an intractable mathematical program to a relatively tractable one. It does so by shifting (or dualizing) some of the constraints that cause the intractability into the objective function (to be minimized or maximized) as penalty terms. The choice of constraints to be dualized normally requires a careful examination of the problem's structure. In the present case, to make the resulting problem trivial, let us choose to dualize the constraints (A2). Accordingly, given a vector $\mathbf{u} = (u_1, u_2, \ldots, u_m) \geqslant \mathbf{0}$ where $m = |J|$, a Lagrangian relaxation of Problem (A1)–(A4) is constructed as follows:

$$\begin{aligned}\min z_{LR}(\mathbf{u}) &= \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} u_j\Big(\sum_{i\in I} a_{ij}x_{ij} - b_j\Big)\\ &= \sum_{i\in I}\sum_{j\in J}(c_{ij} + u_j a_{ij})x_{ij} - \sum_{j\in J} u_j b_j \end{aligned} \tag{A5}$$

subject to (A3) and (A4)

From an economic point of view, each u_j of $\mathbf{u}$—referred to as a Lagrangian multiplier—can be seen as the shadow price of a unit of some restricted resource. In the context in which students are the relevant resource, a Lagrangian multiplier corresponds to the shadow price of a unit of a school's capacity, which has been interpreted as a relative location advantage of a school in Vidale's solution and as a weight of the student population layer in our cartographic model.

The relaxation problem formulated as Eqs. (A3)–(A5) is easily solved by identifying $\min_j\{c_{ij} + u_j a_{ij}\}$ for each cell i, and setting the corresponding $x_{ij} = 1$ and all other $x_{ij} = 0$.

Since the relaxation problem plus Eq. (A2) would have its optimal objective value less than or equal to the optimal value for z (denoted by z^*) and greater than or equal to the optimal value for $z_{LR}(\mathbf{u})$ (denoted by $z^*_{LR}(\mathbf{u})$), it is obvious that:

$$z^*_{LR}(\mathbf{u}) \leqslant z^* \tag{A6}$$

This implies the optimal objective value of a Lagrangian relaxation of a minimization problem provides a lower bound on that of the original problem. Therefore, to find an approximate solution to Problem (A1)–(A4) or tighten the lower bound of its objective value is to find a $\mathbf{u}$ that maximizes $z^*_{LR}(\mathbf{u})$.

One approach to this end is to repeatedly solve the relaxation problem while adjusting $\mathbf{u}$ until $z^*_{LR}(\mathbf{u})$ is maximized. It is known that since $z^*_{LR}(\mathbf{u})$ is concave and differentiable except at points where it has multiple optimal solutions, an iterative procedure called the subgradient method aims to maximize $z^*_{LR}(\mathbf{u})$ by updating $\mathbf{u}$ in the direction of the subgradient of $z_{LR}(\mathbf{u})$ at each iteration (Fisher, 1981, 1985). Given an initial vector $\mathbf{u}^0$, the method updates u_j using the following rule:

$$u_j^{k+1} = \max[0, u_j^k + t^k(\sum_i a_i x_{ij}^k - b_j)] \quad \forall j \in J \tag{A7}$$

where

u_j^{k+1} is the value of u_j at the kth iteration,

x_{ij}^k is an optimal solution to Problem (A3)–(A5) at the kth iteration,
t^k is a positive scalar value called step size at the kth iteration.

While it is not known what step size (if any) guarantees convergence to an optimal solution, there is a formula for determining a good step size that is commonly used in practice (Fisher, 1981, 1985; Held, Wolfe, & Crowder, 1974):

$$t^k = \frac{\lambda^k (z_{LP}(\mathbf{u}) - \bar{z})}{\sum_{j\in J}(\sum_{i\in I} a_{ij} x_{ij}^k - b_j)^2} \tag{A8}$$

where

λ^k is a scalar value satisfying $0 < \lambda^k \leqslant 2$ at the kth iteration,

$\bar{z}$ is the objective value of the best known feasible solution to Problem (A1)–(A4).

It should be noted this formula can be expressed in map algebra, too.

Unfortunately, it cannot be proven that a given $\mathbf{u}^k$ maximizes $z^*_{LR}(\mathbf{u})$ unless $z^*_{LR}(\mathbf{u}^k) = \bar{z}$. Therefore, the subgradient method is usually terminated upon reaching a specified iteration limit (Fisher, 1985). Thus a heuristic solution obtained in this way is only expected to be a good approximation of an exact solution that may (or will) not been found. As implied by Eq. (A6), however, the heuristic solution provides a lower bound on the optimal value z^*. Therefore, if a feasible solution to Problem (A1)–(A4) happens to be found, its objective value should be compared with the greatest $z^*_{LR}(\mathbf{u})$ found so far (*i.e.*, the tightest lower bound on z^*) to see how good that feasible solution is relative to the not-yet-found z^*.

References

Aerts, J. C. J. H., & Heuvelink, G. B. M. (2002). Using simulated annealing for resource allocation. *International Journal of Geographical Information Science, 16*(6), 571–587.

Aerts, J. C. J. H., Eisinger, E., Heuvelink, G. B. M., & Stewart, T. J. (2003). Using linear integer programming for multi-site land-use allocation. *Geographical Analysis, 35*(2), 148–169.

Armstrong, M., Lolonis, P., & Honey, R. (1993). A spatial decision support system for school redistricting. *URISA Journal, 5*, 40–52.

Bação, F., Lobo, V., & Painho, M. (2005). Applying genetic algorithms to zone design, soft computing – a fusion of foundations. *Methodologies and Applications, 9*(5), 341–348.

Belford, P. C., & Ratliff, H. D. (1972). A network-flow model for racially balancing schools. *Operations Research, 20*, 619–628.

Benabdallah, S., & Wright, J. R. (1991). Shape considerations in spatial optimization. *Civil Engineering Systems, 8*(3), 145–152.

Caro, F., Shirabe, T., Guignard, M., & Weintraub, A. (2004). School redistricting: embedding GIS tools with integer programming. *Journal of the Operational Research Society, 55*(8), 836–849.

Cerdeira, J. O., Gaston, K. J., & Pinto, L. S. (2005). Connectivity in priority area selection for conservation. *Environmental Modeling and Assessment, 10*(3), 183–192.

Ceselli, A., & Righini, G. (2005). A branch-and-price algorithm for the capacitated p-median problem. *Networks, 45*(3), 125–142.

Church, R. L., Gerrard, R. A., Gilpin, M., & Sine, P. (2003). Constructing cell-based habitat patches useful in conservation planning. *Annals of the Association of American Geographers, 93*(4), 814–827.

Church, R. L., & Murray, A. T. (1993). Modeling school utilization and consolidation. *Journal of Urban Planning and Development, 119*(1), 23–38.

Cordeiro, J. P. C., Câmara, G., Moura de Freitas, U., & Almeida, F. (2009). Yet another map algebra. *GeoInformatica, 13*(2), 183–202.

Crema, S. C. (1996) Comparison between linear programming and a choice heuristic approach to multi-objective decision making using GIS. In *Proceedings of GIS/LIS '96*, Denver, CO, pp. 954–963.

desJardins, M., Bulka, B., Carr, R., Jordan, E., & Rheingans, P. (2007). Heuristic search and information visualization methods for school redistricting. *AI Magazine, 28*(3), 59–72.

Diamond, J. T., & Wright, J. R. (1987). Multiobjective analysis of public school consolidation. *Journal of Urban Planning and Development, 113*(1), 1–18.

Diaz, J. A., & Fernandez, E. (2006). Hybrid scatter search and path relinking for the capacitated p-median problem. *European Journal of Operational Research, 169*(2), 570–585.

Drexl, A., & Haase, K. (1999). Fast approximation methods for sales force deployment. *Manage Science, 45*(10), 1307–1323.

Duque, J. C., Ramos, R., & Surinach, J. (2007). Supervised regionalization methods: a survey. *International Regional Science Review, 30*(3), 195–220.

Eastman, J. R., Jiang, H., & Toledano, J. (1998). Multi-criteria and multi-objective decision making for land allocation using GIS. In E. Beinat & P. Nijkamp (Eds.). *Multi-criteria analysis for land-use management. Environment and management* (Vol. 9, pp. 227–251). Dordrecht, Netherlands: Kluwer Academic Publishers.

Eastman, J. R., Jin, W., Kyem, A. K., & Toledano, J. (1995). Raster procedures for multicriteria/multiobjective decision. *Photogrammetric Engineering and Remote Sensing, 61*(5), 539–547.

Eastman, J. R., Kyem, P. A. K., Toledano, J. (1993). A procedure for multi-objective decision making in gis under conditions of conflicting objectives. In *Proceedings of EGIS '93*, pp. 438–447.

Ehler, G. B., Cowen, D. J., & Mackey, H. E. Jr., (1996). Development of a shape fitting tool for site evaluation. In M. Molenaar & M. J. Kraak (Eds.), *The proceedings of the 7th international symposium on spatial data handling SDH '96. Spatial data handling* (Vol. 1). London: Taylor & Francis, pp. 4A.1–4A.12.

Fisher, M. L. (1981). The Lagrangian relaxation method for solving integer programming problems. *Management Science, 27*(1), 1–18.

Fisher, M. L. (1985). An applications oriented guide to Lagrangian relaxation. *Interfaces, 15*(2), 10–21.

Fleischmann, B., & Paraschis, J. N. (1988). Solving a large scale districting problem: a case report. *Computers and Operations Research, 15*(6), 521–533.

Forman, S. L., & Yue, Y. (2003). Congressional districting using a TSP based genetic algorithm. In E. Cantu-Paz, J. A. Foster, K. Deb, L. David, & R. Rajkumar (Eds.), *Genetic and evolutionary computation – GECCO 2003. Lecture notes in computer science* (Vol. 2724, pp. 2072–2083). Berlin: Springer.

Frank, A. U. (2005). Map algebra extended with functors for temporal data. In Jacky Akoka, Stephen W. Liddle, Il-Yeol Song, Michela Bertolotto, Isabelle Comyn-Wattiau, Samira Si-Said Cherfi, Willem-Jan van den Heuvel, Bernhard Thalheim, Manuel Kolp, Paolo Bresciani, Juan Trujillo, Christian Kop, Heinrich C. Mayr (Eds.), *Proceedings of perspectives in conceptual modeling, ER 2005 workshops AOIS, BP-UML, CoMoGIS, eCOMO, and QoIS*, Klagenfurt, Austria, October 24–28, 2005, Lecture Notes in Computer Science 3770, Berlin: Springer, pp. 194–207.

Franklin, A. D., & Koenigsberg, E. (1973). Computed school assignments in a large district. *Operations Research, 21*(2), 413–426.

French, K., & Li, X. (2010). Feature-based cartographic modelling. *International Journal of Geographical Information Science, 24*(1), 141–164.

Garfinkel, R. S., & Nemhauser, G. L. (1970). Optimal political districting by implicit enumeration techniques. *Management Science, 16*(8), B495–B508.

Garrison, W. L. (1959). Spatial structure of the economy: II. *Annals of the Association of American Geographers, 49*, 471–482.

Gilbert, K. C., Holmes, D. D., & Rosenthal, R. E. (1985). A multiobjective discrete optimization model for land allocation. *Management Science, 31*(12), 1509–1522.

Gutiérrez-Ándrade, M. A., García, E. A. R. (2009). Redistricting by square cells. In Hernández Aguirre, A., Borja, R. M., García, E. A. R., (Eds.), *MICAI 2009: advances in artificial intelligence: 8th Mexican international conference on artificial intelligence*, Guanajuato, México, November 9–13, Lecture notes in artificial intelligence, Vol. 45. Springer, Berlin, pp. 669–679.

Hess, S. W., & Samuels, S. A. (1971). Experiences with a sales districting model: criteria and implementation. *Management Science Part II, 18*(4), 41–54.

Hess, S. W., Weaver, J. B., Siegfeldt, H. J., Whelan, J. N., & Zitlau, P. A. (1965). Nonpartisan political redistricting by computer. *Operations Research, 13*(6), 998–1006.

Held, M., Wolfe, P., & Crowder, H. P. (1974). Validation of subgradient optimization. *Mathematical Programming, 6*(1), 62–88.

Hitchcock, F. L. (1941). The distribution of a product from several sources to numerous localities. *Journal of Mathematics and Physics, 20*, 224–230.

Holloway, C. A., Wehrung, D. A., & Zeitlin, M. P. (1975). An interactive procedure for the school boundary problem with declining enrollment. *Operations Research, 23*(2), 191–206.

Kalcsics, J., Nickel, S., & Schröder, M. (2005). Towards a unified territorial design approach – applications, algorithms and GIS integration. *Sociedad de Esradistica e Investigacion Operativa TOP, 13*, 1–74.

Lemberga, D. S., & Church, R. L. (2000). The school boundary stability problem over time. *Socio-Economic Planning Sciences, 34*(3), 159–176.

Li, X., & Hodgson, M. E. (2004). Vector field data model and operations. *GIScience and Remote Sensing, 41*(1), 1–24.

Lorena, L. A. N., & Senne, E. L. F. (1994). A column generation approach to capacitated *p*-median problems. *Computers and Operational Research, 31*(6), 863–876.

Lorena, L. A. N., & Senne, E. L. F. (2003). Local search heuristics for capacitated *p*-median problems. *Networks and Spatial Economics, 3*(4), 407–419.

Macmillan, W. (2001). Redistricting in a GIS environment: an optimisation algorithm using switching-points. *Journal of Geographical Systems, 3*(2), 167–180.

Maniezzo, V., Mingozzi, A., & Baldacci, R. (1998). A bionomic approach to the capacitated *p*-median problem. *Journal of Heuristics, 4*(3), 263–280.

Marlin, P. G. (1981). Application of the transportation model to a large-scale "districting" problem. *Computers and Operations Research, 8*(2), 83–96.

Mcdonnell, M. D., Possingham, H. P., Ball, I. R., & Cousins, E. A. (2002). Mathematical methods for spatially cohesive reserve design. *Environmental Modeling and Assessment, 7*(2), 107–114.

McHarg, I. L. (1969). *Design with nature*. Garden City, NY: Natural History Press.

Mehrotra, A., Johnson, E. L., & Nemhauser, G. L. (1998). An optimization based heuristic for political districting. *Management Science, 44*(8), 1100–1114.

Mennis, J., Viger, R., & Tomlin, D. (2005). Cubic map algebra functions for spatiotemporal analysis. *Cartography and Geographic Information Science, 32*(1), 17–32.

Middleton, R. S. (2006). *Geographical distillation: application of the p-median, traveling salesman, and regionalization problems*. PhD thesis, the University of California at Santa Barbara.

Nalle, D. J., Arthur, J. L., & Sessions, J. (2003). Designing compact and contiguous reserve networks with a hybrid heuristic algorithm. *Forest Science, 48*(1), 59–68.

November, S. M., Cromley, R. G., & Cromley, E. K. (1996). Multi-objective analysis of school district regionalization alternatives in Connecticut. *Professional Geographer, 48*(l), 1–14.

Okabe, A., Boots, B., Sugihara, K., & Chiu, S. N. (2000). *Spatial tessellations – Concepts and applications of voronoi diagrams* (2nd ed.). John Wiley.

Önal, H., & Briers, R. A. (2005). Designing a conservation reserve network with minimal fragmentation: A linear integer programming approach. *Environmental Modeling and Assessment, 10*(3), 193–202.

Önal, H., & Wang, Y. (2008). A graph theory approach for designing conservation reserve networks with minimal fragmentation. *Networks, 51*(2), 142–152.

Ploughman, T., Darnton, W., & Heuser, W. (1968). An assignment program to establish school attendance boundaries and forecast construction needs. *Socio-Economic Planning Sciences, 1*(3), 243–258.

Rardin, R. L. (1997). *Optimization in operations research*. Englewood Cliffs, NJ: Prentice-Hall.

Salazar-Aguilar, M. A., Ríos-Mercado, R. Z., & Cabrera-Ríos, M. (2011). New models for commercial territory design. *Networks and Spatial Economics, 6*, 1–21.

Santos-Correa, E., Steiner, M. T., Freitas, A. A., & Carnieri, C. (2004). A genetic algorithm for solving a capacitated *p*-median problem. *Numerical Algorithms, 35*(2–4), 373–388.

Segal, M., & Weinberger, D. B. (1977). Turfing. *Operations Research, 25*(3), 367–386.

Schoepfle, O. B., & Church, R. L. (1991). A new network representation of a "classic" school districting problem. *Socio-Economic Planning Sciences, 25*(3), 189–197.

Shanker, R. J., Turner, R. E., & Zoltners, A. A. (1975). Sales territory design: An integrated approach. *Management Science, 22*(3), 309–320.

Takeyama, M., & Couclelis, H. (1997). Map dynamics: Integrating cellular automata and GIS through geo-algebra. *International Journal of Geographical Information Systems, 11*(1), 73–91.

Tomlin, C. D. (1990). *Geographic information systems and cartographic modeling*. Englewood Cliffs, NJ: Prentice-Hall.

Tomlin, C. D., & Johnston, K. M. (1990). An experiment in land-use allocation with a geographic information system. In Donna J. Peuquet & Duane F. Marble (Eds.), *Introductory readings in geographic information systems* (pp. 159–169). London: Taylor & Francis.

Tomlin, C. D. (1994). Map algebra: One perspective. *Landscape and Urban Planning, 30*(1–2), 3–12.

Upchurch, C., Kuby, M., Zoldak, M., & Barranda, A. (2004). Using GIS to generate mutually exclusive service areas linking travel on and off a network. *Journal of Transport Geography, 12*, 23–33.

Vidale, M. L. (1956). A graphical solution of the transportation problem. *Operations Research, 4*(2), 193–203.

Williams, J. C. (1995). Political redistricting: A review. *Papers in Regional Science, 74*(1), 13–40.

Wright, J., Revelle, C., & Cohon, J. (1983). A multiobjective integer programming model for the land acquisition problem. *Regional Science and Urban Economics, 13*(1), 31–53.

Yeates, M. (1963). Hinterland delimitation: A distance minimizing approach. *The Professional Geographer, 15*(6), 7–10.

Zoltners, A. A., & Sinha, P. (1983). Sales territory alignment: A review and model. *Management Science, 29*(11), 1237–1256.

Zoltners, A. A., & Sinha, P. (2005). Sales territory design: Thirty years of modeling and implementation. *Marketing Science, 24*(3), 313–331.