

A Cautionary Note on the Application of GIS in Spatial Optimization Modeling

Bin Zhou

Department of Geography & GIS, Southern Illinois University Edwardsville, Edwardsville, USA

Email: bzhou@siue.edu

How to cite this paper: Zhou, B. (2024) A Cautionary Note on the Application of GIS in Spatial Optimization Modeling. *Journal of Geographic Information System*, 16, 89-113.

<https://doi.org/10.4236/jgis.2024.161007>

Received: December 15, 2023

Accepted: February 25, 2024

Published: February 28, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Spatial optimization as part of spatial modeling has been facilitated significantly by integration with GIS techniques. However, for certain research topics, applying standard GIS techniques may create problems which require attention. This paper serves as a cautionary note to demonstrate two problems associated with applying GIS in spatial optimization, using a capacitated p-median facility location optimization problem as an example. The first problem involves errors in interpolating spatial variations of travel costs from using kriging, a common set of techniques for raster files. The second problem is inaccuracy in routing performed on a graph directly created from polyline shapefiles, a common vector file type. While revealing these problems, the paper also suggests remedies. Specifically, interpolation errors can be eliminated by using agent-based spatial modeling while the inaccuracy in routing can be improved through altering the graph topology by splitting the long edges of the shapefile. These issues suggest the need for caution in applying GIS in spatial optimization study.

Keywords

Spatial Optimization, GIS, Agent-Based Model, Covariance Function, Interpolation

1. Introduction

Spatial optimization modeling involves the maximization or minimization of an objective function (or a set of objective functions) while satisfying certain constraints. It is an important research subject in operation research and a specialty in geography [1]. Well-known spatial optimization topics include the facility location problem (*i.e.* location-allocation problem), routing, the transportation problem, etc. The last few decades have seen the application of Geograph-

ic Information Systems (GIS) techniques to the study of spatial optimization [2]. While the integration of the two fields has advanced spatial optimization study, the possible pitfalls have not drawn much attention. This paper serves as a cautionary note.

In this study, we use a real-world research example to demonstrate two problems in applying GIS techniques in spatial optimization. The first problem involves large errors in modeling spatial variations of travel costs when using kriging interpolation, a common technique used in generating raster files. Spatial interpolation involves using values at sampled sites to calculate values at unsampled sites. This process is based on a certain autocorrelation algorithm which derives global or local weights from the known values and assigns weighted values to the unknown sites. This is essentially a smoothing scheme which works well in continuous spaces but will generate large errors on spaces with unpredictable and/or discontinuous interruptions.

The second problem is inaccuracy in routing performed on a graph directly created from the polyline shapefile, a common vector file type. Routing refers to searching for the optimal travel paths with desired attributes such as the minimum travel distance or cost, or the fastest travel time or the highest safety. While traditional spatial optimization modeling often obtains optimal travel paths from a distance matrix containing distances among origins and destinations, some optimization studies include routing as part of the algorithm where the routing is performed on road networks. At the heart of creating a distance matrix and a road network is a graph which uses nodes and edges to represent locations and roads connecting locations. With GIS being integrated in spatial optimization, a graph can be created from a polyline shapefile. This significantly changes the topology of a graph. Instead of one edge connecting an origin node and a destination node, there may be many edges in between. This occurs since a polyline shapefile uses separate sub-edges to create a line so that its geometry (*i.e.* the line curvature) resembles that of a road. On each end of a sub-edge is a new node. Thus, a graph created from a polyline shapefile contains more nodes and edges than a graph which only contains origins and destinations as nodes and roads connecting them as edges. The resultant networks, distance matrices, and routing for optimal travel paths may also differ.

To demonstrate the above two problems, we solve a capacitated p -median location-allocation optimization problem as an example. While the paper also suggests remedies, these problems themselves are a reminder of the pitfalls when uncritically adopting GIS techniques and applying them to topics with unique characteristics.

The remainder of the paper is organized as follows. The second section briefly reviews the literature involving the location-allocation problem, agent-based modeling, and the application of GIS in both. The third section discusses the research method and workflow. The fourth section presents the results of modeling, along with detailed discussions of these two problems based on the results. The last section summarizes and concludes the study.

2. Literature Review

In this section, rather than conducting a comprehensive review of all topics involved, we mainly highlight the core issues this paper intends to address. The location-allocation problem locates business or public facilities and allocates customers to them. In the early 1900s, Weber set the problem in the context of a continuous space which had been operationalized by [3] [4]. In subsequent years, the facility location problem has been solved on networks [5] [6] [7]. Most facility location problems involve linear/integer programming. One type of location-allocation problem is the p-median problem which locates a p number of facilities to serve customers distributed at certain locations. The objective is to minimize the total or average shipping cost [8] [9]. An example of the p-median problem is to set up the optimal locations of fire department units to cover different sub-areas within a community so that fire trucks can reach the fire sites as quickly as possible. Since fire events are rare and it is not often that all sub-areas simultaneously experience fire events, a p-median problem may suffice. For a problem where the optimal locations of hospitals or state vehicle offices are involved, the service capacity at each site should be taken into consideration to avoid long lines or delayed service from having the number of customers more than a hospital or an office can efficiently handle. This is the case of the capacitated p-median problem.

GIS techniques enhance solving facility location problems using spatial data processing tools such as data storage, queries, interpolation, and visualization [10] [11]. More importantly, overlaying features layers, identifying finite dominate sets in a network, and skeletonizing polygons play key roles in transforming continuous location problems into discrete location problems [12]. While it is common to couple location search and spatial data processing techniques, the integration of GIS and spatial optimization aims at creating programs with built-in GIS and combinatorial computational capabilities [13].

In standard location-allocation studies, the outcome is often a set of feasible sites where the optimization criteria are met. Typical visual representations are desire lines connecting the facility locations with demand locations, as shown in **Figure 1**. One limitation of using desire lines is that they show a few optimum sites in isolation due to a lack of attention to spatial cost variations in the entire study area. [14] break new ground by bootstrapping a stochastic spatial point process while solving for optimization. Their approach generates contours with declining probability densities from the identified optimal location.

[14] essentially interpolate the spatial variation of opportunities in the facility location problem. This is where GIS can contribute to spatial optimization study since spatial interpolations are standard techniques in modeling spatial variations. At the heart of spatial interpolation is a certain algorithm which specifies the pattern of spatial autocorrelation. In kriging interpolation, this is the covariance function or kernel. Well-known spatial covariances are Matérn, Cauchy, Gaussian, Spherical, etc. (**Figure 2**). The pattern of each function in **Figure 2**

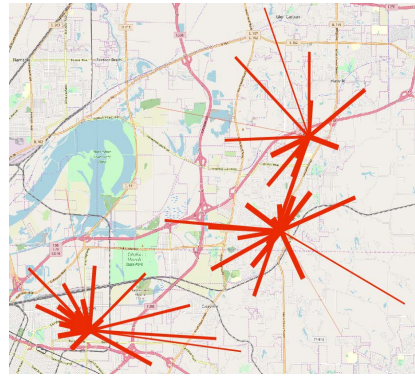


Figure 1. Desire lines. This map shows a standard display of a facility location solution which shows three facilities each connecting to several demand points. The width of the desire lines indicates supply quantities.

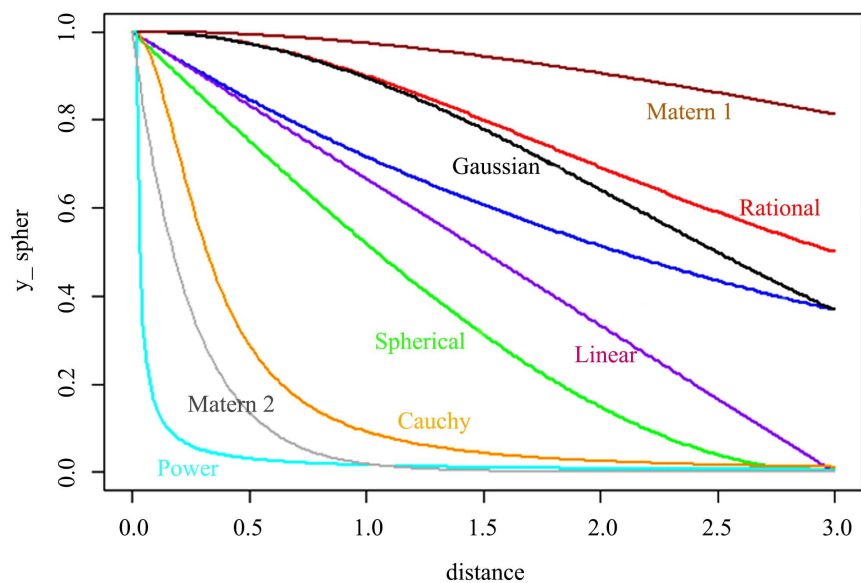


Figure 2. Various spatial kernels can depict patterns of distance decay. The pattern from the same covariance function can vary when the parameters involved change, as seen in Matérn 1 and Matérn 2 where a smoothness parameter is set differently.

can change when relevant parameters are set differently. For example, a smoothness parameter in Matérn 2 is set much higher than in Matérn 1, leading to two different distance decay patterns. However, spatial covariance functions are only our best guesses about the pattern of spatial autocorrelation. They may generate satisfactory approximations in smoothly changing continuous spaces. When the landscape is much more disruptive and complicated than the spatial covariance function used, interpolation errors are inevitable.

The last few decades have seen the rise of agent-based modeling (ABM) in spatial analysis, including the location-allocation problem [15]. More recent contributions include a study [16] which solves a model of charging station locations for electric vehicles. [17] uses a two-stage model to determine the optimal sites to set up printers on a university campus, minimizing the students' travel

distance to the printing service. A study in [18] optimizes biophysically optimal land-use and then compares it with results from an ABM to determine better solutions. In yet another study [19], authors validate the retail distribution pattern obtained from an ABM with that from conventional location-allocation modeling. [20] investigate the optimal vaccination center locations during the COVID-19 Pandemic in New Jersey, USA.

The rising popularity of ABM in spatial modeling and optimization goes hand in hand with the integration of GIS [21] [22]. Most major ABM software programs today (Repast, NetLogo, MASON, GAMA Platform, etc.) have GIS functionalities, including the use of polyline shapefiles to create graphs for routing and transportation assignment studies. Polyline shapefiles use nodes and edges and their arrangements to represent the road topology. The resultant graph should understandably have significant impacts on the outcome of routing. While modifying line features has been a common technique in typical GIS software, the purpose of doing so is mostly confined to creating connected planar networks or appropriate survey line segments [23]. It is surprising that there has been little discussion on how changing graph topology would affect the spatial optimization modeling.

In brief, GIS techniques have become integrated components in both operation research and agent-based modeling. There have been review papers that discuss the state and directions of spatial optimization and spatial agent-based modeling [2] [24]. However, attention to the pitfalls of applying GIS in location-allocation optimization study, either through the operation research methods or ABM, has been lacking.

3. The Method

As stated earlier, this study shows two problems in applying GIS in spatial optimization and suggesting remedies. To achieve this, we will solve a capacitated p -median location-allocation optimization problem and then use the solution to model spatial variations of the travel cost.

The location-allocation optimization problem presented in this paper is based on a real-world problem of setting up four electronic recycling item collection centers to serve 34 local communities in Madison County in southwestern Illinois, USA (Figure 3). Customers from each community can bring in their electronic items to recycle twice a year on given days. The demand from each community is estimated proportionally to their population. On the supply side, four collection centers are sought, each with a capacity based on the number of volunteers available within the community. In other words, both demand quantity and supply capacity vary at the community level. The modeling comes in two stages: modeling location-allocation and simulating spatial variations.

Stage 1. Modeling location-allocation. In this stage, a standard location-allocation problem is solved. Specifically, for a graph $G(V, E)$, a subset of vertices $I \in V$ are defined as the demand points, which implies that demand locations are on a road network. As in a standard facility location problem, this can be expressed as

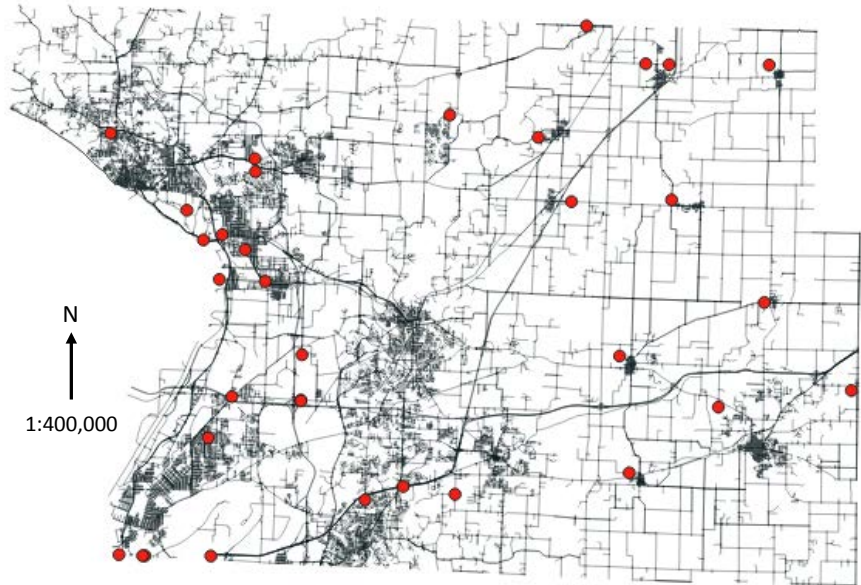


Figure 3. All roads in Madison County, Illinois, USA. The 34 places are used as the demand locations for four electronic recycling items pickup sites.

$$\min Z = \sum_{i \in I} \sum_{j \in V} d_{ij} w_{ij} \tag{1}$$

subject to

$$\sum_{j \in V} w_{ij} = 1 \quad i \in I \tag{2}$$

$$\sum_{i \in I} q_i w_{ij} \leq Q_{y_j} \quad j \in V \tag{3}$$

$$\sum_{j \in V} y_j = p \quad j \in V \tag{4}$$

$$x_{ij}, y_j \in \{0,1\}, w_{ij} \in \{0,1\} \quad i \in I, j \in V \tag{5}$$

where Z is the total distance-weighted travel cost; d_{ij} is the distance on the network between demand vertex i and supply vertex j ; w_{ij} is an allocating variable equal to 1 if demand at i is allocated to j or 0 if not; q_i is the quantity of demand at i ; y_j is a locating variable equal to 1 if a facility is located at j or 0 if not; Q_{y_j} is the facility capacity at j ; and p is the number of facilities. Equation (2) guarantees each demand point is allocated to only one facility. Equation (3) makes sure the total demand allocated to j is within the capacity at j ; and Equation (4) sets the total number of facilities to p . Equations from (1) through (4) show the typical representation and solution for a capacitated p -median process, as presented in [8] [9]. In this paper, the search is performed using an algorithm akin to the tabu algorithm [25].

Stage 2. Simulating spatial variations. In Stage 2, we design the following procedure to simulate the spatial travel cost variations based on results from Stage 1.

The solution from Stage 1 identifies the p (4) number of optimal supply vertices v_p , each serving a set of demand vertices v_t , ($t = 34$) This location-allocation pattern can be expressed as

$$v_p \Rightarrow v_t, p \in (1, 2, 3, 4), t \in (1, 2, \dots, 34) \quad (6)$$

Between p and t there can be $v_u = v_p \cap v_t$ meaning a node can be both a demand and supply location at the same time based on [26].

Given the location-allocation solution in (6), values of the assignment variable w_{ij} are determined between pairs of i and j , and only those conforming to $w_{ij} = 1$ remains in (1). Z can be written into four components each corresponding to one collection facility and the demand vertices it serves. That is,

$$Z = \sum_{i \in I} \sum_{j \in V} d_{ij} w_{ij} = \sum_{i \in I_1} d_{i1} + \sum_{i \in I_2} d_{i2} + \sum_{i \in I_3} d_{i3} + \sum_{i \in I_4} d_{i4} \quad (7)$$

In (7), the four components correspond to the costs of four supply points serving those demand points allocated to them, respectively. Relabeling each component, we have

$$Z = Z_1 + Z_2 + Z_3 + Z_4 \quad (8)$$

In Stage 2, since demand point allocation (*i.e.*, which demand point is served by which facility) is known, the network travel cost (distance) to them from all locations, all nodes, and all links can be calculated. Take Z_1 in (8), which is the cost of the first supply point $v_{p=1}$ serving an n number of demand points allocated to it $v_{t=1} v_{t=2} \dots v_{t=n}$ ($n < 34$). We simulate the cost of serving the n demand points from every location in the study area. This is as if the first facility $v_{p=1}$ takes turns to be at every location. In this study, there are 14,847 locations or patches within the NetLogo space or grids within the GAMA Platform environment (a patch in NetLogo or a grid in GAMA Platform is like a pixel in raster). At any pixel, s , the first facility serves demand vertices $v_{t=1} v_{t=2} \dots v_{t=n}$ while the other three facilities remain at their optimal vertices v_p , $p \in (2, 3, 4)$, the total distance cost is

$$Z_s = \sum_{i \in I_1} d_{is} + Z_2 + Z_3 + Z_4 \quad s \in (1, \dots, 14847) \quad (9)$$

where $\sum_{i \in I_1} d_{is}$, $s \in (1, \dots, 14847)$ denotes the spatial cost variation of the first facility serving the demand vertices $v_{t=1} v_{t=2} \dots v_{t=n}$, and Z_s the total cost of serving all demand vertices by all facilities. At the location where a patch geographically coincides with the first facility's optimal location from the model $v_{p=1}$, $Z_s = Z$, where Z_s is the minimum total travel cost from the model solution.

Using a similar procedure, simulations can be performed for the first facility to vary its location among all nodes (V) or all links (E) on the road network. To avoid cluttering, the paper will not introduce new equations. Since there are 1667 (V) nodes and 1780 links (E) on the road network used in the study, Equation (9) represents the cost when the first facility varies its location among nodes with s being a node, $s \in (1, \dots, 1667)$ and, among links with s being a link, $s \in (1, \dots, 1780)$. The above simulation procedures can similarly be performed for the other three groups of demand vertices, giving spatial cost variations for all facilities.

The Key Technical Details of Model Execution. The modeling is carried out

using NetLogo. To perform routing, a line shapefile is used to create a graph. While NetLogo can import a line shapefile through its GIS Extension, the roads created are just static images (a patch attribute) and do not function as a network of nodes and links. Instead, in this study, we first extract node and edge attributes from the road shapefile and use them to create a graphml network in the Visone software. The Network Extension of NetLogo has the capability to import graphml networks where nodes turn into node turtles and edges into links within NetLogo. These nodes turtles and links form a functioning network.

One problem is that NetLogo cannot handle very large shapefiles. The shapefile in **Figure 3** contains all roads in Madison County and needs to be significantly simplified so that the program can perform routing efficiently. We convert it into a minimum spanning tree polyline shapefile (a network connecting all nodes with the minimum total length), as seen in **Figure 4**. With some additional important improvements as discussed later in the next section, this simplified line file is used in the modeling.

During the modeling stage which searches for the optimal location set, the NetLogo Network Extension is used to move consumer turtles about the network, and their travel distances to alternative supply nodes are calculated using the NetLogo Matrix Extension and Array Extension. Travel data are recorded in NetLogo lists. In addition, similar to performing statistical operations on a raster file, attributes of patches can be recorded and statistical analyses performed on them.

One critical element in this study is the interaction between patches and turtles. When simulating travel cost from every location (*i.e.*, patch) in the study area, every patch “sprouts” (creates) a firm turtle. These firm turtles “wiggle” (move about but not on networks) to the nearest node on the network, from which they move further to demand vertices via the network. The travel data is recorded by each firm turtle and then passed back on to the patch which sprouts

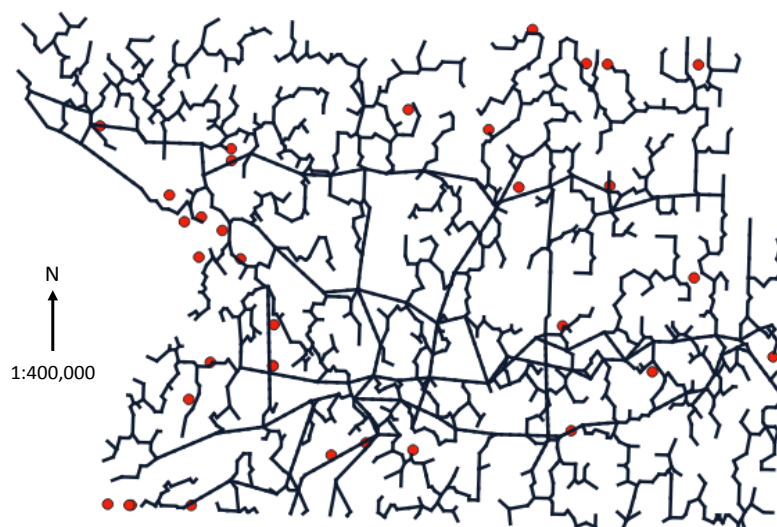


Figure 4. The road system in Madison County, IL, USA as a minimum spanning tree which allows NetLogo to perform routing efficiently.

it. This approximates the travel distance from each patch (as a supply site) to demand vertices via the road network. The patch level travel data is then used to create spatial cost variation at all locations in the study area.

In post-simulation analysis, data (including imagery data) recorded by patches, nodes, and links can be exported and further processed for kriging in standard GIS software or for statistical analysis in other software.

To focus on the main issues in the modeling, we assume away complicating factors such as road quality, road type, speed limits, one-way or two-way driving roads, gasoline prices, and business set-up costs. Bringing in these factors would make a facility location study realistic but would not allow us to distinguish whether the spatial cost variations occur due to the choice of supply locations (which is the focus of the study) or other factors.

4. Results

In this section, we will first present the results of optimization of the capacitated facility location problem, and the simulation of spatial variations of travel cost. Then we will discuss in detail the two problems based on the results.

4.1. Results of the Optimization and Simulation

Stage 1: Modeling the location-allocation. A capacitated p -median electronic recycling item collection model, as described earlier, is run with 1000 iterations in NetLogo. Here “1000 iterations” means that the model stops when there is no improvement within 1000 runs. If the model improves before 1000 runs are over, the tick is set back to 1 and the count starts over. In the end, the model runs 1188 iterations, with 365 iterations generated that meet the capacity constraints. The optimal value of Z is 518 km. The map of desire lines from the solution is shown in **Figure 5**, with four collection centers having arrows pointing to communities they serve.

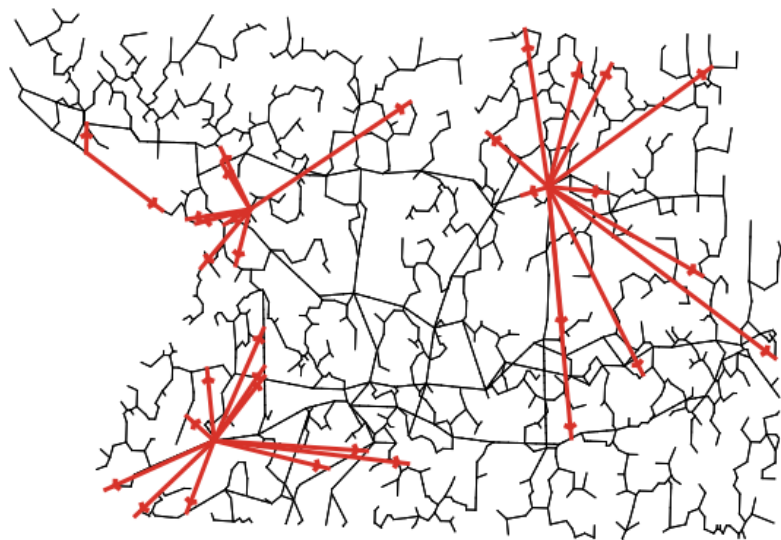


Figure 5. Desire lines from the solution of the capacitated $p = 4$ facility location model.

Stage 2: Simulating spatial variations. Following Equation (9), spatial variations of the four components of Z are simulated and mapped (**Figures 6(a)-(l)**). The maps show distance variations of four facilities in the study area at the pixel, node, or edge level. Take Facility 1 as an example to interpret the maps. **Figure 6(a)** shows that to serve the 12 communities (11 locations at arrows' points and Livingston where the facility is located) the total travel distance ranges from 692 km (lightest shade of blue) to 2990 km (the darkest shade of blue). It would be 692 km if the facility were located at Livingston but up to 2990 km if the facility were located in areas with the darkest shade of blue, such as in the northwestern or southwestern corner of the study area. Other shades of blue show the intermediate total travel distance when serving the 12 communities.

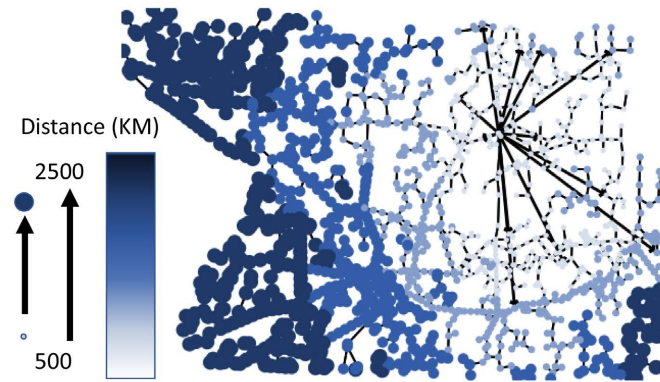
Figure 6(b) shows what the total travel distance would be to serve the 12 communities if the facility were located at different nodes on the roads. The size and shade of the color indicate the total travel distance. Similarly, **Figure 6(c)** shows the total travel distance if the facility were to locate at different edges (mid-point of an edge) on the road. **Figures 6(d)-(l)** show similar information for the other three facilities. The total travel distance may be very different from those for Facility 1 since the number of communities served may be different. For example, in **Figures 6(d)-(f)**, the total travel distance is low since the facility serves only two communities. As discussed earlier, these maps can be interpreted as the rising total travel distance due to shifting the facility away from a solution location while keeping other facility locations unchanged.

The spatial travel distance variation at the pixel level can also be mapped by widely understood contour lines with color fill, as in **Figure 7** which shows the travel distance variation if Facility 1 were to locate at various pixels in the study area.

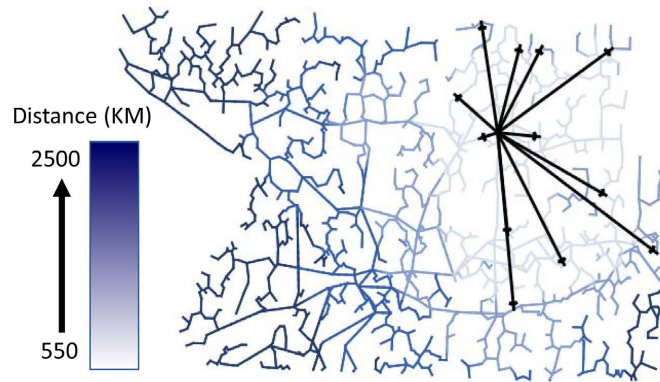
While **Figure 6** and **Figure 7** highlight how spatial simulation works as an alternative to interpolation in modeling spatial variations, there are additional significant points to make in the context of the location-allocation problem. First, maps in **Figure 6** and **Figure 7** give measurable information, compared with desire lines in the standard representation of location-allocation solutions as shown in **Figure 1**. As stated earlier, standard location-allocation modeling looks for a few optimum sites in isolation and does not reveal how much better these sites are than other sites, in any direction, or on and off the roads, in the study area. Furthermore, locational decision is only part of the overall decision-making which often involves trade-offs with other business and public goals. Without measurable comparison among alternative sites and quantifiable data to assess trade-offs between the location decision and other goals, it would be truly difficult to determine whether the business or public interest is best served at these optimal sites derived from modeling. While a full assessment of the facility location modeling is beyond the scope of this paper, the results discussed so far demonstrate the potential to expand the location-allocation modeling.



(a)



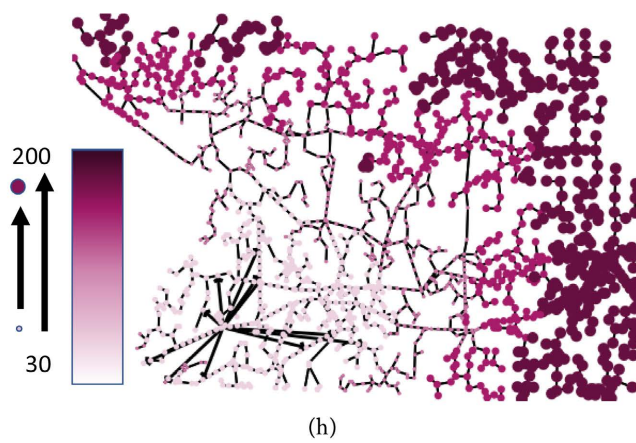
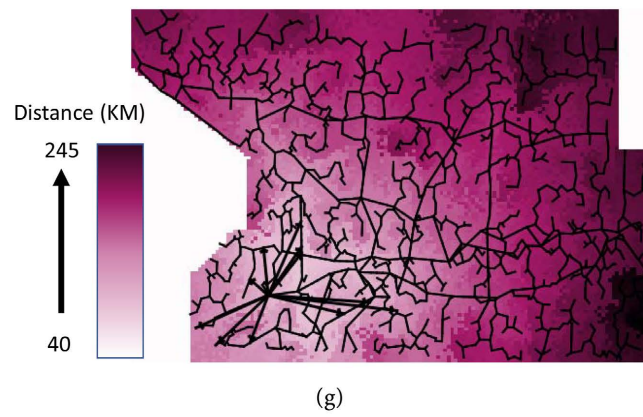
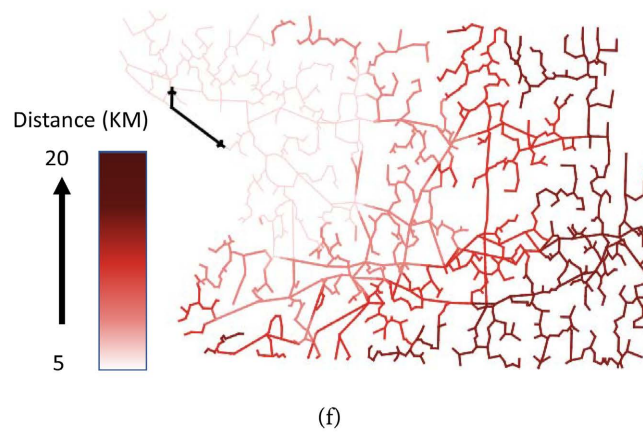
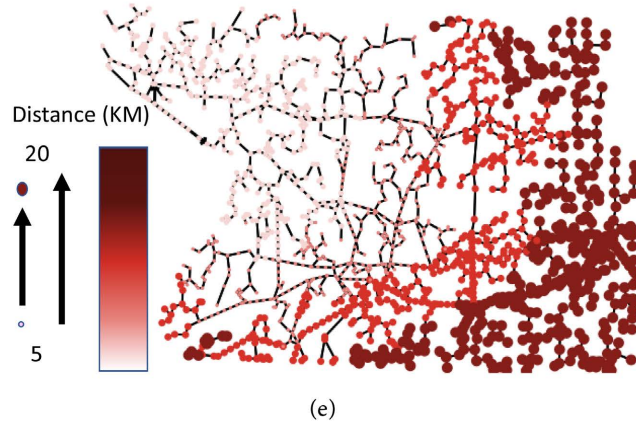
(b)



(c)



(d)



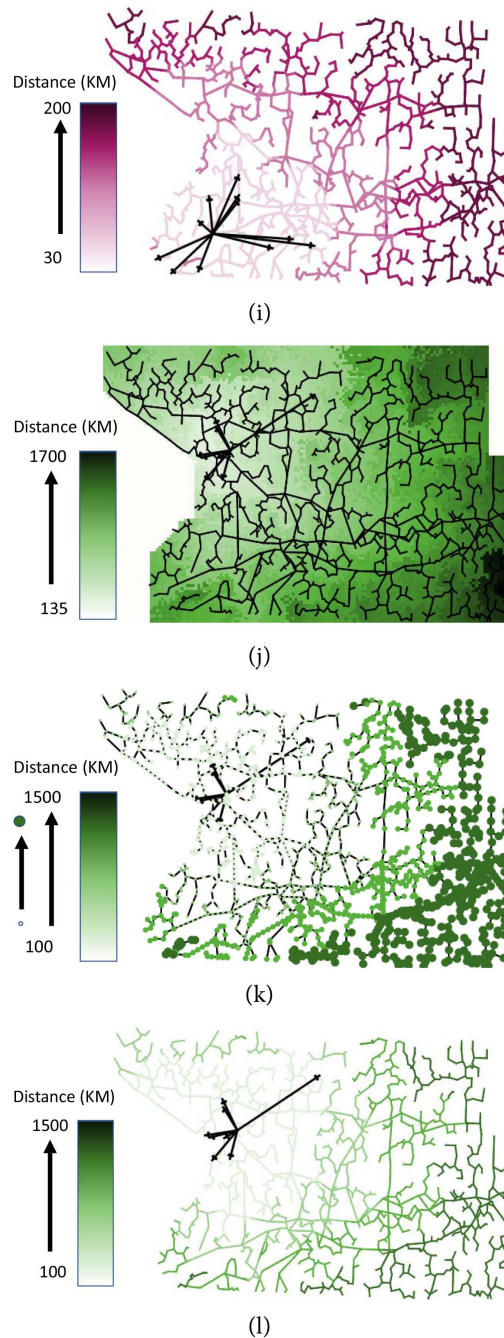


Figure 6. (a) Spatial variation of travel distance to Facility 1 at the pixel level where the closest node is Livingston; (b) Spatial variation of travel distance to Facility 1 at the node level. The node is Livingston; (c) Spatial variation of travel distance to Facility 1 at the edge (edge midpoint) level where the closest node is Livingston; (d) Spatial variation of travel distance to Facility 2 at the pixel level; (e) Spatial variation of travel distance to Facility 2 at the node level; (f) Spatial variation of travel distance to Facility 2 at the edge (edge midpoint) level; (g) Spatial variation of travel distance to Facility 3 at the pixel level; (h) Spatial variation of travel distance to Facility 3 at the node level; (i) Spatial variation of travel distance to Facility 3 at the edge (edge midpoint) level; (j) Spatial variation of travel distance to Facility 3 at the edge (edge midpoint) level; (k) Spatial variation of travel distance to Facility 4 at the pixel level; (l) Spatial variation of travel distance to Facility 4 at the node level; (m) Spatial variation of travel distance to Facility 4 at the edge (edge midpoint) level.

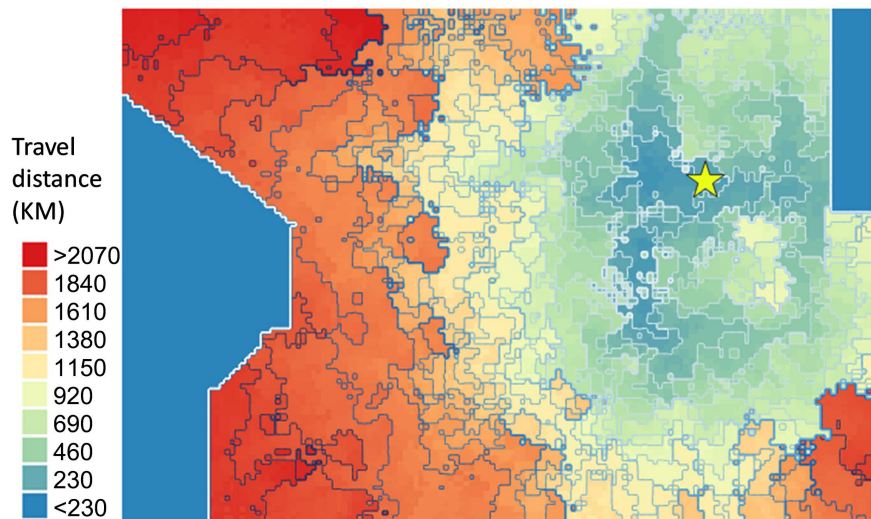


Figure 7. The customers' travel distances to get service from Facility 1 if the facility is located at different locations. The star point is the optimal facility site.

4.2. Spatial Interpolation Errors

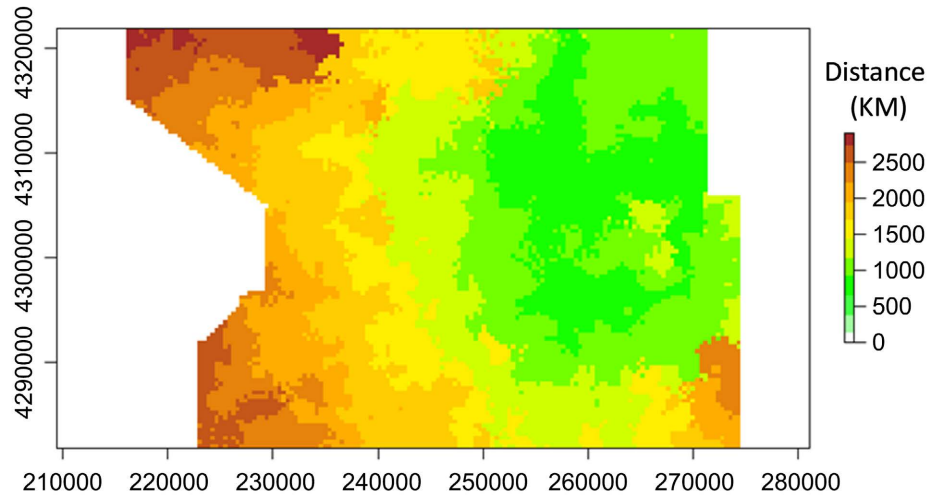
To assess interpolation errors in modeling spatial variation, we re-create **Figure 6(a)** using a GIS program as shown in **Figure 8(a)** without roads and desire lines. We take 30% of the pixel values from **Figure 8(a)** as sampled sites to perform kriging interpolation in order to assess interpolation errors in re-creating it. **Figure 8(b)** displays a result using ordinary kriging. **Figure 8(a)** and **Figure 8(b)** seem to bear significant resemblances.

To avoid bias from using only one random sample, 100 random samples are drawn, and interpolations performed. Thus, there are 100 individual interpolated maps like the one shown in **Figure 8(b)**. Instead of showing all individual maps, we created a composite interpolated map by averaging each pixel values over the 100 interpolations and mapped in **Figure 8(c)**, which still seems to show a strong visual resemblance to the original map in **Figure 8(a)**.

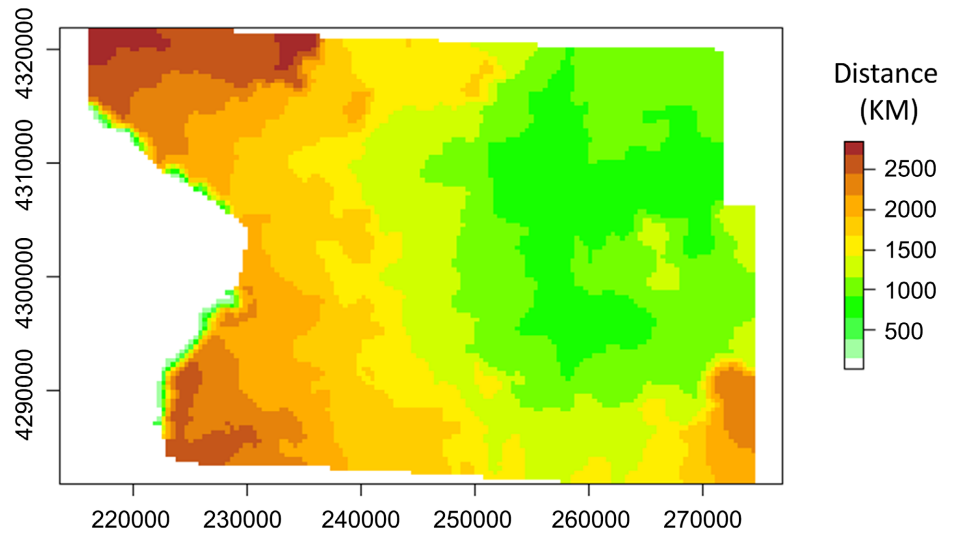
To assess interpolation errors, we used the normalized root-mean-square-errors (or normalized RMSE) calculated in (10).

$$\text{NRMSE} = \frac{\sqrt{\sum_{1 \text{ to } M} (\text{interpolated pixel value} - \text{original pixel value})^2 / M}}{\text{Mean over original pixel values}} \times 100 \quad (10)$$

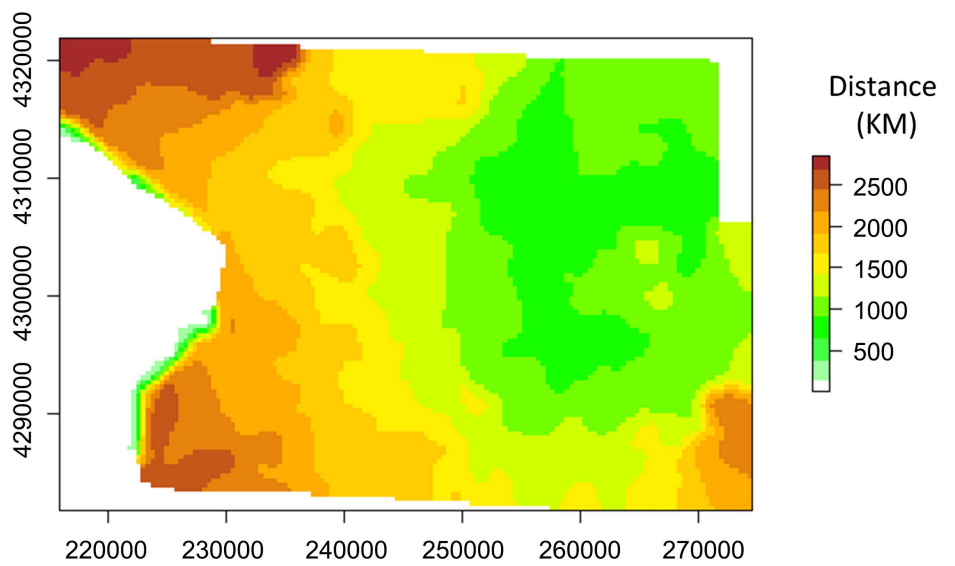
In (10), the *interpolated pixel value* is a pixel value obtained from an interpolation; the *original pixel value* is taken from **Figure 8(a)**, which is used as the interpolating target. The squared deviation between the interpolated and original values is summed over M . Since we perform interpolations 100 times, $M = 100$. The numerator of (10) is the RMSE (parallel to standard deviation except the original pixel value may vary at each pixel). Equation 10 normalizes RMSE by the overall mean value of the pixels in **Figure 8(a)**. Normalized RMSE expresses the deviation of interpolated value from the target, expressed as a percentage of the overall mean pixel value. **Figure 8(d)** shows the distribution of



(a)



(b)



(c)

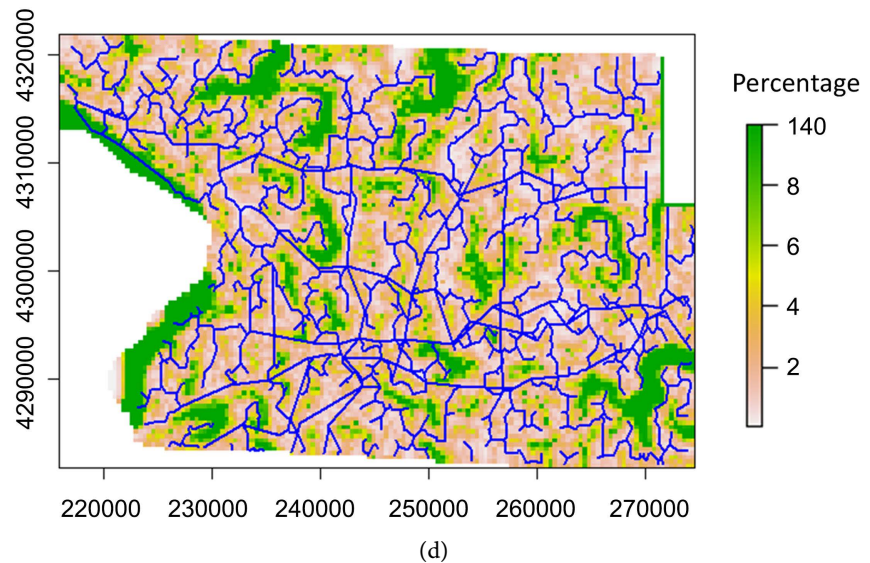


Figure 8. (a) Travel distance variations for Facility 1. The map is an alternative version of **Figure 6(a)** without desire lines and roads and generated from a GIS program. (b) A map generated through ordinary kriging interpolation, using 30%-pixel values randomly taken from (a) as sampled sites. (c) A composite map created by averaging 100 interpolated values for each pixel. In other words, values at each pixel are average travel distance over 100 interpolations for that pixel. (d) Normalized RMSE at the pixel level overlaid with roads. For each pixel, normalized RMSE is calculated from 100 interpolation.

normalized RMSE at the pixel level. The actual range of normalized RMSE is from 0% and all the way up to 140%. To make it easy to identify locations on the map, we rescaled the color scheme by using the same green shade to represent a wide value range from 8% to 140%. As can be seen, many areas have significant levels of interpolation errors. The largest interpolation errors tend to occur in areas between roads.

To obtain the expected interpolation error, we calculate normalized RMSE for each interpolated map. In this case, $M = 14,847$ (the number of pixels in each map). The 100 normalized RMSEs constitute an empirical sampling distribution of sample normalized RMSE, as shown in **Figure 9**. The expected normalized RMSE is 11.33%, with a 95% confidence interval between 10.49% and 12.17%. This is a significant overall level of interpolation errors.

The significant interpolations errors as shown above are obtained regardless of the choice of kriging techniques (simple, ordinary, IDW, etc.). Although there may be room to calibrate interpolation by choosing more sophisticated methods and/or setting parameters differently, it is unlikely to fundamentally alter the general pattern.

To assess the covariance structure which should be used in interpolation, pixel values from **Figure 8(a)** are divided into four quartiles to test for the theoretical (or optimal) covariance structure within each. For values within the first quartile, largely corresponding to the area close to the optimal site (areas with light shading), the optimal structure is the Matérn covariance; for values within the next three quartiles, the optimal covariances are Matérn, Spherical, and Matérn,

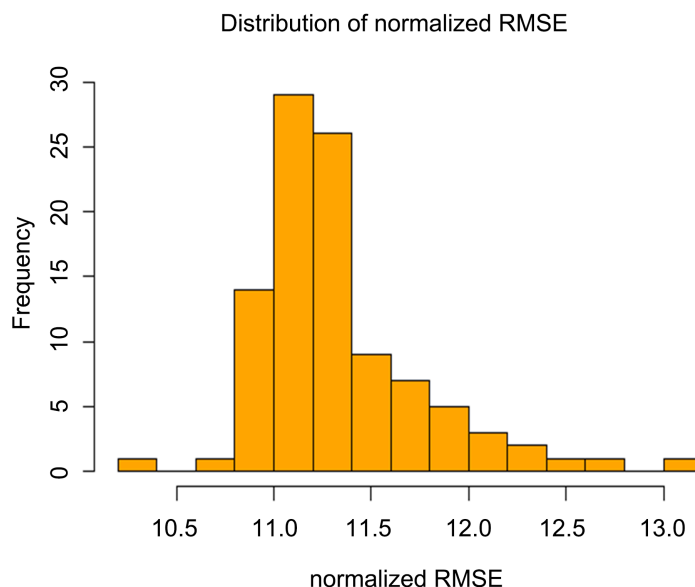


Figure 9. Empirical sampling distribution of sample normalized RMSE from 100 interpolations.

respectively. While 3 out of the 4 groups are optimized by a Matérn function, their ranges (distance within which spatial autocorrelation is significant) and partial sills (measure of spatial autocorrelation) are all different indicating diverse spatial autocorrelation relationships.

When dividing the study area into four spatial quadrants, the optimum theoretical covariance is the Matérn covariance for all four quadrants. However, the northwestern and southwestern quadrants have the same range and partial sills. Their range and partial sills are also the highest. In contrast, the southeastern quadrant has the least range and partial sills. The northeastern quadrant has an intermediate range and partial sills. It is interesting to note that for Facility 1, the northwestern and southwestern quadrants have quite similar travel distance ranges compared with the east. It is not surprising to find that these two quadrants have the same semi-variogram profile. However, for the entire area, the suggested covariance function is Matérn but with partial sills and range smaller than all those in four quadrants separately. The overall spatial autocorrelation is an average of its spatial components. This speaks to the complex and unpredictable nature of the spatial autocorrelation at different spatial levels. Thus, using a uniform or a combination of spatial covariance functions to simulate a complex spatial variation is bound to create significant errors.

The above experiments highlight the problems in using a covariance structure to approximate spatial patterns under certain circumstances. Using the Object View and Field View division [27], the field view sees the world as a continuous space with varying values representing fields of temperature, wind speed, heights, etc. In contrast, the Object View sees the world in terms of point objects, line objects, and area objects. Spatial interpolations essentially are weighting methods which assign weights to sampled values to obtain values for unsampled

sites. The unavoidable outcome is smoothing the values over a space. A significant portion of research using raster data implicitly adopts the field view of the world and models the physical and ecological environment at scales where the impacts from line objects are non-existent (e.g., crop yields at the crop field level [28]) or negligible (impacts of land-use types on land surface temperature [29]). Under these circumstances, it would be appropriate to use a covariance function to approximate continuous and gradually changing landscapes. However, in this study, the study area is crisscrossed by roads into discontinuous spaces of odd shapes, essentially an Object View. The road density and thus the spatial autocorrelation vary significantly making any spatial kernel unsuitable. In addition, the interpolation in the study aims at approximating the network travel cost. Since roads as line features extend in linear or curvilinear fashions, the trend of pixel values should also occur linearly or curvilinearly. In other words, this is a case of one-dimensional interpolation. Most existing spatial covariance functions are designed for two-dimensional space and cannot correctly model linear positions of a road. Interpolation will fill in the missing values over a space in all directions (rather than along a defined line). In the direction which happens to coincide with a line feature, the interpolation may get it close. However, in other directions, the interpolation would be incorrect. This explains why the normalized RMSEs tend to be lower on roads than on other spaces in **Figure 8(d)**. The inaccuracy creates significant cost distortions for a facility, negating the purpose of the modeling. In general, one-dimensional kriging proves to be less successful than some other approaches [30]. Much more needs to be done in this area.

In contrast, ABM approaches do not see the world as contrasting field versus object views. Patches (pixels), turtles (points), links (lines), and clusters of turtles or patches (areas) are all agents. Using a term from [27], these individual agents all have “self-definition”, meaning they have their own attributes. These agents interact with each other at the agent level, and an interaction needs not follow rules defined by a covariance function. Their local level interactions determine the heterogeneity and unpredictability of the autocorrelation in their behavior parameters. Such “random” spatial autocorrelation can hardly be described and generalized by covariance functions. ABM, as adopted in this study, forfeits guessing the pattern of spatial autocorrelation altogether. It focuses on agents’ behaviors which are traveling on road networks. Certain relationships between agents, such as autocorrelation (e.g., relationships in travel costs between nodes on a graph), do not happen by design, but emerge as a result of their positions on the graph and destinations of their travels. While an ABM model can certainly adopt a global kernel function, as an option provided by the ABM software GAMA Platform, it has the option not to do so. This characteristic works well for situations where a spatial pattern is complex involving spatial features resembling both a field view and an object view.

4.3. Impacts of Polyline Shapefiles on Routing

As discussed earlier, to perform routing in NetLogo efficiently, the complete

roads polyline shapefile in Madison County (**Figure 3**) is simplified into a minimum spanning tree, as shown in **Figure 4**. Using this simplified polyline shapefile to create a graph, the initial optimal solution is displayed in **Figure 10**.

One prominent feature in **Figure 10** is that some collection centers provide services in faraway communities, which are closer to other collection centers and thus would supposedly be in the areas of service of other collection centers. This feature also goes against the “First Law of Geography” which states that near things are related more than distant things [31]. A closer observation reveals that those long desire lines tend to extend along or near some long lines within the input shapefile.

The result displayed in **Figure 10** is not meaningful. We believe that this is attributed to the topological structure of the shapefile (and the resultant graph) used in routing. In a road network, local roads tend to cluster in communities with many cycles (closed circuits where a path starts from a point, connects with other points, and returns to the same point) and major roads extend across long distances connecting local communities. In creating a shapefile reflecting such a structure, linear features are created by digitizing between two points over a certain distance in-between. Besides two end points of an edge, additional points are often needed to make the line geometry resemble the shape of an actual road, resulting in polylines. In addition, local communities tend to form road grids manifested as many short edges and nodes in closed circuits in a shapefile while the major roads between communities show up as long edges with fewer nodes. When creating a minimum spanning tree, local cycles are removed, and the major long links remain. This leads to a topological structure of uneven road lengths in the polyline shapefile. In **Figure 11**, the original minimum spanning tree polyline shapefile shows an edge length range of over 60,000 meters (60 km). In **Table 1**, the original shapefile has an edge length standard deviation more than 1.5 times the edge average length (coefficient of variation or CV), and the longest edge is 30 times the edge average length. In **Figure 12**, fourteen roads

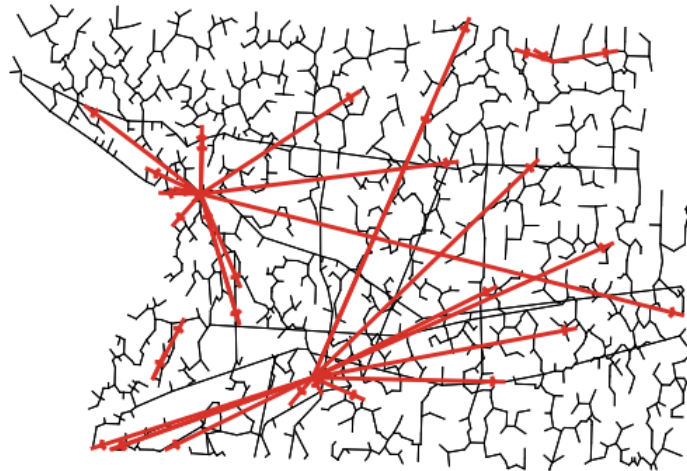


Figure 10. Desire lines from the solution of the capacitated $p = 4$ facility model using unimproved minimum spanning tree polyline shapefile.

The Length Distribution of Edge Lengths

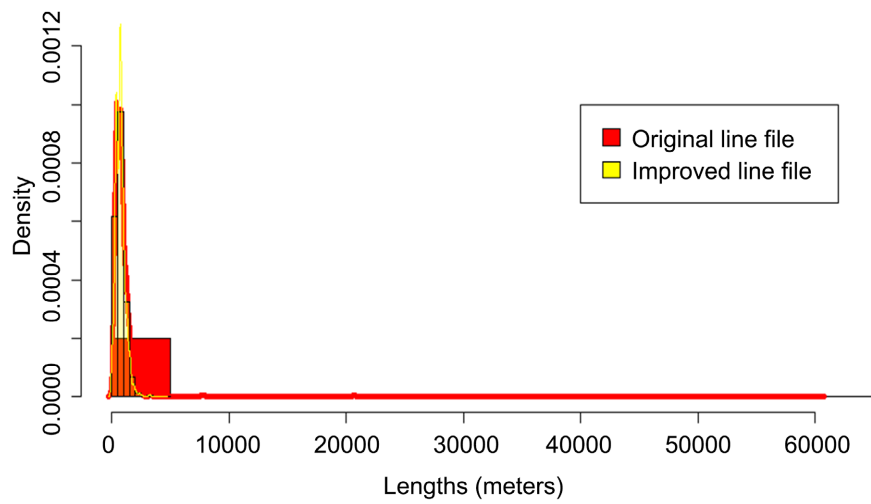


Figure 11. The length distributions of edges in the minimum spanning tree line shapefiles with and without improvement.

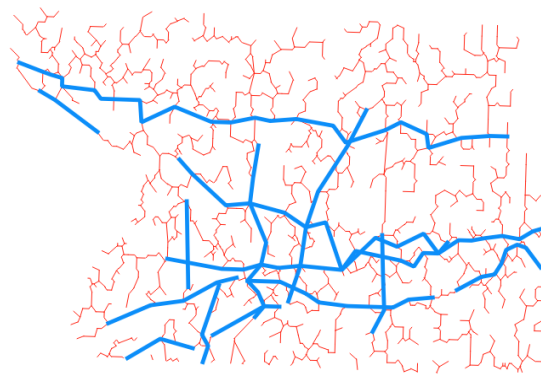


Figure 12. Fourteen roads are in blue, each longer than 5 km and together account for 22% of the total lengths in the road system used in modeling.

Table 1. Topological characteristics of original and improved polyline shapefiles.

Network Attributes	Comparison of Polyline Shapefiles	
	Original	Improved
Number of nodes	1287	1667
Number of edges	1287	1780
Median length of edges/Mean length of edges	0.80	0.98
Standard deviation of length of edges/Mean length of edges*	1.56	0.56
Range of length of edges/Mean length of edges	30.2	6.15
Inter-quartile range of length of edges/Mean length of edges	0.67	0.75
Max length of edges/Mean length of edges	30.2	6.2
Min length of edges/Mean length of edges	0.001	0.003

*The coefficient of variation of edge lengths.

are collectively highlighted in blue, each with a length of more than 5000 meters (5 km) with a total of more than 60,000 meters (60 km). These 14 polylines account for 1% of the total number of lines but 22% of the total road length within the minimum spanning road shapefile. Just as importantly, these long lines extend over long distances connecting communities in different regions.

One problem with using a network with some very long links is that while an optimal location may occur somewhere between the two end nodes, any intermediate location is not available for search since there are no nodes between the end points. More importantly, certain optimal search algorithms, as used in this study, evaluate nodes for optimization in their adjacency sequence (nodes separated by one edge). If two adjacent nodes are the end points of a long link, this would send the search process to a distant region in the study area, eliminating nearby nodes from the sequential evaluation. The original minimum spanning tree line file contains lines with significantly uneven lengths. Since major roads tend to connect with local roads at critical junctions and extend across regions for long distances, they may send the optimal searches back and forth between distant regions, creating cross-region connections, rather than local connections. Nodes close to these long lines may be repeatedly searched but those far from long links may not get enough chance of being assessed during the searches. In general, uneven length distribution and the prominent positions of the long links in a graph may contribute to the optimal search process being dominated by the long links, leading to the long and cross-region serving connections along the long links as seen in **Figure 10**.

To improve the optimal routing, we split the 14 long polylines in **Figure 12**, with all resultant lines being shorter than 1000 meters or 1 km. The road map looks the same, but the shapefile file now contains 380 more nodes and 493 more edges than the original shapefile. This significantly changed the topological structure of the shapefile giving it a much more even edge length distribution than the original polyline shapefile. As shown in **Figure 11**, the improved shapefile has a length range of less than 5000 meters (5 km), which is only 8% of the range in the original line file. **Table 1** shows that the improved shapefile has a much smaller coefficient of variation, a drastically smaller mean-normalized range, and a smaller ratio of the maximum road length to the mean, than the original shapefile. Without an extreme line structure, the routing and the search process are less influenced by the long lines. In addition, there are more nodes between end nodes along a road. This improves the chance that a better solution can be found since those locations are now available for optimization. Indeed, using the improved line shapefile, the total travel cost of serving all communities by four facilities is 518 km, a 14% reduction from 600 km from using the original line shapefile. Using the improved polyline shapefile in the study, the resultant spatial pattern of location-allocation, as shown in **Figure 5**, is meaningful in terms of service areas being delineated based on distances travelled.

In general, compared with a graph with limited numbers of edges/nodes and a

skewed topological structure, a graph with more edges and nodes, shorter edge lengths, and balanced topological organizations provides more feasible sites for optimal locations and allows more flexibility in routing search in terms of directions and lengths. More nodes on a network also make it easier to create a connected graph for routing as origin and destination nodes can more easily snap to the closest available network nodes with less distortion of the network. Upon creation, a polyline shapefile contains nodes at intersections, ends of lines, and between nodes. The former two types of nodes become external when a graph is created. They can be used for optimal routing search. Many of the last type of nodes remain as internal and not available for routing search. Splitting polyline shapefile edges strategically creates more external nodes, which eventually improves the topological structure of a graph and contributes to better routing performance.

5. Summary and Concluding Remarks

This study reveals two problems in applying GIS in spatial optimization modeling, using a capacitated p -median location-allocation problem. First, it shows that standard spatial interpolation generates significant errors with errors rising between roads. This occurs when the study area is segmented by roads into highly discontinuous and unpredictable spaces which make it difficult for any spatial covariance function to specify the appropriate spatial autocorrelation pattern. Given such a complex space, spatial interpolation is an unrealistic choice for the problem at hand. In contrast, using an agent-based approach eliminates the need for adopting any spatial covariance function. Rather than imposing a certain spatial autocorrelation structure on the model, the ABM focuses on the individual agents' behaviors and allows spatial relationships to emerge. Although this does not mean ABM can create the perfect spatial variation, spatial simulation does provide a viable option to model spatial variation without using interpolation.

This study also shows that routing can be significantly impacted by the topological structure of a polyline shapefile. When the distribution of road lengths is very uneven and skewed, the routing may be dominated by long edges and the search for optimization may not be meaningful. This study corrects this problem by splitting very long edges into shorter segments, which leads to significant improvement in the model in terms of separating service areas for different facilities following the well-known distance decay effect. While the skewed topological structure of the original graph is a result of using an unimproved minimum spanning tree polyline shapefile, this still should remind the spatial modeler of the potential impacts of the topological structure of the polyline shapefile to avoid undesirable modeling outcomes.

The above findings are related to the adoption of GIS techniques in spatial optimization modeling. Before GIS was integrated with facility location study, researchers used distance matrices and network graphs without questioning. With

GIS techniques having become toolkits in locational modeling, modelers should be aware that many distance-related parameters (distance, shipping cost, travel time, etc.) in the model may be subject to the topological structure of the polyline shapefile used, as demonstrated in this study. To a certain extent, distance and routing are artefacts of the polyline shapefile topological structure. The modeler may benefit from exploring the best topological structure when performing routing on any graph created from polyline shapefiles.

The contributions from this study are unique in that issues addressed in the paper have not received much attention in the field of spatial optimization. While spatial interpolation remains as a significant technique used to model spatial variation, other options such as ABM should be explored and made available in situations where the old procedures fail. Together with the impacts of polyline shapefile topology on routing, these issues are a reminder that the integration of GIS with spatial optimization study is not fool proof. Caution should be exercised to make sure that GIS techniques are used appropriately.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Tong, D. and Murray, A.T. (2012) Spatial Optimization in Geography. *Annals of Association of the American Geographers*, **102**, 1290-1309. <https://doi.org/10.1080/00045608.2012.685044>
- [2] Lei, T.L. (2021) Integrating GIS and Location Modeling: A Relational Approach. *Transactions in GIS*, **25**, 1693-1715. <https://doi.org/10.1111/tgis.12804>
- [3] Frenk, J.B.G., Melo, M.T. and Zhang, S. (1994) The Weiszfeld Method in Single Facility Location. *Investigacao Operacional*, **14**, 35-59.
- [4] Miehle, W. (1958) Link-Length Minimization in Networks. *Operations Research*, **6**, 165-302. <https://doi.org/10.1287/opre.6.2.232>
- [5] ReVelle, C.S. and Swain, R.W. (1970) Central Facility Location. *Geographical Analysis*, **2**, 30-42. <https://doi.org/10.1111/j.1538-4632.1970.tb00142.x>
- [6] Erlenkotter, D. (1978) A Dual-based Procedure for Uncapacitated Facility Location. *Operations Research*, **26**, 937-1094. <https://doi.org/10.1287/opre.26.6.992>
- [7] Teitz, M. and Bart, P. (1968) Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph. *Operations Research*, **16**, 901-1091. <https://doi.org/10.1287/opre.16.5.955>
- [8] Moon, D. and Chaudhry, S.S. (1984) An Analysis of Network Location Problem with Distance Constraints. *Management Science*, **30**, 263-394. <https://doi.org/10.1287/mnsc.30.3.290>
- [9] Daskin, M.S. and Maass, K.L. (2015) The p-Median Problem. In: Laporte, G., Nickel, S. and Saldanha da Gama, F., Eds., *Location Science*, Springer, Cham, 21-45. https://doi.org/10.1007/978-3-319-13111-5_2
- [10] Church, R.L. and Murray, A.T. (2009) Business Site Selection, Location Analysis, and GIS. John Wiley & Sons, Hoboken. <https://doi.org/10.1002/9780470432761>

- [11] Hari Shankar, M. (2017) Geographic Information System Based Solution for Location Allocation Problem for Finding High Quality Service Locations. *International Journal of Advanced Remote Sensing and GIS*, **6**, 2377-2394. <https://doi.org/10.23953/cloud.ijarsg.302>
- [12] Murray, A.L. (2016) Maximal Coverage Location Problem: Impacts, Significance, and Evolution. *International Regional Science Review*, **39**, 5-27. <https://doi.org/10.1177/0160017615600222>
- [13] Lei, T.L., Church, R.L. and Lei, Z. (2016) A Unified Approach for Location-Allocation Analysis: Integrating GIS, Distributed Computing and Spatial Optimization. *International Journal of Geographical Information Science*, **30**, 515-534. <https://doi.org/10.1080/13658816.2015.1041959>
- [14] Bonneu, F. and Thomas-Agnan, C. (2009) Spatial Point Process Models for Location-Allocation Problems. *Computational Statistics and Data Analysis*, **53**, 3070-3081. <https://doi.org/10.1016/j.csda.2008.10.016>
- [15] Bruno, G., Genovese, A. and Sgalambros, A. (2010) An Agent-Based Framework for Modeling and Solving Location Problems. *TOP*, **18**, 81-96. <https://doi.org/10.1007/s11750-009-0116-1>
- [16] Han, J., Zhang, J., Zeng, B. and Mao, M. (2021) Optimizing Dynamic Facility Location-Allocation for Agricultural Machinery Maintenance Using Benders Decomposition. *Omega*, **105**, Article ID: 102498. <https://doi.org/10.1016/j.omega.2021.102498>
- [17] Sun, X., Yu, H. and Solvang, W.D. (2020) Solving the Location Problem of Printers in a University Campus Using p-Median Location Model and AnyLogic Simulation. In: Wang, Y., Martinsen, K., Yu, T. and Wang, K., Eds., *IWAMA 2019: Advanced Manufacturing and Automation IX*, Springer, Singapore, 577-584. https://doi.org/10.1007/978-981-15-2341-0_72
- [18] Bartkowski, B., Beckmann, M., Drechsler, M., Kaim, A., Liebelt, V., Müller, B., Witing, F. and Strauch, M. (2020) Aligning Agent-Based Modeling with Multi-Objective Land-Use Allocation: Identification of Policy Gaps and Feasible Pathways to Biophysically Optimal Landscapes. *Frontiers in Environmental Science*, **8**, 543832. <https://doi.org/10.3389/fenvs.2020.00103>
- [19] Zhang, J. and Robinson, D.T. (2022) Investigating Path Dependence and Spatial Characteristics for Retail Success Using Location Allocation and Agent-Based Approaches. *Computers, Environment and Urban Systems*, **94**, Article ID: 101798. <https://doi.org/10.1016/j.compenvurbsys.2022.101798>
- [20] Yin, X., Bushaj, S., Yuan, Y. and Büyüktaktın, İ.E. (2023) COVID-19: Agent-Based Simulation-Optimization to Vaccine Center Location Vaccine Allocation Problem. *IISE Transactions*. <https://doi.org/10.1080/24725854.2023.2223246>
- [21] Crooks, A., Malleson, N., Manley, E. and Heppenstall, A. (2019) Agent-Based Modelling & Geographical Information Systems: A Practical Primer. Sage Publications Ltd, Los Angeles. <https://doi.org/10.4135/9781529793543>
- [22] Chen, Y.C., Yao, H.L., Weng, S.D. and Tai, Y.F. (2022) An Analysis of the Optimal Facility Location of Tourism Industry in Plain Region by Utilizing GIS. *SAGE Open*, **12**, page. <https://doi.org/10.1177/21582440221095020>
- [23] Law, M. and Collins, A. (2022) Getting to Know ArcGIS Desktop 10.8. ESRI Press, Redlands.
- [24] Heppenstall, A., Crooks, A., Malleson, N., Manley, E., Ge, J. and Batty, M. (2021) Future Developments in Geographical Agent-Based Models: Challenges and Opportunities. *Geographical Analysis*, **53**, 76-91. <https://doi.org/10.1111/gean.12267>

-
- [25] Arostegui Jr., M.A., Kadipasaoglu, S.N. and Khumawal, B.M. (2006) An Empirical Comparison of Tabu Search, Simulated Annealing and Genetic Algorithms for Facilities Location Problems. *International Journal of Production Economics*, **103**, 742-754. <https://doi.org/10.1016/j.ijpe.2005.08.010>
- [26] Hakimi, S.L. (1965) Optimal Distribution of Switching Centers in a Communication Network and Some Related Theoretic Graph Theoretic Problems. *Operations Research*, **13**, 343-514. <https://doi.org/10.1287/opre.13.3.462>
- [27] O'Sullivan, D. and Unwin, D.J. (2010) *Geographic Information Analysis*. John Wiley & Sons, New York. <https://doi.org/10.1002/9780470549094>
- [28] Pritsolas, J. (2018) *Principal Component Analysis and Spatial Regression Techniques to Model and Map Corn and Soybean Yield Variability with Radiometrically Calibrated Multitemporal and Multispectral Digital Aerial Imagery*. Master Thesis, Southern Illinois University Edwardsville, Edwardsville.
- [29] Khan, M.S., Ullah, S., Sun, T., Rehman, A. and Chen, L. (2020) Land-Use/Land-Cover Changes and Its Contribution to Urban Heat Island: A Case Study of Islamabad, Pakistan. *Sustainability*, **12**, Article 3861. <https://doi.org/10.3390/su12093861>
- [30] Naprstek, T. and Smith, R.S. (2019) A New Method for Interpolating Linear Features in Aeromagnetic Data. *Geophysics*, **84**, 15-24. <https://doi.org/10.1190/geo2018-0156.1>
- [31] Tobler, W. (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, **46**, 234-240. <https://doi.org/10.2307/143141>