An Optimization Model for the P-Compact-Regions Problem and Its Application to Regional Economic Modeling

Wenwen Li[1, 2], Richard L. Church[1], Michael F. Goodchild[1, 2]

[1] Department of Geography, University of California, Santa Barbara, CA 93106-4060

[2] Center for Spatial Studies (Spatial@UCSB), University of California, Santa Barbara, CA 93106-4060

The p-compact-regions problem involves the search for an aggregation of n atomic spatial units into p compact regions to fit a specific purpose. This paper reports our efforts in designing a specialized randomized greedy and edge-reassigning heuristic to solve the problem to near optimally. We apply this model to support urban economic simulation, in which activities need to be aggregated from the 4109 Transportation Analysis Zones (TAZs) of six southern California counties into 100 model zones to achieve desired computational feasibility of the economic simulation model. Spatial contiguity, physiography, political boundaries, the presence of local centers, and intra-zonal and inter-zonal traffic are considered, and efforts are made to ensure consistency of selected properties between the disaggregated and aggregated regions. The moment of inertia introduced in a novel approach to computing the compactness of the model zones was shown to be robust and to perform better than the familiar and often used isoperimetric quotient (IPQ). The average compactness of the generated model zones reached 0.89 (the theoretical upper limit is 1.00). Our proposed approach provides a substantial extension and a computationally efficient process as compared to alternative approaches such as location-allocation models, automated zonation, political redistricting, and patch-growing processes. A Web application was developed based on ArcGIS Server to allow users to view and work interactively with the optimized solution.

Keywords: Spatial optimization, greedy, heuristic, compactness, moment of inertia, zoning

Correspondence: Dr. Wenwen Li, Spatial@UCSB, University of California, Santa Barbara, CA 93106-4060.
e-mail: wenwen@spatial.ucsb.edu

Submitted: April 18, 2011.

Introduction

The *p*-compact-regions problem involves aggregating or combining n spatially contiguous units into *p* compact regions ($p \leq n$) while optimizing a defined objective function. According to the literatures of geography, geographic information science (GIScience), computer science, and regional science this type of aggregation can also be termed zonation, districting, and regionalization (Duque, Church, and Middleton 2011) and finds widespread application in many domains. In GIScience, this problem was first addressed by Openshaw (1977), in discussing the design of algorithms to aggregate data from small areas to larger zones, and in investigating the effects of such aggregations on statistical and mathematical analyses -- effects that he collectively termed the modifiable areal unit problem.

There are a substantial number of real-world problems that require solutions to *p*-region problems (problems that do not necessarily include compactness as an objective or constraint). For example, emergency medical service providers aim to partition a service area into zones in order to provide satisfactory coverage (Marianov and ReVelle 1996; Alsalloum and Rand 2006; Sorensen and Church 2010). In conservation planning, decision makers may want to identify regions that maximize the coverage of selected species (Church, Storm, and Davis 1996), perhaps combined with the minimization of cost (Moilanen 2007). Partitioning an area into compact electoral districts is an effective alternative to political gerrymandering (Young 1988; Pang et al. 2010). In human geography, a single set of reporting zones for projects such as the census is no longer sufficient to meet researchers' needs (Young, Martin, and Skinner 2009), so flexible aggregation has become increasingly important. For microeconomic modeling of urban land use and transportation, such as the project that provides the case study for this paper, it is necessary to aggregate economic activity in order to achieve computational feasibility, and at the same time to preserve certain properties between the atomic and aggregated zones (Anas and Liu 2007).

Over the past several decades, researchers have developed a number of methods for providing solutions to versions of the *p*-regions problem. Optimal solutions have been obtained using integer programming and integer goal programming to contiguous, but non-compact form. However, the computational complexity and the non-linear nature of "compact" versions of the *p*-

region problem make it difficult to formulate and solve using general-purpose optimization software. This has led researchers to develop heuristic algorithms, such as greedy methods, tabu search, and simulated annealing. In this paper, we propose a heuristic that combines the randomized greedy and edge-reassignment methods to solve a large *p*-compact-regions problem for the potential application of a microeconomic model- RELU-TRAN. RELU-TRAN (Anas and Liu 2007) is a dynamic general-equilibrium model of a metropolitan economy and its uses of land. It equilibrates floor space, land and labor markets, and the market for the products of industries, treating development (construction and demolition), spatial interindustry linkages, commuting, and discretionary travel. Mode choices and equilibrium congestion on the highway network are treated by integrating an algorithm of stochastic user equilibrium.

Our goal is to aggregate the 4109 Transportation Analysis Zones (TAZs) in six counties (Los Angeles, Riverside, San Bernardino, Orange, Imperial, and Ventura, as Fig. 1 shows) of Southern California to approximately 100 model zones, the maximum number considered computationally feasible for subsequent work in microeconomic simulation of the region. Meanwhile, factors including spatial contiguity, model-zone compactness, the coincidence of model-zone boundaries with physiographic features and political boundaries, and the patterns of intra-zonal and inter-zonal traffic need to be considered. We call this problem "supreme" because while realistic it nevertheless requires iterative computation among more than 4000 atomic units, which is large in comparison with the literature on similar problems; because of the many constraints and the complexity of the objective function; and because several elements of the objective function are non-linear.

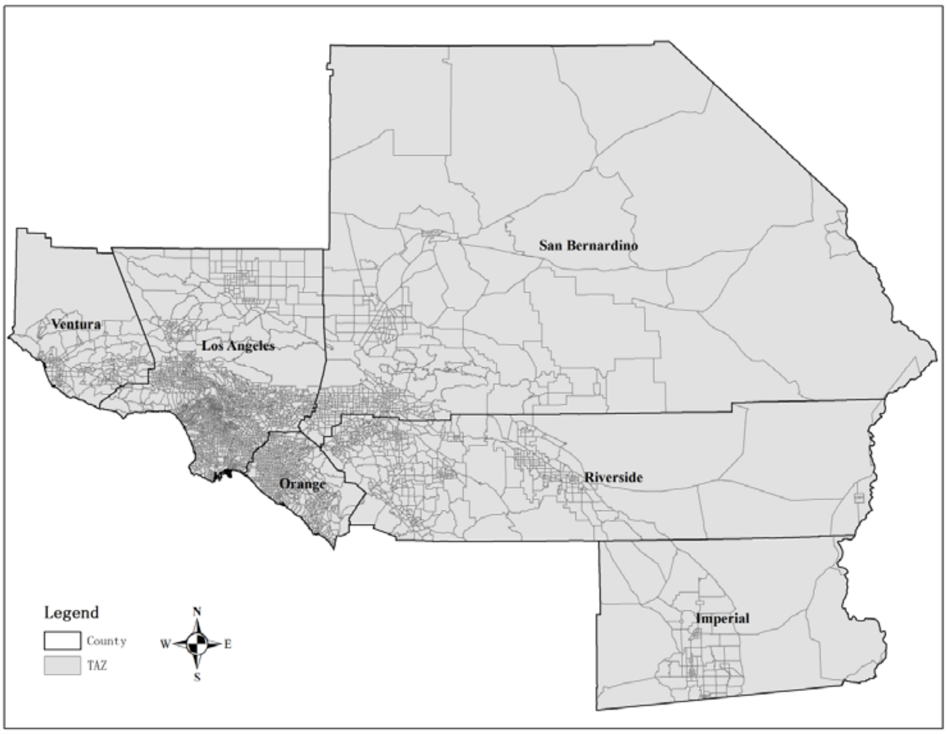

Figure 1. Case study for the *p*-compact-regions problem: grouping more than 4,000 TAZs in 6 counties of Southern California into approximately 100 compact regions.

Previous studies have acknowledged the importance of compactness -- "one of the most intriguing and least-understood properties of geographic shapes" (Angel, Parent, and Civco 2010), in zoning algorithm design. Compactness is seen as desirable for two reasons: it implies maximum accessibility to all parts of the zone; and according to Tobler's First Law of Geography (Tobler 1970) a compact zone is likely to be homogeneous, sharing common attributes and properties. Metrics of compactness often compare a shape to a circle of the same area as the shape, on the grounds that a circle minimizes the sum of squared distances of all parts of its area from its centroid. However, many studies do not provide an explicit and quantitative measurement of compactness, relying instead on indirect objectives, such as minimizing travel distances within a zone (Wie and Chai 2004) and minimizing the perimeter lengths of zones (Fischer and Church 2003).

In this paper we use an effective and quantitative measure of compactness based on the moment of inertia (Massam and Goodchild 1971), a quantity well-known in physics. The moment of inertia of a two-dimensional body about an axis through a point and perpendicular to the body is defined as the second moment of the body about the point, in other words the integral over the body's area of the squared distance from the point. Techniques for employing the moment of inertia in this way are discussed in detail below. The moment of inertia is easily calculated from the shape's defining vertices, and has the advantage of being additive, so that compactness of a proposed aggregation can be readily calculated from properties of the building blocks.

In summary, our purpose in this paper is to explore various aspects of the *p*-compact-regions problem, using a real, practical example of substantial size. Rather than a one-size-fits-all solution, we believe that any implementation of the *p*-compact-regions problem must respond to the specific circumstances. The circumstances we explore in this paper may well be applicable more generally, and we believe that the experience we have gained in this project, and the solutions we have found to many of its issues, may provide useful guidance to others. The paper also makes a substantial and original contribution in the use of a compactness measure based on the moment of inertia of a shape about a vertical axis through its centroid.

Literature

The family of *p*-region problems involves the aggregation of basic spatial units into groups to meet a range of criteria, including maximizing both intra-group homogeneity and inter-group heterogeneity. Openshaw (1977) defined a general model of the zone-design process by aggregating *n* zones into *p* contiguous regions, where *p* is usually a small fraction of *n*. A closely related variant is the location-allocation problem (Goodchild 1979; Goodchild 1988), which seeks to locate *p* facilities optimally, and simultaneously to allocate distributed demand to them, forming *p* zones. Models of this type have been applied in many fields and are used to solve a variety of problems, ranging from public-facility location-allocation (Murray, O'Kelly, and Church 2006; Church and Scaparra 2007; O'Hanley and Church 2010) to business service-center allocation (Aboolian, Sun, and Koehler 2009).

Morrill (1976) proposed to use a computer-based capacity-constrained location-allocation model to generate equally populous congressional districts such that the aggregated travel in each district to its center was minimized. The algorithm starts by manual selection of initializing point centers for each district. Then an application of the transportation problem is defined, to assign building blocks to each center such that the total distance of assignment is minimized. This results in a set of regions surrounding each center (though in principle a center may find itself outside its allocated region). The centers can now be relocated to minimize distance within each region. By alternating assignment (allocation) and relocation until no further changes occur, the approach yields compact regions of approximately equal population. This procedure is relatively simple, being based on straight-line travel rather than the real transportation network, and not considering similarity among the atomic geographic units in the aggregation process. Similar works in the literature include those of Mehrotra, Johnson, and Nemhauser (1998), Bozkaya, Erkut, and Laporte (2003), Ricca and Simeone (2008), and Altman and McDonald (2010).

Church et al. (2003) proposed a patch-growing process (PGP) to identify an optimum habitat patch for the San Joaquin kit fox. This process starts by initializing a seed cell which has a minimum specified suitability value. It then repeatedly calculates a composite suitability score for each cell, based on the number of edges that the cell shares with the current patch, and edge weights for each cell. At each iteration, the process tends to select the most suitable cells to add into the patch; therefore it can be considered a greedy algorithm. Meanwhile, by limiting the number of cells eligible to join the patch at any iteration, the algorithm provides a flexible strategy to control the growing speed. The PGP algorithm achieves a substantial improvement in comparison to Brookes' (Brookes 1997) PRG (parameterized region-growing) algorithm, but it does not provide solutions for problems requiring the growing of multiple regions simultaneously, such as the problem discussed in this paper.

Our problem is one of zonation with multiple objectives and constraints. First, the compactness of the model zones should be maximized. Second, in order to achieve homogeneity, model zones are not allowed to cross either county boundaries or physiographic barriers. Third, another goal is to keep model zones from being too large as well as assisting in keeping the errors in modeling transportation flows on the highway network small as compared to a disaggregated TAZ based flow estimation. To achieve this goal, we require that intra-zonal traffic occurring within each model zone be equal to or less than a certain portion (for example, 10%) of the total inter-zonal traffic. This problem is substantially different from those described in the literature both conceptually and computationally, because of the large number of building blocks and zones, the complexity of the objective function and constraints, and the innovative approach to compactness. In the next sections, we formalize the model and discuss our greedy and edge-reassigning heuristic algorithm for solving the *p*-compact-regions problem.

Problem statement and model formalization

Our goal is to aggregate 4109 TAZs into approximately 100 model zones, the maximal number considered feasible for the applied urban economic model of RELU-TRAN. Conceptually, we seek to maximize overall *compactness* of the model zones and at the same time keep the *spatial contiguity* of each zone. We define a number of linear *physiographic features*, aligned along major mountain-range barriers, and constrain model zone boundaries so as not to cross these features. We also constrain model zones not to cross county boundaries. Using data on traffic between TAZs, we compute intra-zonal and inter-zonal traffic and constrain the solutions such that intra-zonal traffic is less than or equal to a defined proportion of total traffic. Traffic is measured by the total number of zone-to-zone trips in multiple modes, such as drive-alone trips, shared-ride trips, etc. We also identify a set of economic *subcenters* (large shopping centers and large concentrations of employment) by manual inspection, and constrain the solution so that subcenters are not split between zones (see Section "Selection of seeds"). To formulate the *p*-compact-regions problem, consider the following problem parameters:

i , j : Index of units which are either subcenters (a small aggregation of TAZs) or independent TAZs (TAZs not part of a subcenter)

k : Index of zones

u : Index of zones

$t_{i,j}$: Traffic between unit i and unit j ;

Cty_i : The county that unit i belongs to;

$A_i = \sum_j (t_{ij} + t_{ji})$: A_i is the traffic between unit i and all other units;

$PS = \{(i, j) \mid$ unit i and unit j are separated by a physiographic boundary$\}$;

$S = \{i \mid i$ is a subcenter unit and not an independent TAZ$\}$;

The decision variables are:

$$X_{i,u} = \begin{cases} 1, \text{ if unit } i \text{ is assign to zone } u \\ 0, \text{ if not} \end{cases}$$

$$T_{i,j,u} = \begin{cases} 1, \text{ if both } i \text{ and } j \text{ are assigned to zone } u \\ 0, \text{ if not} \end{cases}$$

Consequently, can define the model as follows:

Maximize:

$$\sum_k C_k = \sum_k \frac{A_k^2}{2\pi I_k^G} \tag{1}$$

Subject to:

$$X_{i,u} + X_{j,u} \leq 1 + T_{i,j,u} \tag{2}$$

$$X_{i,u} + X_{j,u} \leq 1, \text{ for } \forall u, \text{ and for } \forall (i, j) \in PS \tag{3}$$

$$Cty_i = Cty_j, \text{ for } \forall i, j, \text{ where } X_{i,u} = X_{j,u} = 1 \tag{4}$$

$$\sum_i \sum_j t_{i,j} T_{i,j,u} \leq \theta \sum_i A_i X_{i,u} \tag{5}$$

$$X_{i,u} + X_{j,u} \leq 1, \text{ for } \forall i, j \in S \tag{6}$$

The objective function (1) aims to maximize the overall compactness of all model zones. Compactness of a model zone k is measured by the moment of inertia I_k^G, defined as the second moment of the zone about an axis perpendicular to it and passing through its centroid G:

$$I_k^G = \int_{A_k} d^2 da \tag{7}$$

where da is an infinitesimal part of the area and d is the distance from da to G. It can be shown that the moment of inertia of a circle of area A about its centroid is $A^2/2\pi$. Thus our measure of compactness of model zone k is:

$$C_k = \frac{A_k^2}{2\pi I_k^G} \tag{8}$$

and ranges from 0 in the case of an infinitely extended shape to 1 in the case of the most compact figure, a circle. Further details on the calculation of compactness of model zones are included in Section “Compactness measure”.

Constraint (2) is always true given the above definition. Inequation (3) is used to formulate the physiography constraint and equation (4) with conditions defines the county boundary constraint. Constraint (5) limits the fraction of traffic that is intra-zonal to be less than or equal to a fraction θ of the zone’s total intra- and inter-zonal traffic. We elaborate on this constraint below in Section “Modeling zonal traffic”. Inequation (6) formulates the subcenter constraint.

Methodological framework for the model

In order to find a near-optimal solution to the *p*-compact-regions problem, we propose a method based on the combined use of a novel form of randomized greedy and edge-reassigning algorithms. Due to the nature of the objective and constraints, there is a need to tailor and address the following issues when implementing the region-growing procedure. First, our initial condition, dictated by the need to solve a massive equilibrium model, is 100 manually selected seed TAZs (the selection procedure is discussed in detail in Section “Selection of seeds”), and the purpose of the greedy algorithm is to assign all the other TAZs in our study area into the 100 zones with the strict restriction that each TAZ can only be assigned to one model zone. Second, at each iteration only one TAZ will be assigned to a model zone, meaning that only one model zone has a chance to grow in a given iteration. The TAZ that is assigned must lead to an improvement in the objective function that is at or close to the greatest possible improvement in that iteration. Always choosing the greatest improvement may not yield a good final solution. However, introducing a probabilistic selection strategy, such as randomly selecting one of the best *N* choices, will generally lead to a better solution if the procedure is executed a large number of times. To support this, the algorithm should maintain a candidate list that saves the best *N* choices for each of the model zones.

Third, before the assignment of any TAZ is complete, the objective function of each model zone must be updated if the availability of a TAZ is changed. Note that instead of updating the status of all model zones after a TAZ is assigned, only the model zones that are affected need to be updated. For example, when TAZ i is assigned at an iteration, the status of the model zone to which it is assigned must be updated, by recalculating the objective, the quantitative value for each constraint, and the objective function. In addition, the algorithm needs to identify new candidate TAZs that can improve composite objective function of the model zone. Other model zones that may be affected are those that already specify TAZ i in their candidate lists. For such zones, if TAZ i becomes unavailable, another candidate TAZ should be identified to replace it in the zone's candidate list (if any are adjacent and unassigned).

Finally, as the greedy algorithm does not include backtracking from a suboptimal solution after the assignment of all the TAZs is completed, we introduce an edge-reassigning algorithm to examine each TAZ that lies at the edge of a model zone, and to evaluate moving it from its current model zone to the one across the edge/boundary, if the objective function sum over all model zones is thereby improved and if no constraints are violated. The procedure will cease when no further improvement can be obtained by edge reassignment.

Fig. 2 shows the greedy algorithm for constructing initial model zones (MZs) in the form of pseudo code. Before explaining the design of the heuristic algorithm, it is necessary to introduce the following important data structures used in the algorithms:

(1) ***M***: A HashMap of TAZ-TAZ adjacency: records all the spatially adjacent TAZs of each TAZ, indexed by the ID of each TAZ. Length(*M*)=4109;

(2) ***ZstateProperties***: A two-dimensional array maintaining the status information of partial solutions for each model zone. The size of the first dimension of the array is 100, equaling the number of model zones to be produced; the size of the second dimension of the array is 8, which includes the greedy function, area, perimeter, moment of inertia, *x* and *y* coordinates of the centroid, and intra-zonal and inter-zonal traffic of the model zone.

(3) ***roundGreedyParameter***: A three-dimensional array maintaining the new status of each model zone if any of the *N* TAZs that would provide the *N* best improvements to the objective function were added to the zone. The size of the first dimension is 100, the number of model zones to be produced. The size of the second dimension is *N*, and the

third dimension is defined by the properties of the model zones. Besides the eight properties mentioned in (2), one other item is saved: the ID of the TAZ to be potentially added. Note that the TAZ must be currently unassigned to be considered as a candidate. A list is built to save the TAZs that are already assigned. Both *roundGreedyParameter* and *ZstateProperties* include the status information of model zones; the difference is that *roundGreedyParameter* contains the possible future states of a model zone and *ZstateProperties* contains the current state of a model zone.

```
# step0
initialize each MZ with one or more seed TAZs and the status information of the MZ;
while (not all the TAZs are taken):
    for each MZ: # step1
        deal each MZ n TAZs to let them all reach an acceptable initial condition;
    end for
    for each MZ: # step2
        if this MZ grew last time or
          the TAZ that is desired by this MZ was taken in the previous iteration:
            for each TAZs e in MZ:
                neighsOfe ← M[e] # get all the neighbor TAZs of e;
                for each neighbor TAZ in neighsOfe:
                    if the neighbor TAZ i is not taken and the TAZ lies in the
                      same county as all other TAZs in MZ and adding the
                      TAZ will not cross the physiographic barrier:
                        Suppose TAZ i is going to be added to the MZ
                        (1) update the greedy function as Obj_new;
                        (2) update other property information of the MZ;
                        Obtain current objective from ZstateProperties as Obj;
                        if (Obj_new-Obj) is in the N plans that best improve
                          the objective:
                            update entry in roundGreedyParameters;
                        end if
                    end if
                end for
            end for
        end if
    end for
    randomly select one plan out of the N best plans for each model zone;
    allow MZ k to grow where k has the greatest increase over all the selected plans;
    the TAZ which has just been assigned is marked as unavailable;
end while
```

Figure 2. Greedy algorithm

As Fig. 2 shows, the algorithm starts by designating one or more TAZs as the seeds for growing the 100 model zones (the selection of seeds is discussed in Section “Selection of seeds”). As the model zones grow, the status information, such as area, centroid, and moment of inertia, is updated. Areas and the coordinates of the centroids are used to calculate the moments of inertia of model zones, as discussed in the section titled “Compactness measure”.

To ensure that all model zones grow initially to a viable size, we first deal a fixed number *K* of TAZs to each zone, selecting them on the basis of the objective function. We call this “dealing,” as it involves something akin to a card dealer in poker, where the dealer deals in a clockwise fashion 1 card to each player for a set size of the card hand. Here, we “deal’ one of the best TAZs to a seed, and continue dealing until each seed has one of the best neighboring TAZs assigned to it. We then “deal” out a second set of TAZs, a third set, etc. until each seed cluster has a size of K+1 units (K assigned TAZs plus the original seed unit). In our study we set the number of TAZs dealt in this way to *K*=10, thus accounting for roughly a quarter of all TAZs assigned. Without this dealing step, some model zones may fail to grow beyond two or three TAZs if adding another TAZ will decrease rather than increase the objective function. We suggest that some experimentation with *K* may be needed in other applications of this heuristic if the values of *n* and *p* are significantly different from those in our study. This dealing step is deceptively simple, but extraordinarily powerful, as a model zone is quite awkward in size and shape when it contains only 2 or 3 TAZs, and adding a TAZ at this point often makes a cluster less compact rather than more compact. But, allowing all zones to grow (by adding TAZs) beyond an “awkward” shape helps to tune the heuristic towards generating a high percentage of compact zones.

After the dealing step, the randomized greedy algorithm is applied globally over the 100 model zones. For each model zone, the *N* best plans for adding a TAZ are selected by looking at all possible TAZs that can be added. One of the *N* is then randomly selected. At this point, a TAZ may be eligible to be added to several different model zones. After the plans for all of the model zones are identified, the model zone with the best objective improvement is given the chance to grow. Note each Zone has a selected candidate, but that candidate was one of the N best possibilities. This means that there is a degree of randomness associated with which TAZ that will be added to a zone, and which zone will be selected for the next assigned TAZ. This is a form of randomized adaptive search. After adding the selected TAZ, the plans for the other model zones are updated as necessary.

```
while True:
    Generate an array randomArr that contains all the TAZ ids
    while randomArr is not null:
        Randomly pick one TAZ i from randomArr;
        Remove this TAZ from randomArr;
        Get all MZs that contain neighboring TAZs of TAZ i except for MZ k that TAZ i
        belongs to and put them in list neigbrMZ;
        if TAZ i is not an inner TAZ (neigbrMZ is null) or moving TAZ i from MZ k will
        not break the connectivity of the remaining TAZs in MZ k:
            find the greatest gain by moving TAZ i to the any MZ in neigbrMZ;
            if the greatest gain increases the objective value:
                move TAZ i from MZ k to MZ u, including
                update ZstateProperties for both MZ k and MZ u;
                reset Count to 0;
            end if
        end if
    end while
    Count increased by 1;
    if the number of TAZs that have been continuously inspected is more than the
    number of TAZs:
        the optimized solution found, program quits;
    end if
end while
```

Figure 3. Edge-reassigning algorithm

As the greedy algorithm does not include backtracking to find a global optimal solution, the result may only be a suboptimal solution. To overcome this limitation, the edge-reassigning algorithm (as Fig. 3 shows) is designed to tune the results. The input of the edge-reassigning algorithm is the zone partition plan obtained from the greedy algorithm. A number of variables are reserved, such as the TAZs assigned to a model zone, and the current objective values and other properties of the model zones. Each TAZ is then examined in turn and compared to its adjacent TAZs. If at least one of its adjacent TAZs is assigned to a different model zone, then the TAZ is identified as an edge TAZ and becomes a candidate for reassignment. In order for each model zone to have an equal opportunity to grow, the sequence for checking potential reassignments of edge TAZs is randomized. If moving an edge TAZ from its current model zone MZ_k to the adjacent model zone MZ_u improves the overall objective, then this TAZ is removed from MZ_k and reassigned to MZ_u. The mathematical expression to decide whether the overall objective increases or not is as shown in Eq. (9). Let $O(MZ_k)$, $O(MZ_u)$, and $O\_new(MZ_k)$, $O\_new(MZ_u)$ be the contributions of MZ_k and MZ_u before and after edge reassignment respectively. Then reassignment is made if:

$$O(MZ_k) + O(MZ_u) > O\_new(MZ_k) + O\_new(MZ_u) \quad (9)$$

and if the reassignment will not violate any of the constraints. This process is repeated until no further improvement can be obtained.

In the greedy algorithm, the TAZ that is added to a model zone is selected from a pool of all the unassigned TAZs adjacent to the model zone. Therefore, contiguity of model zones is preserved without the need for a specific strategy. In the edge-assigning process, however, removing one TAZ from a model zone may potentially violate the contiguity constraint by leaving a residual zone that is split into two unconnected parts. To address this issue, we could check whether there is a path between any of the TAZ pairs in the model zone, but this would be a very time-consuming process. The strategy we use here is to select one TAZ within a model zone, and to find its adjacent TAZs within the same zone until no more TAZs can be added. If the size of the identified TAZ set is less than the size of the model zone, it means there are one or more TAZs that are not reachable after the edge reassignment. Therefore, the edge TAZ cannot be reassigned.

Implementation of the computational model

In this section, we will discuss in detail how the objective and constraints of the model are computed quantitatively.

- **Selection of seeds**

The seeds are the starting points for growing model zones. As noted earlier, we define a set of subcenters, which are areas where the employment concentration is relatively high (McDonald 1987). In this paper, these employment subcenters are contiguous (not necessarily compact) areas composed of TAZs, as shown in Fig. 4(a) and (b). Based on Giuliano and Small's procedure (1991), 51 subcenters are identified in our study area, including 19 in Los Angeles, 11 in Riverside, 10 in Orange, 6 in San Bernardino, 4 in Imperial, and 1 in Ventura Counties. One of the constraints is that one subcenter cannot be assigned to more than one model zone, and one model zone cannot contain more than one subcenter. Therefore, the subcenter itself can be a natural seed. As we need 100 model zones in total, another 49 TAZs are selected manually as the supplemental seeds to the existing subcenter seeds. To select these supplemental seed TAZs we first identified junctions in the major-highway network, which tend to be more frequent in areas of high population density. We also selected some points arbitrarily in order to augment the subcenters and major-highway junctions. Although some of the seed TAZs are selected manually, they serve only to initialize the growth of model zones. During the edge-reassigning process, these TAZs can be swapped out of the original model zone and therefore will not have a great effect on the final result.

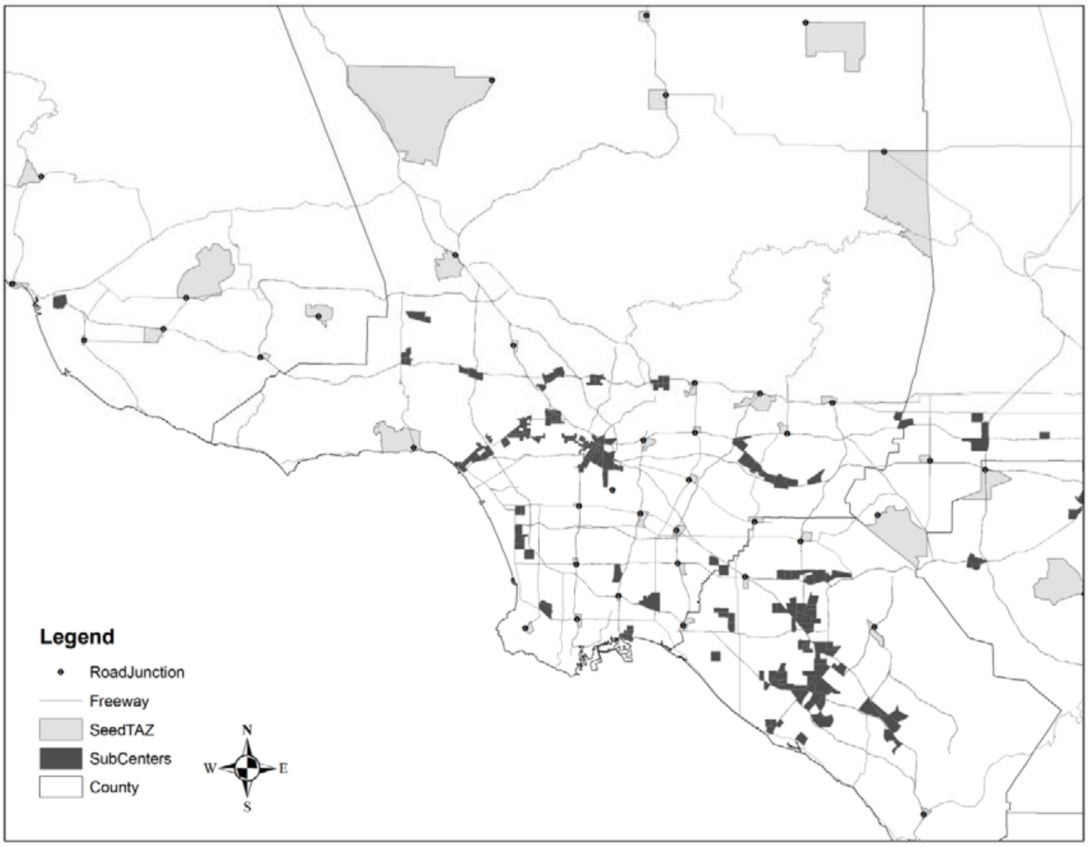

Figure 4(a). Seed TAZs in Ventura, Los Angeles, and Orange Counties.

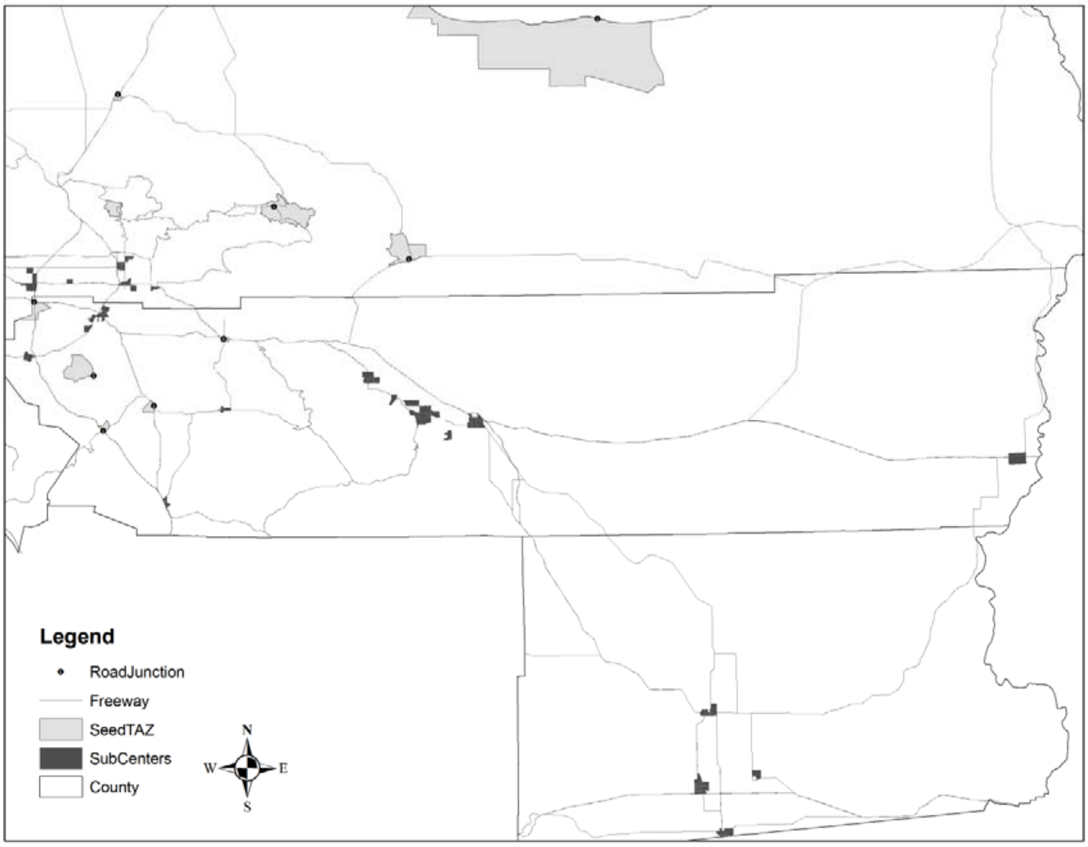

Figure 4(b). Seed TAZs in San Bernardino, Riverside, and Imperial Counties.

- **Compactness measure**

The computation of moment of inertia includes four parts: (1) moment of inertia of 4109 TAZ shapes, computed about the shape's centroids; (2) moment of inertia of a model zone, computed from its constituent TAZs; (3) the change in moment of inertia of a model zone when a new TAZ is added during the dealing and randomized greedy algorithms; and (4) the change of this property when a TAZ is detached during the edge-reassignment phase.

Moment of inertia is minimal about the centroid, by definition. Therefore, to compute the moment of inertia for each TAZ *i* it is necessary to compute the coordinates of its centroid *g*:

$$x_g^i = \int_{a_i} x da_i \Big/ a_i \tag{10}$$

$$y_g^i = \int_{a_i} y da_i \Big/ a_i \tag{11}$$

where x in (10) is the coordinate of da and a is the area's measure, and similarity for (11). To compute these properties it is necessary to represent the area as a polygon, and to partition the polygon into primitive elements for which calculation of moments of inertia and centroids is simple. The properties of these primitive elements can then be combined using straightforward principles. In the literature (Massam and Goodchild, 1971), triangles were used as the primitive elements. However, the triangle method is tricky for this study because there is no simple, obvious, and robust way to divide the TAZ polygons into triangles uniquely. Instead, we use a trapezium-based strategy for computing a polygon's moment of inertia. As the Fig. 5 shows, for any adjacent pair of

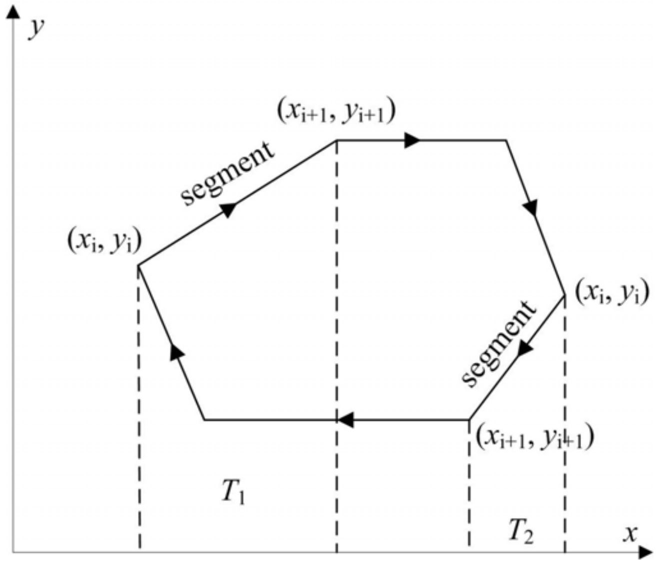

Figure 5. A polygon with points defined clockwise, showing the trapezia generated from two of its segments.

vertices, such as the segment $(x_{i,}y_i)$ to $(x_{i+1,}y_{i+1})$, when dropping perpendiculars to the x axis, a trapezium is formed with two vertical sides, one horizontal side, and one diagonal side. The calculation of moment of inertia of the trapezium T formed from segment $(x_{i,}y_i)$ to $(x_{i+1,}y_{i+1})$, about its centroid, can be reduced to calculating the sum of the moments for the triangle part (D) and the rectangle part (R) about their respective centroids, and then combining them appropriately. Mathematically, it can be expressed as:

$$I_T = I_D + I_R + a_D d_{DT}^2 + a_R d_{RT}^2 \tag{12}$$

$$a_T = a_D + a_R \tag{13}$$

$$x_D = \left(x_i + 2x_{i+1}\right)/3 \tag{14}$$

$$y_D = \left(2y_i + y_{i+1}\right)/3 \tag{15}$$

$$I_D = a_D\left[\left(x_{i+1} - x_i\right)^2 + \left(y_{i+1} - y_i\right)^2\right]/18 \tag{16}$$

$$x_R = \left(x_i + x_{i+1}\right)/2 \tag{17}$$

$$y_R = y_i / 2 \tag{18}$$

$$I_R = a_R\left[\left(x_{i+1} - x_i\right)^2 + y_i^{\,2}\right]/12 \tag{19}$$

$$x_T = \left(a_D x_D + a_R x_R\right)/a_T \tag{20}$$

$$y_T = \left(a_D y_D + a_R y_R\right)/a_T \tag{21}$$

where I denotes the moment of inertia of a shape about its centroid, a denotes the area of the shape, x and y denote the coordinates of the shape's centroid, and d denotes the distance between two centroids. The calculation of the moment of inertia and area for the whole polygonal area requires adding trapezia covering the inside of the polygon (such as T_1) and subtracting the trapezium pieces that lie outside of the polygon, such as T_2. Note that the calculations of centroids and moment of inertia in Equations (14)-(19) are for trapezia covering the inside of the polygon (such as T_1). For calculating trapeze lying outside of the polygon, the subscripts $_{i+1}$ and $_i$ should be exchanged.

$$I_T = I_{T_1} + I_{T_2} + a_{T_1} d_{T_1 T}^2 + a_{T_2} d_{T_2 T}^2 \tag{22}$$

$$a_T = a_{T_1} + a_{T_2} \tag{23}$$

$$x_T = \left(a_{T_1} x_{T_1} + a_{T_2} x_{T_2}\right)/a_T \tag{24}$$

$$y_T = \left(a_{T_1} y_{T_1} + a_{T_2} y_{T_2}\right)/a_T \tag{25}$$

In the computer implementation, a rule needs to be defined to determine which trapezia fall inside the polygon area and which fall outside. For a clockwise polygon, such as the one shown in Fig. 5, the trapezium should be added when $x_{i+1} \geq x_i$ and should be subtracted when $x_{i+1} < x_i$. For a counterclockwise polygon, the opposite applies. In this way, the moment of inertia of any regular polygon can be computed. In our study, the point arrays of each TAZ shape are read and the initial moment of inertia of TAZs is computed and obtained from the above equations. Meanwhile, Equations (22) - (25) can be used for calculating the variation of moment of inertia when attaching a TAZ to a model zone in the greedy phase and detaching a TAZ from a model zone in both greedy and edge-reassigning phases. Whenever the shape of a model zone is changed, the moment of inertia is recalculated. Then by substituting this value in Eq. (1), the shape index is obtained.

- **County boundaries and physiographic features**

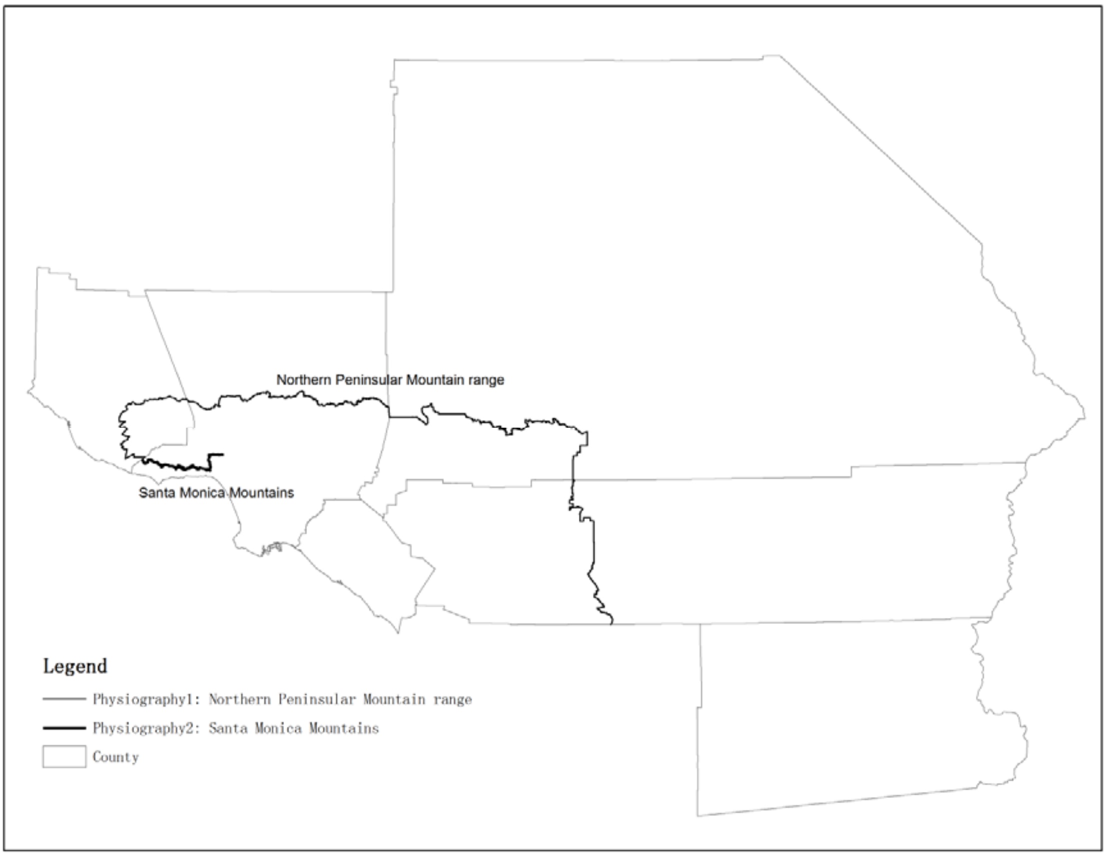

Figure 6. Physiographic features and county boundaries

One of the constraints in this optimization model is that the model zone cannot cross county boundaries or physiographic barriers. As shown in Fig. 6, TAZs are partitioned within counties, therefore, no single TAZ crosses a county boundary. To force model zones not to cross county boundaries, a topological analysis of (TAZ, County) containment was conducted and a hash-table was established with the ID of the TAZ as the key and the county that it belongs to as the value. Physiographic features act as natural boundaries for model zones as well, since it makes little sense that a local zone or gained for the purposes of modeling land use, economic, and transportation planning have a mountain range running through it as compared to forming a boundary on a side. In this project, two ranges of mountains in Southern California are identified as physiographic barriers: the Northern Peninsular Mountain range and the Santa Monica Mountains. The Northern Peninsular Mountain range crosses four counties and forms closed boundaries with county boundaries, dividing the four counties into sub-counties. Since neither county boundary nor physiographic boundary can be crossed, the values in the hash-table can be defined as sub-county codes instead of the original county codes. When a candidate TAZ is selected to be added to a model zone, the sub-county code will be fetched from the hash-table lookup. If this TAZ is not in the same sub-county as all the other TAZs in the model zone, this TAZ will be disregarded for consideration in adding to that zone.

The Santa Monica Mountain barrier requires a different approach since it does not form a closed boundary with county boundaries (specifically Los Angeles County). To incorporate this physiographic barrier into the zoning process, conflict groups were constructed of TAZs that cannot be assigned to the same model zone because they fall on opposite sides of the physiographic barrier. Each time a TAZ is considered for assignment to a model zone, the program will check whether it has conflict with any TAZ that has already been included in the model zone. The above strategy guarantees that both county boundary and physiographic feature constraints are satisfied.

- **Modeling zonal traffic**

To restrict the size of each model zone and to keep the errors in modeling transporation flows on the highway network small enough, the intra-zonal and inter-zonal traffic are considered. Define

MZ_k as the set of TAZs comprising model zone *k*. Let t_{ij} denote the traffic observed between TAZ *i* and TAZ *j* (we discuss the nature and source of such data below). Finally, let A_k represent the intra-zonal traffic in model zone *k*, B_k represent the inter-zonal traffic into and out of zone *k*, and α_{ik} represent the increase of intra-zonal traffic by adding TAZ *i* to zone *k*:

$$\mathrm{A}_k = \sum_{i \in MZ_k} \sum_{j \in MZ_k} t_{ij} \tag{26}$$

$$\mathrm{B}_k = \sum_{i \in MZ_k} \sum_{j \notin MZ_k} \left(t_{ij} + t_{ji} \right) \tag{27}$$

Our goal is to limit the ratio between intra-zonal trips and the sum of intra- and inter-zonal trips to be less than or equal to a threshold θ, that is:

$$A_k / \left(A_k + B_k \right) \leq \theta \tag{28}$$

The assumed value of θ is 0.1. When all the model zones reach this intra-zonal to inter-zonal traffic limit but there are still some TAZs unassigned, θ is automatically increased by 0.05. This operation is repeated until all the TAZs are assigned. To calculate variables A_k and B_k, a TAZ by TAZ origin-destination matrix containing the number of trips between each pair of TAZs was obtained from the Southern California Association of Governments.

Experiments

The above section discussed the computational issues of implementing the zoning process. In the next section, we describe several experiments to explore variations of these methods and to assess their effectiveness in solving the zoning problem.

- **Randomized greedy vs. greedy heuristics**

Fig. 7 shows the distribution of average compactness of the generated model zones over 333 runs. The difference of the average compactness is caused by the random choices of which model zone should grow. At each step, the candidate TAZ to attach to each model zone is randomly selected from the *N* best candidates (*N*=3) TAZs which most improve the compactness of the model zone. The model zone with the most potential improvement over the 100 model zones will be selected to grow at the step. Out of the 333 runs, in 171 (51.4%) the average compactness of the generated model zones is improved (closer to 1) compared to that generated from the non-randomized procedure The poorest average compactness (77.8) was obtained from the 94th run and the best (82.1) from the 264th run. Clearly multiple runs with the randomized greedy procedure are capable of yielding a solution that is closer to the global optimum.

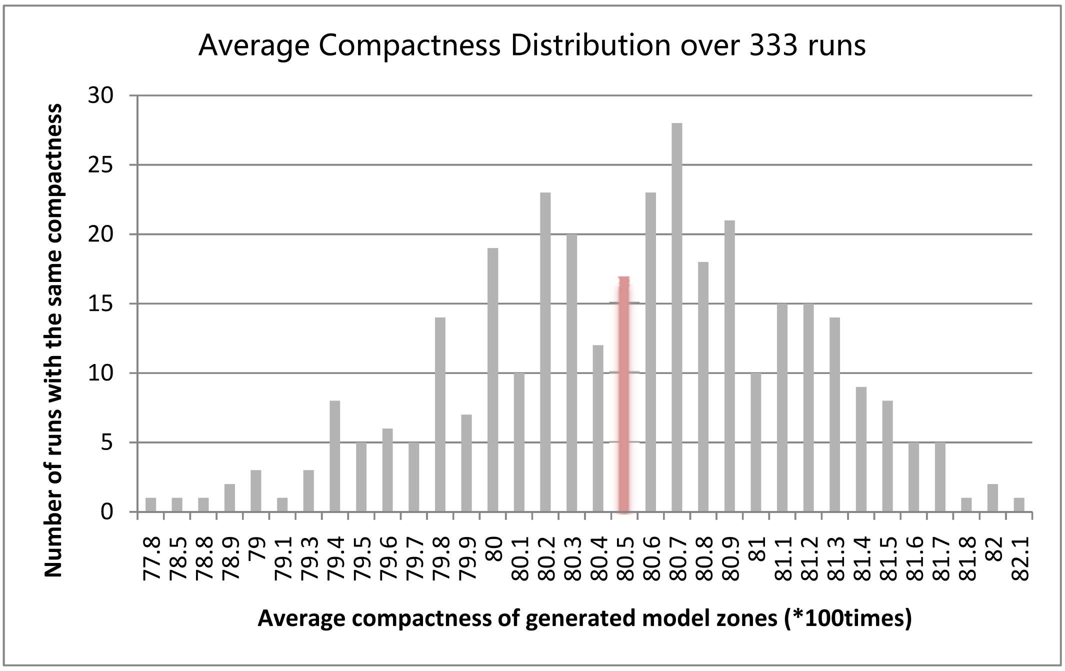

Figure 7. The distribution of average compactness for 333 runs of the randomized greedy heuristic

- **Comparison of compactness measures**

This section compares the average compactness of model zones using the moment of inertia introduced in this paper and that obtained using a popular compactness measure, the isoperimetric quotient (IPQ; Osserman, 1978). Mathematically, the objective changes to:

$$Max\sum_{k}\frac{4\pi A_k}{P_k^2} \tag{29}$$

where A_k is the area of MZ_k and P_k is its perimeter. In principle, the IPQ suffers from the well-known bias introduced by estimating the lengths of real geographic curves from the lengths of their polyline representations (Longley et al. 2010). Moreover the moment of inertia is directly related to the average accessibility of the entire area, rather than the geometry of its perimeter. The contrast between the two measures is most obvious when a roughly circular area that nevertheless has a long, undulating perimeter several times as long as the perimeter of a circle is compared to a long, thin area. The IPQ can return similar values in these two cases, whereas the moment of inertia will be very different, indicating the marked difference in accessibility and circularity of the two cases.

We ran the randomized greedy algorithm 333 times using the IPQ (Fig. 8) and moment of inertia (Fig. 9) as compactness measures from the same set of seed TAZs and compared the best result from the 333 runs. Intuitively, the model zones produced by the proposed compactness measure are much rounder in shape and insensitive to the detailed form of the edge. By calculating the shape index (using moment of inertia), the averaged compactness of model zones using IPQ is 0.711, which is 0.10 less than that obtained by using the proposed moment of inertia.

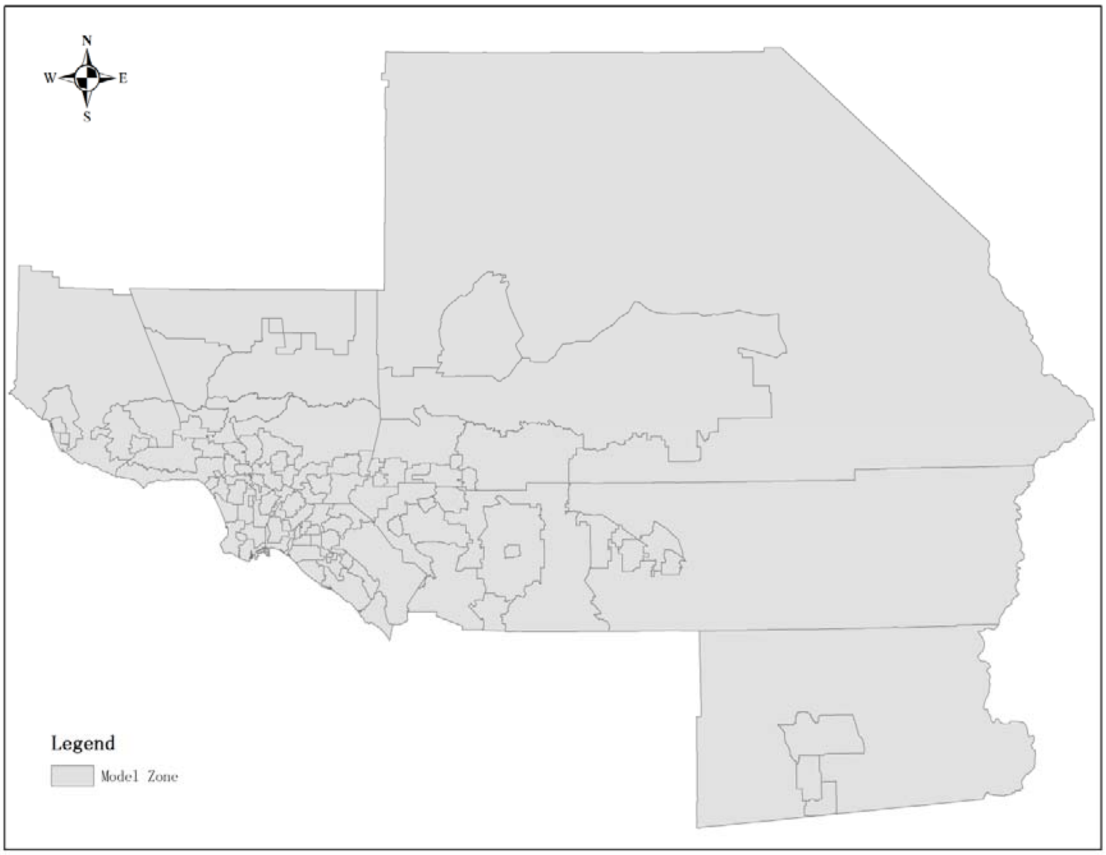

Figure 8. Zoning result using IPQ as the compactness measure

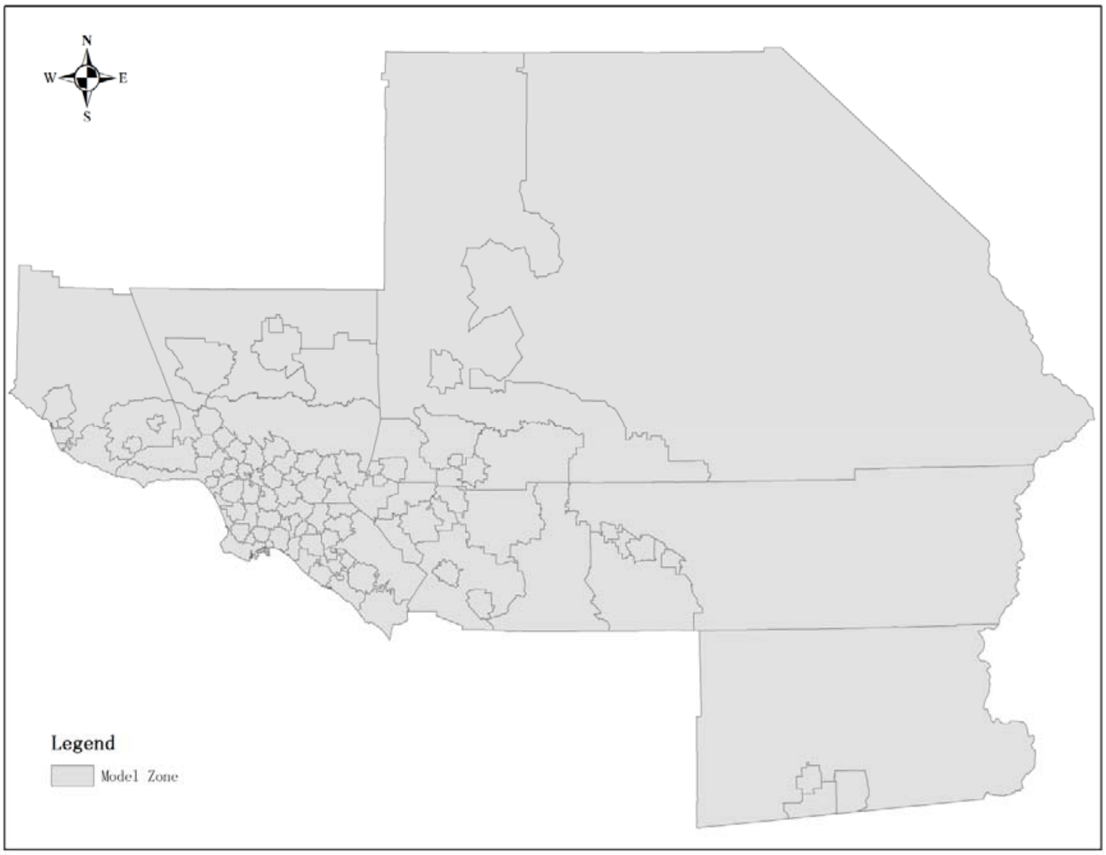

Figure 9. Zoning result using moment of inertia as the compactness measure

- **Improvement due to the edge-reassigning strategy**

We compared solutions with (Fig. 10) and without (Fig. 9) the edge-reassigning component of the heuristic. In 333 runs we found an improvement in average compactness of 8.7%, to 0.893. The maximum increase in compactness for a single model zone was 0.60, and the minimum -.06. We conclude that edge reassignment provides a substantial improvement in the heuristic.

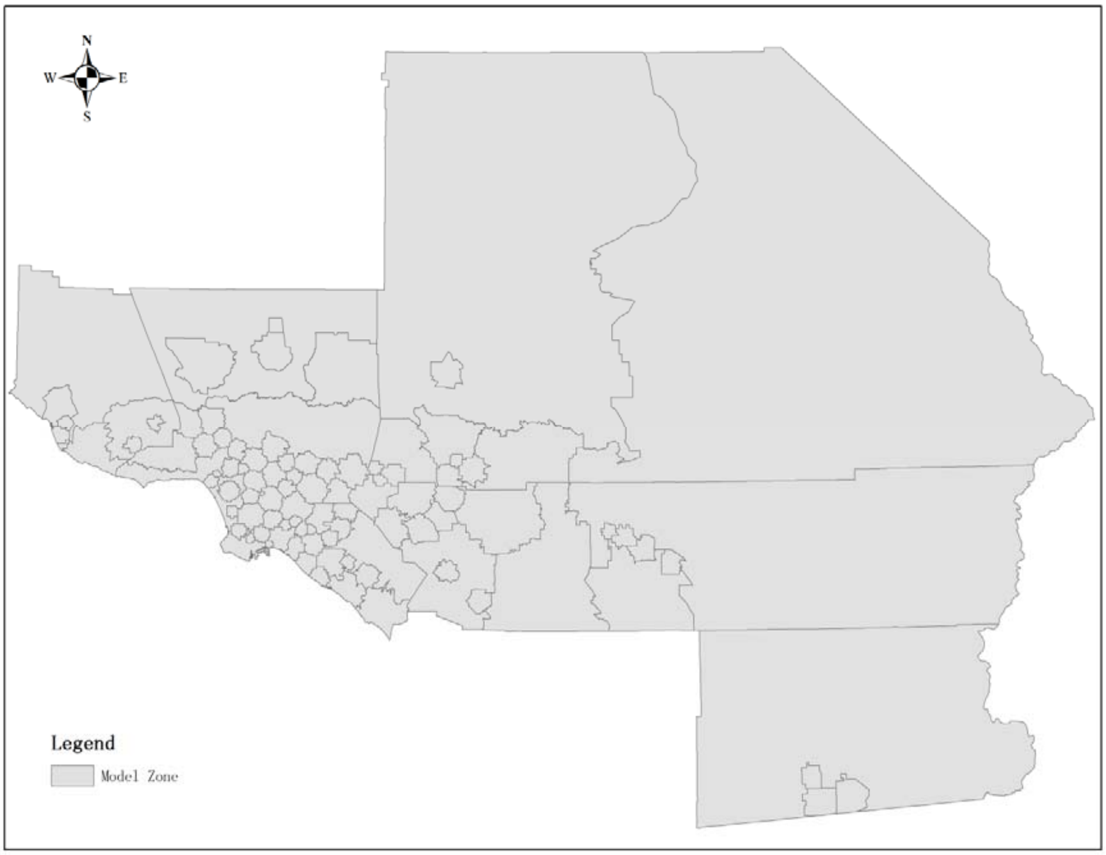

Figure 10. Edge-reassignment improvement to the result shown in Fig. 9

Conclusion and Discussion

This paper presents an approach to solving the *p*-compact-regions problem using heuristic methods. We introduced the model and the computational issues for objectives and constraints, and proved the advantages of the proposed randomized greedy and edge-reassignment algorithm through a number of experiments using a large and real case study motivated by the need to model the economy of six counties in Southern California.

Measuring compactness with the moment of inertia provides clear advantages over the more traditional IPQ. In the Section “Compactness measure” we provided the essential equations needed to compute the measure for aggregate areas, based on the properties of constituent building blocks. These equations demonstrate the additive nature of the measure, which makes it especially convenient for applications of this nature. We showed how the measure provides a substantial visual improvement in compactness for the case study as well.

As we noted at the outset, we do not believe a one-size-fits-all approach is suited to real-world applications of the *p*-compact-regions problem. Applications will vary in the number and nature of the constraints, the computational issues raised by the size of the problem, and the geography of the application area. Instead, we argue that by investigating a large, real problem we have been able to explore many of the issues that will arise, in varying combinations, in finding practical solutions to the problem.

More specifically, our approach requires the problem-solver to experiment with a number of parameters. The number K of building blocks added to each seed during the dealing phase of the algorithm is one such parameter; in this project we found that setting it so that approximately 25% of building blocks were dealt provided acceptable results. Similarly the number N of candidate building blocks maintained during the greedy phase, and randomly assigned, is under user control. We found that N=10 gave acceptable results. The number of runs of randomized assignment is also user-controlled, and in our case taking the best of 333 runs provided a substantial improvement. Finally, we defined a parameter θ to define the threshold of the traffic constraint, setting it to 0.10 and increasing it by increments of 0.05. Again, users might experiment with this parameter and its effects on the growth of model zones.

The RELU-TRAN project is still under development, this generation of model zones being one of the first stages in a multi-year project. At this time the modelers in the research team are satisfied with the model zones, finding on detailed inspection that they satisfy all of the requirements both in principle, in terms of the objectives and constraints, and in practice in terms of the detailed positioning of zone boundaries. The project will now proceed to the next steps of model calibration and application.

Acknowledgment

The application of RELU-TRAN to Southern California is funded by the University of California's Office of the President, and is being carried out as a collaboration between researchers at the University of California's Riverside and Santa Barbara campuses, and at the University at Buffalo. The cooperation of collaborators at Harvard University and the staff of the Southern California Association of Governments in providing access to data is also acknowledged.

References

Alsalloum O. L., and Rand G.K. (2006). "Extensions to Emergency Vehicle Location Models." *Computers and Operations Research* 33 (2006), pp. 2725-43.

Anas A., and Liu Y. (2007). "A Regional Economy, Land Use, and Transportation Model (RELU-TRAN): Formulation, Algorithm Design and Testing." *Journal of Regional Science* 47(3), 415-55.

Aboolian R., Sun Y., and Koehler G. J.(2009). "A Location-Allocation Problem for a Web Services Provider in a Competitive Market." *European Journal of Operational Research* 194(1), 64-77.

Altman M., and Mcdonald M. (2010). "The Promise and Perils of Computers in Redistricting." *Duke Journal of Constitutional Law & Public Policy* 5, 69-111.

Angel S., Parent J., and Civco D. (2010). "Ten Compactness Properties of Circles: Measuring Shape in Geography." *Canadian Geographer* 54(4), 441-61.

Brookes, C. J. (1997). "A Parameterized Region Growing Programme for Site Allocation on Raster Suitability Maps." *International Journal of Geographical Information Science* 11, 375-96.

Bozkaya B., Erkut E., and Laporte G. (2003). "A Tabu Search Heuristic and Adaptive Memory Procedure for Political Districting." *European Journal of Operational Research* 144(1), 12-26.

Church R. L., Storms D. M., and F. W. Davis (1996). "Reserve Selection as a Maximal Covering Location Problem." *Biological Conservation*, 76(2), 105-12.

Church R. L., Gerrard, R. A., Gilpin M., and Stine P. (2003). "Constructing Cell-Based Habitat Patches Useful in Conservation Planning." *Annals of the Assoication of American Geographers* 93(4), 814-27.

Church R. L., and Scaparra M. P. (2007). "Protecting Critical Assets: the r-Interaction Median Problem with Fortification." *Geographic Analysis* 39(2), 129-46.

Duque W.C., Church R. L., and Middleeton R. S. (2011). "The p-region Problem." *Geographic Analysis* 43, 104-126.

Fischer D., and Church R. L. (2003). "Clustering, and Compactness in Reserve Site Selection: an Extension of the Biodiversity Management Area Selection Model." *Forest Science* 49(4), 555-65.

Giuliano G., and Small K. A. (1991). "Subcenters in the Los Angeles Region." *Regional Science, and Urban Economics* 21(1991), 163-82.

Goodchild M. F. (1979). "The Aggregation Problem in Location Allocation." *Geographic Analysis* 11(3), 240-55.

Goodchild M. F. (1988). "Location-Allocation on a Microcomputer: The PLACE Package." In *Desktop Planning: Microcomputer Applications for Infrastructure, and Services Planning, and Management*, 166-71, edited by P.W. Newton, M. A. P. Taylor, and R. Sharpe. North Melbourne: Hargreen.

Longley, P. A., Goodchild M. F., Maguire. D. J., and Rhind D. W. (2010). *Geographic Information Systems, and Science*, 3rd ed. Hoboken, N.J.: Wiley.

Massam B. H., and Goodchild M. F. (1971). "Temporal Trends in the Spatial Organization of a Service Agency." *Canadian Geographer* 15, 193–206.

Morril R. L. (1976). "Redistricting Revisited." *Annals of Associations of American Geographers* 66 (1976), 548-66.

McDonald J. F. (1987). "The Identification of Urban Employment Subcenters." *Journal of Urban Economics* 21(2), 242-58.

Marianov V., and ReVelle C. (1996). "The Queuing Maximal Availability Location Problem: a Model for the Sitting of Emergency Vehicles." *European Journal of Operational Research* 93 (1996), 110–20.

Mehrotra A., Johnson E. L., Nemhauser G. L. (1998). "An Optimization based Heuristic for Political Districting." *Management Science* 44(8), 1100-14.

Murray A. T., O'Kelly M. E., and Church R. L. (2006). "Regional Service Coverage Modeling." *Computers & Operational Research*, 35(2), 339-55.

Moilanen A. (2007). "Landscape Zonation, Benefit Functions, and Target-based Planning: Unifying Reserve Selection Strategies." *Biological Conservation* 134(4), 571-79.

Openshaw S. (1977). "A Geographical Solution to Scale, and Aggregation Problems in Region-building, Partitioning, and Spatial Modeling." *Transactions of the Institute of British Geographers, New Series* 2, 169-84.

Osserman R. (1978). "The Isoperimetric Inequality." *Bullentin of American Mathematical Society* 84 (6), 1182–238.

O'Hanley J. R. , and Church R. L. (2010). "Designing Robust Coverage Network to Hedge Against Worst-case Facility Losses." *European Journal of Operational Research* 209(1), 23-36.

Pang S., He H., Li Y., Zhou T., and Xing K. (2010). "An Approach of Redistricting based on Simple, and Compactness." *Advanced in Swarm Intelligent* 6145 (2010), 415-24.

Ricca F., and Simeone B. (2008). "Local Search Algorithms for Political Districting." *European Journal of Operational Research* 189(3), 1409-26.

Sorensen P., and Church R. L. (2010). "Integrating Expected Coverage, and Local Reliability for Emergency Medical Services Location Problems." *Socio-Economic Planning Science* 44(1), 8-18.

Tobler W. (1970). "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46(2), 234-40.

Wie B., and Chai W. (2004). "An Intelligent GIS-based Spatial Zoning System with Multiobjective Hybrid Metaheuristic Method." In *Proceedings of the 17th International Conference on Innovations in Applied Artificial Intelligence*, 769-78, edited by B. Orchard, C. Yang, and A. Moonis. Berlin Heidelberg: Springer-Verlag.

Young H. P. (1988). "Measuring the Compactness of Legislative Districts." *Legislative Studies Qarterly* 13(1), 105-15.

Young C., Martin D., and Skinner C. (2009). "Geographically Intelligent Disclosure Control for Flexible Aggregation of Census Data." *International Journal of Geographic Information Science* 23(4), 457-82.