Spatial Optimization Approaches for Solving the Continuous Weber and Multi-Weber Problems

by

Jing Yao

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2012 by the
Graduate Supervisory Committee:

Alan Murray, Chair
Pitu Mirchandani
Michael Kuby

ARIZONA STATE UNIVERSITY

December 2012

ABSTRACT

Facility location models are usually employed to assist decision processes in urban and regional planning. The focus of this research is extensions of a classic location problem, the Weber problem, to address continuously distributed demand as well as multiple facilities. Addressing continuous demand and multi-facilities represents major challenges. Given advances in geographic information systems (GIS), computational science and associated technologies, spatial optimization provides a possibility for improved problem solution.

Essential here is how to represent facilities and demand in geographic space. In one respect, spatial abstraction as discrete points is generally assumed as it simplifies model formulation and reduces computational complexity. However, errors in derived solutions are likely not negligible, especially when demand varies continuously across a region. In another respect, although mathematical functions describing continuous distributions can be employed, such theoretical surfaces are generally approximated in practice using finite spatial samples due to a lack of complete information. To this end, the dissertation first investigates the implications of continuous surface approximation and explicitly shows errors in solutions obtained from fitted demand surfaces through empirical applications.

The dissertation then presents a method to improve spatial representation of continuous demand. This is based on infill asymptotic theory, which indicates that errors in fitted surfaces tend to zero as the number of sample points increases to infinity. The implication for facility location modeling is that a solution to the discrete problem with greater demand point density will approach the theoretical

optimum for the continuous counterpart. Therefore, in this research discrete points are used to represent continuous demand to explore this theoretical convergence, which is less restrictive and less problem altering compared to existing alternatives.

To Xiaoxiang and Shuyao

ACKNOWLEDGEMENTS

Over the last three years, I have received much support and encouragement from many people.

I would like to acknowledge the three members of my advisory committee, Professors Alan Murray, Pitu Mirchandani and Michael Kuby. I would like to thank especially my advisor Prof. Alan Murray for believing in this project and offering me the opportunity to pursue my Ph.D. degree at ASU. He always encouraged me to explore on my own and gave me excellent guidance whenever necessary. His wisdom, patience and support have helped me overcome many difficulties in producing this dissertation. Prof. Pitu Mirchandani and Prof. Michael Kuby offered help in improving my analytical skills and expanding my academic vision in research.

I would like to extend my gratitude to Prof. Luc Anselin, Prof. Sergio Rey and Dr. Julia Koschinsky for providing me the chance to work with the research groups focused on software development of PySAL and GeoDaSpace in the GeoDa Center at ASU. This unique experience greatly deepened my understanding of spatial analysis theory and methodology, and significantly improved my GIS coding and software design skills.

I also wish to thank Prof. Victor Agadjanian at the School of Social and Family Dynamics at ASU and Dr. Michael Cudnik at College of Medicine at the Ohio State University for providing me the opportunity to work with them on their research projects. Such experience and assistance definitely enhanced my many inter-disciplinary and cross-disciplinary thinking and will benefit my future research.

I would like to thank the timely funding support from National Natural Science Foundation of China (Young Scientists Fund, Grant 41201117), enabling me to continue and deepen my research over the next three years.

Finally, I owe a debt of gratitude to my family. In the past three years, my husband, Xiaoxiang, made enormous sacrifices and gave me tremendous support, both mentally and materially. I want to say sorry to my 5-year-old daughter, Shuyao, for not accompanying her when she started preschool in China. I am indebted to my parents, parents-in-law and many friends both in China and United States for their unlimited love and steadfast confidence in me.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

Chapter 1

INTRODUCTION

1.1 Background

In ancient China, the discipline of spatial arrangement relies on knowledge of astronomy and geography, known as Feng Shui, and is inherently considered in location selection to achieve harmony and fortune, and remains very popular in Southern China. In fact, it has long been recognized in geography that location related decisions play an important role in human activities since “all human activities involve the choice, either explicit or implicit, of location” (Church and Sorensen 1996). Nowadays, location science continues to be a very active research area involving people from diverse disciplines, including mathematics, operation research, management science, geography, urban planning, and industrial engineering among others.

1.1.1 Facility Location Problems

Location problems usually concern determining one or more sites for facilities under certain constraints, sometimes involving demand allocation, to optimize certain objectives. Current objective functions in location models can be classified as minimization, maximization, min-max and max-min (Brandeau and Chiu 1989). A minimization model can be used to minimize total or average travel distance or cost (Hakimi 1964, 1965, ReVelle and Swain 1970). A maximization model can be applied to optimize service coverage (Church and ReVelle 1974). Min-max models are appropriate when seeking locations that minimize the distance the farthest customer to its closest facility (Hakimi 1964, 1965). Finally, max-min models reflect a decision to maximize the closest customer to its closest facility. Of interest in this dissertation are minimization models, the Weber problem and the multi-facility Weber problem, where demand is continuously distributed.

As one of the first formalized location problems ever posed, the Weber problem involves:

Placing a single facility in continuous space in order to serve a finite set of demand points, where the goal is to minimize total transportation costs.

In addition to the classic context seeking the best location for a factory, it has been widely applied in other situations as well. For example, it can be use to find the best location for an emergency center to minimize average response time. Since it was first proposed in the 17th century, considerable research effort has been devoted to the Weber problem and related models (Wesolowsky 1993, Drezner *et al*. 2002), largely attributed to its capability and potential for extension. In fact, many location models can be considered built upon the Weber problem (Drezner *et al*. 2002), such as the continuous location-allocation problem (Cooper 1963, 1964) and the p-median problem (Hakimi 1964, 1965).

One extension of the Weber problem of interest in this dissertation is continuous demand instead of discrete demand proposed in the original formulation, referred to as the continuous Weber problem (Drezner 1995). Generally, demand is modeled as discrete points in facility location problems due to the lack of detailed data or the need to simplify model formulation and solution

(Miller 1996, Francis *et al*. 2009). However, it is more reasonable in practice to conceive of the underlying demand as continuously distributed such as population and risk. The solution methods for the Weber problem are not appropriate for the continuous counterpart. In fact, the objective of the latter involves a double integral which makes it quite difficult to solve.

Another concern in this dissertation further extends the continuous Weber problem by taking into account siting several facilities simultaneously, which is known as the continuous location-allocation problem or multi-Weber problem (MWP) (Cooper 1963, Plastria 1995). In addition to finding locations for multiple facilities, the problem involves an allocation process as well with the goal that each demand is served by the nearest facility. This generalization is more complicated in both model formulation and solution because of the need to address both location and allocation decisions.

While many solution approaches have been proposed to solve the continuous Weber problem (Drezner 1995, Carrizosa *et al*. 1998, Fekete *et al*. 2005) and the continuous multi-Weber problems (Marucheck and Aly 1981, Suzuki and Okabe 1995, Murat *et al*. 2010), unfortunately, they are all subject to simplified assumptions regarding demand region or continuous distribution. Given the geographic nature inherent in location related decision-making, this dissertation explores spatial optimization approaches to solve those two extensions of the Weber problem.

1.1.2 Spatial Optimization

Spatial optimization can be simply considered as the science of optimal spatial arrangement (Church 2001). It is rooted in classic geographic theories, such as Von Thunen's land rent theory to explain the efficiency of observed agricultural activities, Alfred Weber's theory of industrial location and cost minimization, and Christaller's central place theory concerned with the organization patterns of cities (Church 2001, Murray 2010). Generally, spatial optimization involves the attempt to identify the best locations for facilities, as well as arrange land use or other resources spatially with respect to some objectives, often relying on distance or cost constraints. The development of spatial optimization is mostly attributed to the availability of more accurate spatial data, progress in geographic information science, the evolution of optimization algorithms, and the advance of computational technologies.

Recent years have seen a proliferation in application of spatial optimization approaches involving GIS in location problems, largely attributed to the advances in geographic information science and computer technologies (Church 2002, Church and Murray 2009, Murray 2010). GIS is an information system designed to collect, store, manipulate, analyze, manage and present all types of geographically referenced data (Longley *et al*. 2011). It has played an important role in spatial optimization in terms of data input and visualization, as well as model formulation and model analysis (Longley and Batty 1996, Church 1999, Murray 2010).

Traditionally, GIS has been used to collect and prepare spatial information as input for spatial optimization models and visualize results through desktop mapping. For instance, data aggregation, which is often applied in optimization models to reduce problem size or facilitate model formulation, can be easily achieved in GIS by extracting data at a certain geographic scale like census tracts. Another example related to data acquisition is that distance as a common measure of accessibility in many models can be obtained through spatial analysis functionality of GIS. Further, model solutions can be presented by the powerful visualization capability of GIS, such as the Voronoi diagram depicting trade areas (Suzuki and Okabe 1995) and the spider diagram describing the scheme of location-allocation (Bender *et al*. 2002). Visualization is also important in understanding objective, decision and model spaces (Densham 1994) and detecting underlying problems that otherwise cannot be identified (Murray 2005).

In addition to data provision and visualization, GIS can also aid in model formulation and analysis. In one respect, GIS helps model formulation by evaluating spatial properties and relationships. Murray (2010) highlighted the importance of spatial search in the case where an area of interest needs to be identified using complicated spatial relationships, such as adjacency, contiguity, containment and intersection. For example, constraints based on adjacency can be used to avoid the simultaneous selection of adjacent sites (Murray and Church 1996, Downs *et al*. 2008). Contiguity or connectivity is often necessary in land acquisition (Cova and Church 2000, Wu and Murray 2007, Kim and O'Kelly 2009). In another respect, sometimes models can be solved just through basic GIS operations when the available options are limited. A well known example is the work of McHarg (1969) identifying suitable roads through comparison of a series of map layers, each with a different theme. Such an operation is a basic function of GIS known as overlay and is common in suitability analysis. Another GIS function often used in model solution is the Voronoi diagram. For example, Suzuki and Okabe (1995) developed a heuristic using a Voronoi diagram to solve the continuous p-center problem, which was later extended in Wei *et al*. (2006) for emergency warning sirens location.

Obviously, GIS is an indispensable component in spatial optimization and can help to better understand the problem under study and assist decision-making processes. Therefore, it is crucial in spatial optimization that GIS is integrated to the greatest extent possible with location models in formulation, solution and evaluation in addition to simple data management and graphic display.

1.2 Research Objectives

There are a number of goals and objectives associated with this dissertation research. The primary goals may be summarized as follows:

(1) Examine alternatives for effectively representing continuous demand;

(2) Develop better solution approaches for the continuous Weber problem;

(3) Develop better solution approaches for the continuous multi-Weber problem.

The intent is to address these goals and objectives through integration of GIS, both to better represent continuously distributed demand as well as to develop enhanced spatial optimization techniques that exploit geographical knowledge.

1.3 Organization of Research

The aim of this research is to employ spatial optimization approaches that combine GIS and operation research methods to solve the continuous Weber and continuous multi-Weber problems. The dissertation is structured as follows.

Given the importance of spatial representation in facility location problems and the fact that continuous representation in practice is usually approximated by a finite set of sample points, Chapter 2 investigates the implications of continuous surface approximation in spatial analysis, particularly in location analysis. First, the process of how continuous surfaces are approximated in a GIS environment is described. The errors introduced in fitted surfaces are then explored. An empirical study using such surfaces in facility location modeling is employed to explicitly demonstrate the cumulative errors in analysis results.

Following the discussion of continuous representation, Chapter 3 focuses on solving the continuous Weber problem. Simplifying assumptions in existing solution approaches are discussed. Then, the representation method employed in this research is addressed, built on infill asymptotic theory widely applied in spatial data analysis. Further, the mathematical formulation of the continuous Weber problem is offered and a spatial optimization method based on the proposed continuous representation is presented. Empirical results are provided and discussed, highlighting the advantage of integration with GIS.

Chapter 4 extends the continuous Weber problem to take into account multiple facilities - the continuous multi-Weber problem. It is more complex due to the consideration of allocating demand to facilities in addition to seeking optimal locations. Relevant solution techniques are reviewed and their assumptions for continuous demand are discussed. Following the mathematical formulation of the continuous multi-Weber problem, a spatial optimization solution approach is detailed, building on the continuous representation method presented in Chapter 3. This is followed by application results and associated discussion.

Finally, Chapter 5 concludes the dissertation by summarizing research findings and implications, as well as the contribution to theories and methods of location science. Directions for future research are also discussed.

Chapter 2

CONTINUOUS SURFACE REPRESENTATION AND APPROXIMATION: SPATIAL ANALYTICAL IMPLICATIONS*

2.1 Introduction

Spatial representation has long been a critical issue in spatial analysis (Miller and Wentz 2003, Goodchild and Haining 2004, Church and Murray 2009). It is widely accepted in the geographic information systems (GIS) community that continuous fields and discrete objects are the two basic conceptual models of geographical space (Openshaw 1983, Peuquet 1988, Couclelis 1992, Frank 1992, Goodchild 1992, Burrough 1996, Worboys and Duckham 2004, Longley *et al*. 2011). The field view considers the phenomena continuously distributed across space, such as rainfall, air quality, elevation, temperature, population density and land use. In contrast, the object view conceives of the real world as an empty geographic space littered with discrete entities, such as roads, rivers, lakes, parcels and buildings.

There are basically two spatial representation approaches to reflect these perspectives: raster and vector (Goodchild 1992, Burrough and McDonnell 1998, Worboys and Duckham 2004, Longley *et al*. 2011). In a vector-based model, geographical entities are described by geometry having distinct locations and boundaries: point, line and polygon. Each entity can have one or more attributes. In a raster representation, geographical space is usually delineated by a

* This chapter is a slightly modified version of a paper published in International Journal of Geographical Information Science, co-authored with Alan Murray.

continuous grid of square cells, with each raster cell having an associated attribute(s). It should be noted that the two conceptual models (object and field) and the two representation approaches (vector and raster) do not have an exact one-to-one relationship (Longley *et al*. 2011). Of course, the choice of representation is usually not straightforward, but rather dependent on the spatial phenomena under study, the analysis context, data availability, computational efficiency and so on.

Of interest here is the representation of the continuous field. In principle, the set of locations contained in continuous geographic space is infinite, while representation in a digital computer based environment is finite (Winter 1998, Cova and Goodchild 2002). Further, although in mathematical terms a field can be modeled as a function mapping a location in geographic space to a value in an attribute domain (Worboys and Duckham 2004, Kjenstad 2006), in practice such functions are never simply known or given with certainty due to a lack of complete information. As a result, representations of fields are necessarily abstracted or approximated at a certain spatial scale, and continuous geographic space in GIS is usually fitted or approximated through spatial interpolation based on sample points. Thus, errors and uncertainty are unavoidable as it is impossible to collect and store the attribute value about each and every location (Goodchild 2004).

When an approximated continuous surface is used in spatial analysis, resident errors will inevitably be introduced into subsequent results, since spatial analysis implementation is closely related to how we represent spatial phenomena

(Goodchild *et al*. 1992, Miller and Wentz 2003). In other words, the results of spatial analysis are reliant on the representation methods employed. Of course, theoretically, as the number of sample points tend to infinity, greater accuracy in the approximation is possible, something known as infill asymptotic theory (Cressie 1993, Stein 1999). In practice, however, it is hoped that a finite number of sample points would be sufficient to reasonably approximate a continuous field.

The aim of this chapter is to explore the significance of approximate representation of a continuous surface as well as its implications in certain spatial analyses. Section 2.2 discusses spatial representation and related issues. Then, methods that are employed to fit continuous surfaces are discussed, followed by a designed experiment to explicitly show the errors in an approximated surface. Propagation of errors in spatial analysis is then examined. The chapter concludes with a discussion of results and practical implications.

2.2 Background

How to represent infinite geographic space in a finite digital computing environment has been a great challenge in GIScience, receiving significant attention over the past few decades. Discrete-object and continuous-field distinctions, introduced in late 1980s and early 1990s, are well known perspectives for conceptualizing geographical space (Couclelis 1992, Goodchild 1992, Worboys and Duckham 2004, Longley *et al*. 2011), and are also considered as formal spatial concepts that have been adopted in GIScience (Goodchild 2010).

The object/field view assumes that geographic phenomena are either discrete or continuous. Due to the complexity of reality and application needs, scholars have argued that both perspectives are not necessarily exclusive and that they can coexist or be integrated (Peuquet 1988, Couclelis 1992, Winter 1998, Egenhofer *et al*. 1999, Worboys and Duckham 2004). Winter (1998) introduced a hybrid representation, valuable for data integration and interoperability. Yuan (2001) offered a hierarchy space-time framework that included both object and field properties to represent complex geographic phenomena such as precipitation. Another example is the field object that integrates both perspectives by mapping locations in a field to objects (Cova and Goodchild 2002, Goodchild *et al*. 2007). Kjenstad (2006) provided a common base-model for the two conceptual models using the unified modeling language (UML). Further, Goodchild *et al*. (2007) suggested that all of these spatial concepts could be unified in a general theory relying on the atomic form of geographic information: the geo-atom and the geo-dipole. Though spatial representation has become more complex given these new concepts, discrete objects and continuous fields remain the foundation for geographic representation (Goodchild *et al*. 2007).

The focus in this study is the representation of continuous fields. The concept of "field" originates from classical physics and has been employed to describe physical properties such as electricity, magnetism and gravity (Yuan 2001, Kjenstad 2006). Fields in GIS have been extended as a perspective for conceptualizing geographic space, assuming phenomena varies continuously across space. There are many ways to implement a field view in current GIS, such as spatial tessellations consisting of regular or irregular polygons and uniformly or non-uniformly distributed point sets. Comprehensive discussions of modeling continuous fields abound in the literature, including that of Goodchild (1992), Goodchild *et al*. (1992), Frank (1992), Burrough (1996), Kemp (1997a, b), Worboys and Duckham (2004) and Longley *et al*. (2011).

An important issue that arises when modeling a field is the error and uncertainty inherent in any fitted continuous surface. First, uncertainty in imperfect sample data from various sources can potentially impact the accuracy of estimated values. For example, Kyriakidis and Goodchild (2006) studied the propagation of sampling errors through spatial interpolation and derived prediction errors in fitted surfaces. Also, all techniques used to construct surfaces representing continuous phenomena are subject to assumptions that are usually not satisfied in practice, and the resulting errors have been investigated in many applications, such as elevation models (Wood and Fisher 1993; Gong *et al*. 2000), monitoring data of sulfur dioxide concentrations (Host *et al*. 1995), areal interpolation between zonal systems (Fisher and Langford 1995) and estimation of rainfall magnitude (Tomczak 1998).

As errors/uncertainties in approximated surfaces are inevitable, increasing effort has been devoted to error modeling and propagation research (Goodchild and Gopal 1989, Heuvelink 1998, Cressie and Wikle 2011). This work is largely numerical, built on the theory of error analysis in statistics, such as Monte Carlo simulation (e.g. Fisher and Langford 1995) and regression models (e.g. Carlisle 2005). Further, most approaches analyzing error propagation have focused on the impacts on common spatial operations like intersection (e.g. Heuvelink and Burrough 1993) or attribute calculations from approximated continuous data like a digital elevation model (DEM) (e.g. Oksanen and Sarjakoski 2005). With regard to the implications of such error propagating through more advanced spatial analysis, however, related research is almost non-existent. An exception is the uncertainty work of Murray *et al*. (2008) focused on facility siting where they explicitly consider the impacts of different spatial representations of continuous demand space. Needless to say, cumulative error and uncertainty using approximated surfaces in spatial analytical methods is generally not well understood, yet can be significant. We will explore this issue in the remainder of the paper.

2.3 Approximations of Continuous Surfaces

To facilitate visualization and analysis of continuously distributed phenomena, a surface is usually approximated based on observed values at sampled locations. Sampling procedures are adopted for dealing with a very large or infinite population, concerned with selection of a subset of individuals from which some characteristics of the whole population can be estimated (Burt *et al*. 2009). Specifically, spatial sampling is a process that determines a finite set of locations in geographic space (Berry and Baker 1968) based on underlying principles of geographic phenomena – spatial dependency and spatial heterogeneity. The former indicates that similar phenomena can be found at nearby locations so a sample of observations over space will be reasonable, while the latter suggests a dependency relationship can vary across space (Anselin 1989). A similar idea is also inherent in Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970). These theories suggest that a finite sample of points is sufficient to obtain reliable inferences about values at unsampled nearby locations. Basic spatial sampling schemes include random sampling, stratified sampling, systematic sampling, clustered sampling and contour sampling (Berry and Baker 1968, Burt et *al*. 2009, Longley *et al*. 2011). The choice of sampling method depends on the known distribution of the phenomena, but also the time and cost of collecting the data.

Many methods and procedures have been developed to fit or approximate continuous surfaces using sample data (Lam 1983, Myers 1994, Burrough 1996, Mitas and Mitasova 1999). These are generally referred to as spatial interpolation methods, enabling estimation of attribute values at unobserved locations in geographic space based on observed/measured attribute values for a sampled set, generating a coverage (usually a raster grid) of the study region (Lam 1983, Cressie 1993, Longley *et al*. 2011). In other words, if $g(x, y)$ is the true attribute value at a particular location (x, y), the goal of spatial interpolation is to obtain an estimate $\hat{g}(x, y)$ that has as little error as possible at all locations (x, y). Consider the following notation:

k = index of sample points

(x_k, y_k) = location of sample point k

α_k = observed attribute value at sample point k

m = number of sample points

$g(x, y)$ = true attribute value at location (x, y)

$\hat{g}(x, y)$ = estimated attribute value at location (x, y)

$\varepsilon(x, y)$ = estimation error at location (x, y)

Then, the estimated attribute value, $\hat{g}(x, y)$, is a function of the observed sample data:

$$\hat{g}(x, y) = f(\alpha_1, \alpha_2, \alpha_3, ..., \alpha_m) \tag{2-1}$$

where $f(\alpha_1, \alpha_2, \alpha_3, ..., \alpha_m)$ is some function. In general, the intent is to minimize the estimation errors:

$$Minimize \quad E = \iint_{(x,y)\in R} \varepsilon(x, y)dxdy \tag{2-2}$$

where R is the study area and $\varepsilon(x, y) = |g(x, y) - \hat{g}(x, y)|$.

Spatial interpolation is currently available in most commercial GIS packages, including a host of techniques to support and fit a continuous surface. Common interpolation approaches include inverse distance weighting (IDW), natural-neighbor, splines, polynomial regression, and Kriging, among others, which can be classified in a number of ways according to their inherent characteristics, such as global or local, exact or approximate, deterministic or stochastic (Lam 1983, Myers 1994, de Smith *et al*. 2009). For example, natural-neighbor is a local method as it only uses neighboring data rather than all observations, as well as an exact approach because it preserves the original known

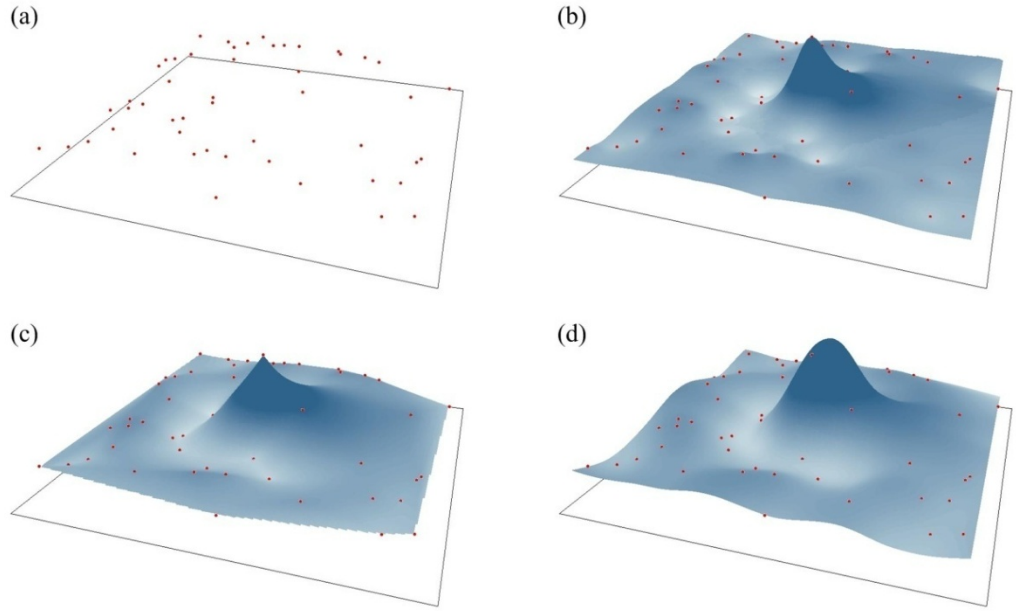


Figure 2.1: Different approximated surfaces based on the same sampled points: (a) sampled data points, (b) surface estimated using IDW, (c) surface estimated using natural-neighbors and (d) surface estimated using a spline

values at sample points. IDW is a deterministic method since it does not provide any measure of prediction accuracy. Kriging, which is widely used in geostatistics, relies on a random function and employs statistical methods to analyze sample data and errors in attribute value prediction.

Choosing an appropriate interpolation method is challenging as it depends on many factors, including the spatial phenomenon under study, attribute value type, desired accuracy, spatial variability, computational capability and assumptions employed. Given a set of sampled locations, several possible surfaces can be obtained, depending on the interpolation approaches used. For example, Figure 2.1 shows three surfaces fitted using IDW (Figure 2.1b), Natural-neighbor (Figure 2.1c) and Spline (Figure 2.1d) based on the sampled points shown in Figure 2.1a. Regardless of the method employed, errors and uncertainty are unavoidable since all interpolation techniques are subject to certain assumptions that may or may not be true (Lam 1983). While much can be done to improve spatial interpolation method accuracy through the use of ancillary information or supplementary datasets, such as remote sensing data (Wu and Murray 2007), errors and uncertainty remain. In other words, $g(x, y)$ is never likely to be known, and any approximation, $\hat{g}(x, y)$, will no doubt have error, $\varepsilon(x, y) > 0$, for most locations (x, y).

2.4 Error in Approximated Surfaces

To illustrate errors that arise due to spatial interpolation approaches, a known surface is assumed and relied upon. Mathematical surfaces are generally adopted to assess the performance of different interpolation methods (Morrison 1971, Zimmerman *et al*. 1999). In this study, Figure 2.2 represents the actual field of the continuous phenomena of interest and is defined by the following function:

$$g(x, y) = f_1(x, y) + f_2(x, y) + f_3(x, y) + f_4(x, y) + f_5(x, y) + f_6(x, y) \qquad (2\text{-}3)$$

where

$$f_1(x, y) = exp(-((x^2 + y^2)/1.8));$$

$$f_2(x, y) = 0.8 * exp(-(((x+3)^2 + (y-1)^2)/1.9));$$

$$f_3(x, y) = 0.87 * exp(-(((x-2)^2 + (y-1)^2)/1.4));$$

$$f_4(x, y) = 0.72 * exp(-(((x-1)^2 + (y+2.3)^4)/2.4));$$

$$f_5(x,y) = 0.42 * exp(-|x|^{1.2} * (sin(x+4))^2 / 0.89$$
$$-|y|^{0.85} * (cos(1.6*(y+3.3)))^4 / 0.72);$$

$$f_6(x,y) = 0.68 * exp(-(((x-4.1)^6 + (y-3.8)^2) / 0.8));$$

If the actual surface is sampled, it is possible to investigate errors that would result from the use of an approximated surface derived using an interpolation approach. Several discrete point instances are sampled here, and then used to fit an interpolation process. The number of points range from 100 to 5,000. One realization is shown in Figure 2.3 for 100 points. 30 different instances are generated for each number of points. In total, 210 different samples are generated and examined. The sampling process follows a random pattern because the intent here is to investigate errors in the estimated surface rather than comparing the performance of various interpolation techniques. In addition, a set of 810,000 points regularly spaced are used to evaluate predictive capabilities and resulting errors.

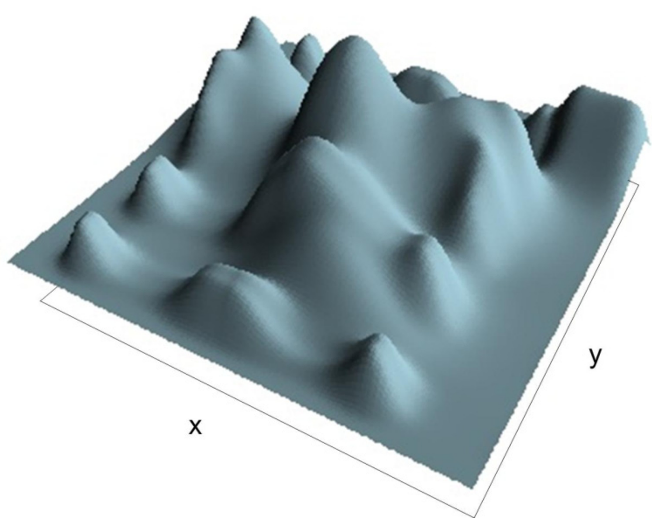


Figure 2.2: Actual attribute surface

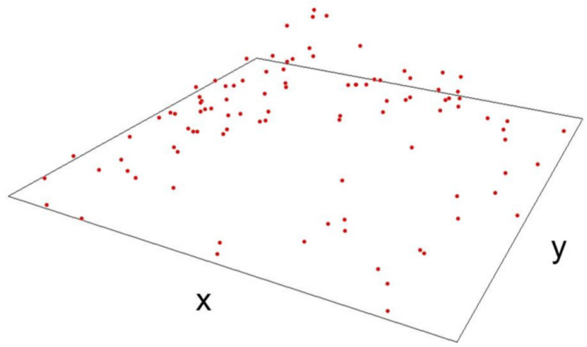


Figure 2.3: Illustrative point sample

Several spatial interpolation approaches, including IDW using inverse of distance square, Natural-neighbor, ordinary Kriging, regularized Spline and second order Trend surface, are applied to each of the 210 different sample point instances. Interpolation is carried out using ArcGIS 10.0 (ESRI). The fitted surfaces are then used to estimate attribute values at the 810,000 evaluation locations. Error assessment is based on the set of evaluation points, measured as the difference between the true value given by the actual surface and the estimated value given by the fitted surface. The average estimation error for each set of observations and each interpolation method are summarized in Table 2.1.

The errors vary in terms of interpolation method and point sample density. It is obvious in one respect that, for any interpolation technique, the errors decrease when the number of sampling points increases. For example, the IDW mean relative estimation error is 637.6% when using 100 sample points, but drops to 26.8% when using 5,000 sample points. It is not difficult to understand this trend since the more sample points, the more likely the characteristics of the underlying surface can be captured, implying potential to reduce the subsequent

Table 2.1: Relative mean error (%)

	Number of points						
Methods	100	200	300	400	500	1000	5000
IDW	637.6	372.4	290.6	232.8	190.9	93.1	26.8
	(300.7, 1158.2)	(238.9, 587.9)	(149.9, 403.1)	(139.6, 359.3)	(119.7, 300.1)	(65.3, 126.9)	(17.4, 41.3)
Natural-neighbor	149.7	110.3	87.8	72.8	60.8	35.3	8.2
	(84.9, 262.7)	(77.6, 199.1)	(64.4, 127.8)	(53.0, 108.1)	(44.7, 95.2)	(28.5, 46.3)	(7.5, 9.3)
Kriging	544.5	303.0	232.4	164.6	91.2	28.1	6.2
	(208.2, 1772.8)	(93.2, 947.5)	(62.0, 697.5)	(47.7, 637.0)	(40.5, 484.1)	(20.6, 37.2)	(5.7, 6.8)
Spline	413.9	168.0	104.0	75.3	52.8	20.3	3.7
	(123.8, 1075.7)	(65.4, 371.9)	(31.7, 291.0)	(24.5, 274.1)	(23.9, 186.6)	(12.6, 49.5)	(3.5, 4.1)
Trend	660.7	644.1	643.6	640.7	640.5	635.8	497.0
	(608.9, 1045.0)	(570.1, 792.1)	(509.5, 777.9)	(496.5, 773.5)	(492.4, 757.3)	(475.0, 668.2)	(437.2, 623.6)

Note: (min, max) relative error shown as well.

inference error (Burt *et al*. 2009). In fact, many interpolation approaches will generate similar fitted surfaces if very dense sample points are used. However, selecting an appropriate method is more critical if less dense points are relied upon (de Smith *et al*. 2009). Of course, it is often the case that fewer sample points are used in practice due to cost and time associated with sample data collection and processing (Berry and Baker 1968, Burt *et al*. 2009). The implications of this are very clear in Table 2.1.

In another respect, the errors differ among various interpolation approaches given the same number of sample points. For instance, when using 100 sample points, the largest relative mean error is about 660.7% for the Trend surface method, while Natural-neighbor produces a surface with the smallest relative mean error, 149.7%. Also, estimation errors vary among an interpolation technique. For example, the Trend surface approach gives the largest prediction error for any sample size, and remains consistently high even for larger point samples. The relative merits of different interpolation methods have been extensively studied (Lam 1983, Cressie 1993, Myers 1994, de Smith *et al*. 2009), so are not a focus here. However, Weber and Englund (1992) found IDW gave better results than those given by Kriging through the estimation of contaminant concentrations using multiple sample datasets, while Zimmerman (1999) concluded that ordinary/universal Kriging is superior to IDW by a designed experiment subjected to different surface types and sampling patterns. The fact remains that there is not an interpolation method that is perfect in all contexts, and the quality of an interpolation approach can be different depending on the underlying phenomena and the sample points used (Mitas and Mitasova 1999).

Regardless of the sample points or the interpolation techniques employed, the essential implication of Table 2.1 is that errors in approximated surfaces are always present, though in varying degrees. This leads to a critical issue of how such uncertainty and errors will influence any spatial analyses involving continuous data. How will errors propagate through analyses? The next section will further explore the implications of surface approximation error in spatial analysis using an empirical study.

2.5 Significance of Error Propagation

As discussed above, spatial interpolation approaches rely on sample points to fit a surface describing continuous fields. Due to the lack of complete information about the true distribution of continuous phenomena, precise estimation is often impossible. Instead, asymptotic theories are generally relied on to derive approximate results. The most straightforward way is to consider the sample size increasing to infinity to obtain asymptotic inference of a continuous distribution. A related asymptotic theory for spatial data is infill asymptotics (Cressie 1993) or fixed-domain asymptotics (Stein 1999), widely adopted in spatial interpolation for a fixed and bounded domain. According to infill asymptotics, using previous notation, we have:

$$\lim_{m \to \infty} \hat{g}(x, y) = g(x, y) \tag{2-4}$$

This implies:

$$\lim_{m \to \infty} E = 0 \quad (2\text{-}5)$$

That is, when the number of sample points increases to infinity, the estimation error tends to zero as more information reduces uncertainty in the inference of unknown attribute values (Stein 1999). In other words, a discrete set $(\alpha_1, \alpha_2, \alpha_3, ..., \alpha_m)$ with dense point samples can better describe a continuous distribution $f(\alpha_1, \alpha_2, \alpha_3, ..., \alpha_m)$ in an asymptotic sense. The implication for spatial analysis, however, remains unknown.

To explore this further, consider the continuous Weber problem. The Weber problem is one of the first proposed spatial analytical techniques, where the intent is to site a single facility in continuous space in order to serve a finite set of demand points at minimum total transportation cost (Wesolowsky 1993, Church and Murray 2009). The continuous Weber problem is an extension of the Weber problem involving continuously distributed demand. A mathematical formulation of the continuous Weber problem is given in Chapter 3. Current efforts to solve the continuous Weber problem generally employ some sort of mathematical function, $g(x, y)$, to represent the distribution of continuous demand, and focus on developing advanced optimization algorithms to solve this problem (Carrizosa *et al*. 1998, Fekete *et al*. 2005, Church and Murray 2009, Drezner and Suzuki 2010, Murat *et al*. 2010). Of course, the actual continuous surface is never simply known or given with certainty. Instead, it is usually fitted using a point sample, so $\hat{g}(x, y)$ is actually relied upon in practice. Errors are

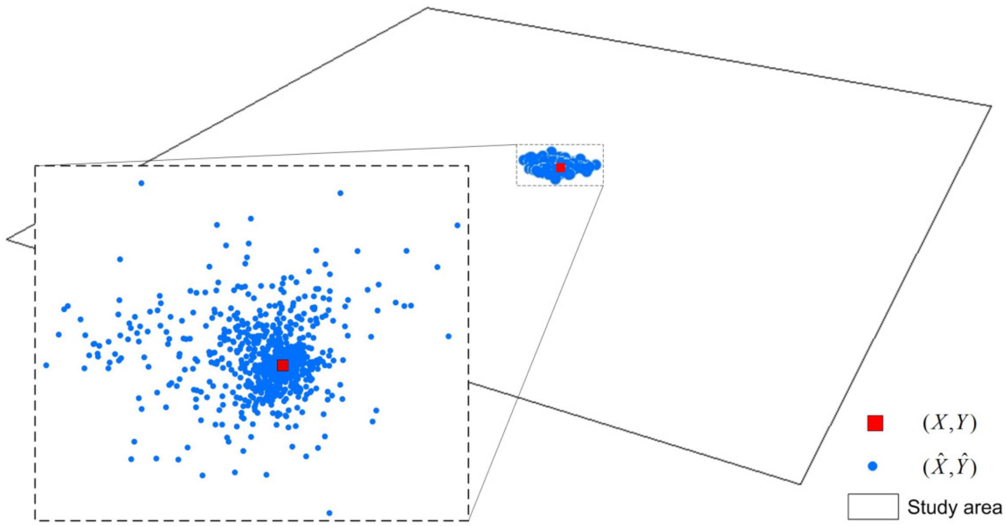


Figure 2.4: Weber solutions for approximated surfaces compared with the actual optimal location

therefore unavoidable, and may potentially affect the analysis and decision making in various ways.

In this study, the continuous Weber problem is solved using the approach detailed in Chapter 3. Comparison is made between the actual optimal facility location (X,Y) for $g(x,y)$ and that obtained using $\hat{g}(x,y)$, $(\hat{X},\hat{Y})$. Figure 2.4 shows the distribution of $(\hat{X},\hat{Y})$ obtained using the previously detailed fitted surfaces relative to the actual optimal location (X,Y). As is evident in Figure 2.4, the approximated locations $(\hat{X},\hat{Y})$ derived from the fitted surfaces are littered around the actual optimal location. Figure 2.5 illustrates that the distribution of the $(\hat{X},\hat{Y})$ solutions varies for the different approximated surfaces based on the point sample size. By visual inspection, it is not difficult to observe trends. Specifically, for smaller samples (e.g., 100 or 200), the derived locations $(\hat{X},\hat{Y})$

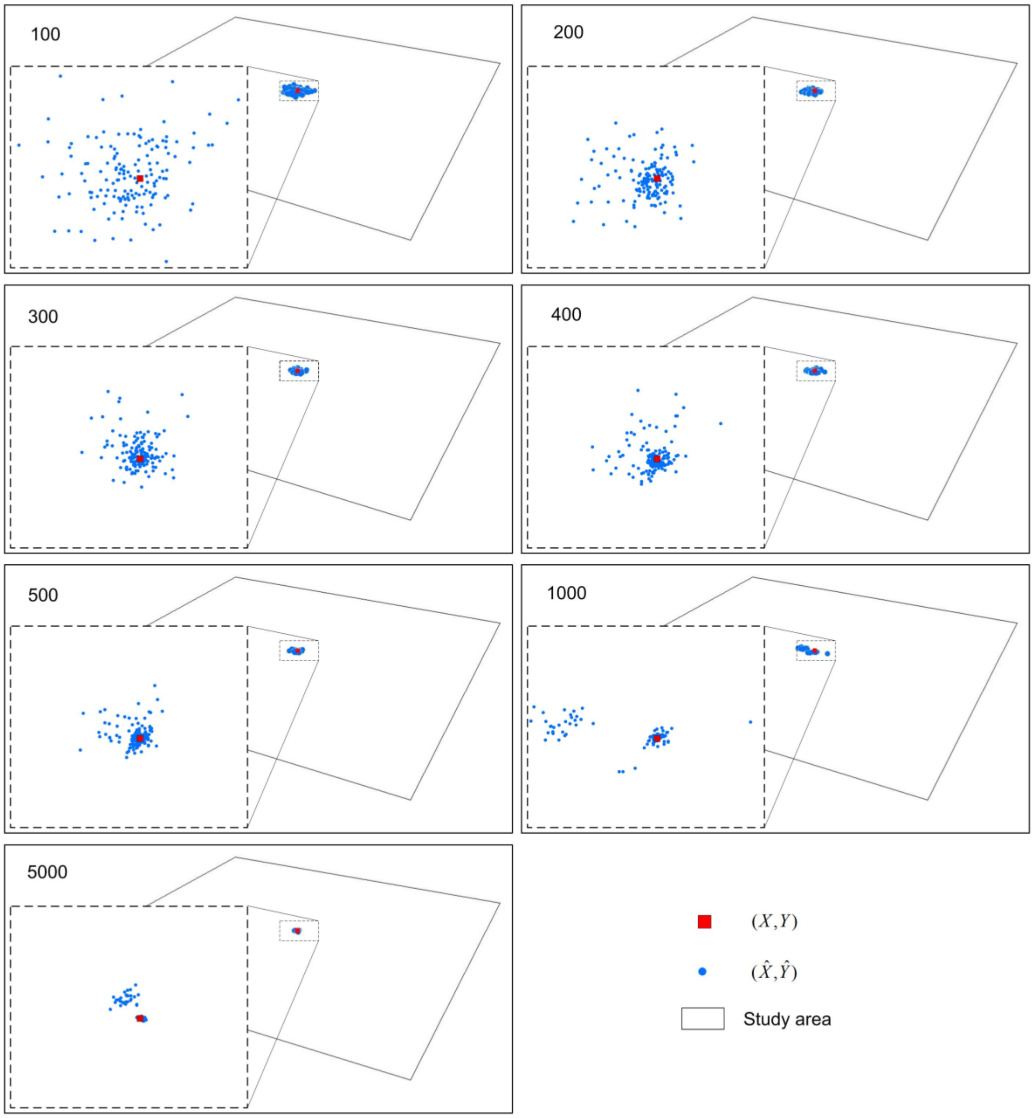


Figure 2.5: Weber solutions by surface sample size

for the approximated surfaces $\hat{g}(x,y)$ are more dispersed, while for larger sample sizes (i.e., 5000) the derived locations $(\hat{X},\hat{Y})$ are more compact and closer to the true optimum.

The spatial variability shown in Figures 2.4 and 2.5 is significant as well. If the derived locations $(\hat{X},\hat{Y})$ are evaluated using the actual surface $g(x,y)$, the objective value can be computed and the associated error can be measured. The error is the difference between the optimal objective value based on (X,Y) and that for $(\hat{X},\hat{Y})$, divided by the former. Table 2.2 gives the objective value percent errors. While in general the error decreases when larger point samples are used, the distribution of error varies among interpolation techniques. For example, the surfaces fitted by the Spline method generates the smallest mean error in objective value, whereas those from the Trend surface approach always produce the largest mean error, regardless of the sample size. Further, due to different sample distributions, errors differ even for an interpolation technique when using the same sample size. The Spline method with sample size 1000 is one example. The maximum relative error in Table 2.2 for this case is 2.7%, 27 times higher than the mean value of 0.1%. Though the average errors may seem relatively small in Table 2.2, the observed maximums are generally a concern in all cases except when the number of sample points is large.

The results shown in Figure 2.4 and Table 2.2 indicate that errors contained in approximated surfaces absolutely affect spatial analysis. In this case the spatial analysis is erroneous, both in terms of resultant location as well as its associated efficiency. That is, the identified locations using the approximated surface $\hat{g}(x,y)$ are not the same or equivalent to the actual optimal location associated with the actual surface $g(x,y)$. Further, the locational errors are significant because they are highly inefficient as measured by the continuous Weber objective function. Thus, error is propagated through spatial analysis making any findings questionable and uncertain.

Table 2.2: Relative mean errors (%) in objective values for the continuous Weber problem

Methods	Number of points						
	100	200	300	400	500	1000	5000
IDW	26.0	7.5	3.9	2.6	1.5	0.7	0.1
	(0.6,140.3)	(0.2,26.3)	(0.0,14.0)	(0.1,10.5)	(0.0,5.1)	(0.0,2.8)	(0.0,0.3)
Natural-neighbor	12.7	2.2	1.5	1.0	0.6	0.1	0.0
	(0.3,52.0)	(0.0,7.1)	(0.0,4.5)	(0.0,4.0)	(0.0,2.8)	(0.0,0.4)	(0.0,0.0)
Kriging	17.4	6.8	3.4	1.6	0.7	0.1	0.0
	(0.4,107.5)	(0.3,51.6)	(0.1,29.2)	(0.0,15.8)	(0.0,8.3)	(0.0,0.2)	(0.0,0.0)
Spline	12.6	2.2	0.7	0.6	0.2	0.1	0.0
	(0.2,87.8)	(0.1,10.8)	(0.0,4.4)	(0.0,3.3)	(0.0,1.0)	(0.0,2.7)	(0.0,0.0)
Trend	46.9	25.0	18.4	15.4	10.4	8.8	6.2
	(4.0,153.8)	(0.8,81.5)	(0.5,54.2)	(1.8,47.5)	(0.2,34.7)	(3.9,15.9)	(1.8,10.8)

Note: (min, max) relative error shown as well.

2.6 Discussion and Conclusions

Spatial analysis is greatly dependent on how we represent geographic phenomena. Discrete objects and continuous fields are two distinct views for describing geographical space. Continuous surfaces are usually relied upon for a range of spatial analyses. Though it is possible to depict a continuous surface using an exact mathematical function, this is not typically possible in practice. Instead, approximated surfaces derived using spatial interpolation are usually adopted, but are subject to errors and uncertainty. As a result, a fitted surface presents problems for spatial analytical methods that assume error free inputs.

In this study, a designed experiment is employed to investigate prediction errors in approximated surfaces. Though varying in terms of sample sizes and interpolation approaches, the presence of errors is ubiquitous. Of course, the errors are influenced by many factors, such as sample size, interpolation technique and sample schemes. In general, the error decreases as the sample size increases. Also, various interpolation approaches have different assumptions that are not necessarily satisfied by the sample data. Finally, sample schemes influence the distribution of sample points, which contribute to distribution errors even when the sample size and the interpolation method are the same.

When approximate surfaces are used as inputs, errors propagate through spatial analysis. Application results based on the continuous Weber problem demonstrated that analysis and derived results are significantly affected by approximated surfaces. Specifically, errors are present in derived facility locations and associated objective values. The implication is that the identified location can be far away from the actual optimal location, which means that total transportation costs would be substantially higher than they would be if properly sited, and this is due solely to erroneous input data.

Spatial analysis requires a reliable representation of geographic space. It is highly problematic to assume that a continuous surface $g(x, y)$ is known and accurate in any planning and analysis context. This research shows that errors in an approximated continuous surface based on sample points are inevitable, which leads to cumulative errors and uncertainty in spatial analysis. It appears essential then that future research focus on developing new methods that take into account the errors inherent in approximated surfaces in order to reduce their impacts on spatial analytical results. Work by Kyriakidis and Goodchild (2006) provides a capacity to understand and quantify surface errors. Making use of this information in spatial analytical methods remains the challenge.

Chapter 3

THE CONTINUOUS WEBER PROBLEM[†]

3.1 Introduction

Our daily life relies on various public and private service facilities. For example, we need schools for education, supermarkets to purchase groceries, public transit stations to connect different transportation modes, fire stations to respond to emergencies, and others as well. Decisions related to where those services should best be located are therefore essential. As an example, for the well-being of a region, hospitals must be well placed in order to maximize accessibility for medical care and emergencies, and distribution centers should be appropriately located to minimize shipment costs. Where these facilities are located can affect the quality of services provided and profitability. Location modeling therefore continues to be important in urban and regional planning as well as other socio-economic contexts.

Of particular interest in this paper is a location planning problem that concerns siting a single facility to serve continuously distributed demand across a region at minimum total transportation cost. A special case is the Weber problem, having received considerable attention since first proposed in the 17th century (Wesolowsky 1993, Drezner *et al*. 2002). This seemingly simple problem has attracted a lot of interest largely attributed to its broad applicability and potential for extension (Drezner *et al*. 2002). Consequently, many location models can be

---

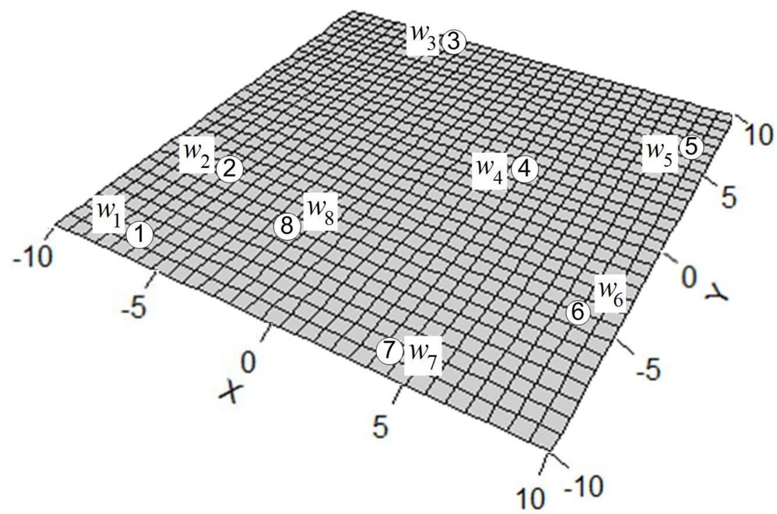
[†] A modified version of this chapter has been submitted for publication, co-authored with Alan Murray.

traced back to the Weber problem, such as the continuous location-allocation problem involving multiple facilities (e.g. Cooper 1963, 1964) and the p-median problem given fixed locations for potential facilities (e.g. Hakimi 1964).

Model formulation of this problem generally requires a geographic representation of demand. Originally, demand in the Weber problem is conceived to be a finite set of fixed points. It is appropriate when demand is in fact discrete, but in practice it is likely that the underlying demand is continuously distributed over space. Figure 3.1 illustrates this contrast between discrete and continuous demand in a region. Figure 3.1a shows eight fixed points abstracted to represent demand across the region, each with an associated demand weight or value. Figure 3.1b depicts demand continuously distributed over space. In fact, many problems in location research involve continuous demand, such as Hotelling's problem seeking to locate two competing suppliers to serve continuous uniform demand (Hotelling 1929, Ghosh and Craig 1984) or the design of supply chain networks (Bhattacharya and Bandyopadhyay 2010, Tsao and Lu 2012). Errors are inevitably introduced into any location model relying on a discrete demand point abstraction when it is actually continuously distributed (Murray 2003, Francis *et al*. 2009). Therefore, it is essential that demand be represented continuously in such a case, but this complicates both model formulation as well as model solution.

While much attention has been devoted to the Weber problem with continuous demand (Love 1972, Bennett and Mirakhor 1974, Drezner and Wesolowsky 1980, Drezner and Drezner 1997, Wang et al. 1997, Carrizosa *et al*.

(a)



(b)

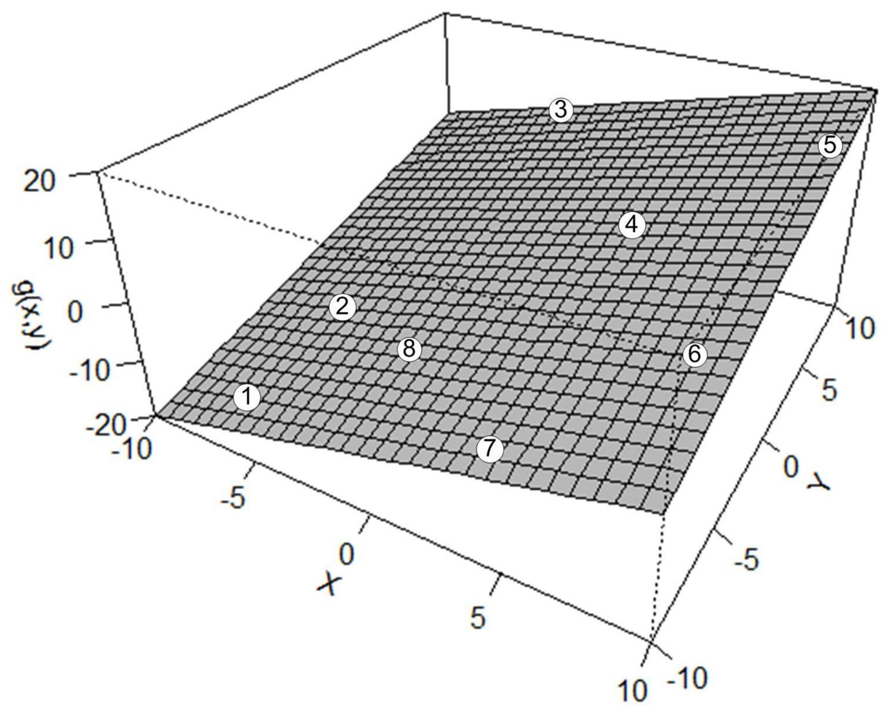


Figure 3.1: Demand representations:

(a) Discrete demand and (b) Continuous demand $g(x, y) = x + y$

1998), the complexity of the objective function and computational difficulty has necessitated simplification in demand region shape, the distribution of demand or travel distance. Model solutions for the continuous Weber problem are either problem specific or not computationally efficient. Developing an effective solution approach for the Weber problem involving continuous demand therefore remains a major research challenge.

In this chapter we are interested in solving the Weber problem with continuously distributed demand where the facility may be located anywhere so as to minimize the average user distance to access the facility, hereafter referred to simply as the continuous Weber problem. The next section reviews existing research related to the continuous Weber problem. Problem specification and spatial representation are then covered. An approach is proposed to site a single facility to best serve continuously distributed demand in a region. This is followed by an empirical evaluation and assessment. The final section gives discussion and conclusions.

3.2 Background

The Weber problem was originally proposed for locating a factory with the intent of minimizing transportation costs. A range of disciplines, such as mathematics, geography, industrial engineering, regional science and transportation, among others, have studied this problem, each referring to it differently. Common references include: the Weber problem, the Fermat problem, Torricelli point and the minisum problem. Recent reviews can be found in Wesolowsky (1993) and Drezner *et al*. (2002).

As noted previously, demand is assumed to be a finite set of fixed points. To achieve this, data aggregation involving the abstraction of geographical space as a finite set of points is commonly relied upon. This is done because it is often how space is perceived in the most simplistic terms but also because of publicly accessible data, like that from the Census (Miller 1996, Murray 2003, Church and Murray 2009). Further, such discretization offers computational efficiencies as well. Miller (1996) addressed the importance of adopting non-point based demand representations in location modeling, but others too have recognized this need. Wesolowsky (1977) employed density functions to model a probability distribution of demand over space. Love (1972) and Drezner and Wesolowsky (1978) represented demand as regions rather than points, where demand is considered continuously distributed over the study area. While a point based simplification is attractive, errors and uncertainty are introduced into the objective function, and ultimately the solution (Drezner and Wesolowsky 1978; Hillman and Rhoda 1978; Vaughan 1984) when demand is actually continuously distributed.

Attempts to get around this complication abound in the literature. Instead of using distance between a point-based proxy and a facility, an alternative is to approximate it by the average distance to the demand area from the facility. For example, assuming uniform demand, the "distance correction" approach proposed by Drezner (1995) applied the average distance of a circle to study a competitive facility location problem in a rectangular demand area. Further, Carrizosa *et al*. (1998) approximated non-uniformly distributed demand by replacing the demand area with disks, which in turn are substituted with triangles. Rather than exploring a single value of expected distance, Carmi *et al*. (2005) proved that for a convex demand region the average distance from the optimal facility location is within $[d/7, d/7]$, where d is the region diameter. Abu-Affash and Katz (2009) improved the bounds to $[4d/25, 2d/3\sqrt{3}]$. More recently, Puerto and Rodríguez-Chía (2011) obtained the geometrical characterizations of the entire set of optimal solutions given demand defined by some probability distribution. Algorithms to solve the continuous Weber problem by approximating average distance are relatively easy to implement for regular demand regions such as circles (Koshizuka and Kurita 1991) and rectangles (Love 1972, Vaughan 1984), but cumbersome otherwise.

Another group of approaches solve a double integral representing continuous demand directly using numerical procedures. Drezner and Wesolowsky (1980), Drezner and Drezner (1997), Wang *et al*. (1997), Chen (2001) and Franco *et al*. (2008) extended the Weiszfeld algorithm given the differentiability of the objective function associated with continuous demand. Fekete *et al*. (2005) proposed an exact algorithm relying on computational geometry for rectilinear distances. Gugat and Pfeiffer (2007) derived bounds on the optimal objective value for the Weber problem with continuous regional demand traveling along a network. Church and Murray (2009) discussed an iterative enumeration process. These approaches, again, are only appropriate for certain regularly shaped demand areas, distance norms or assumed convex solution surfaces.

Work on the continuous location-allocation problem (LAP) or the continuous Weber problem involving multi-facilities by Dasci and Verter (2001, 2005), Ouyang and Daganzo (2006) and Murat *et al*. (2010, 2011) is also relevant. Assumptions central to work in this area are approximate uniform demand (e.g. Dasci and Verter 2001, 2005) or a non-homogeneous demand distribution specified by exact, known mathematical function (e.g. Ouyang and Daganzo 2006; Murat *et al*. 2010, 2011). For example, Murat *et al*. (2011) developed a multi-dimensional shooting algorithm to solve a continuous LAP which can tackle dense demand within non-convex regions. It was found to be more efficient than a steepest descent algorithm proposed earlier by Murat *et al*. (2010). However, the algorithm is limited to two facilities and specific demand distribution functions.

As discussed above, methods for solving the continuous Weber problem to date are developed for problem instances under certain assumptions, either for regular demand area shapes, rectilinear distances, or specific density functions. The efficiency of applying these approaches to irregular regions as well as demand that is generally distributed remains unknown. The objective of this chapter is to develop an efficient and general solution procedure for the continuous Weber problem that is not restricted by assumed demand distributions or regional shape.

3.3 Problem Specification

When demand is represented by discrete points, the Weber problem objective function can be readily structured to reflect minimizing the sum of weighted distances over a finite set of demand points. The following notation is adopted in the model formulation:

i = index of demand points

(x_i, y_i) = location of demand point i

n = number of demand points

w_i = weight of demand associated with point i

(X, Y) = facility location

The Weber problem with discrete demand using the Euclidean distance norm is as follows (Wesolowsky 1993):

$$\textit{Minimize} \qquad \sum_{i=1}^{n} w_i \sqrt{(X - x_i)^2 + (Y - y_i)^2} \qquad (3\text{-}1)$$

The goal is to find the best location, defined by variables *X* and *Y*, that minimizes the total weighted distance demand is from the sited facility. This is equivalent to minimizing average distance. By definition, the facility location may be anywhere in continuous space.

Many algorithms have been developed to solve the Weber problem, including early efforts like the "Torricelli Point" for the unweighted problem with three demand points and the Varignon Frame for a finite set of weighted demand points. The Weiszfeld (1936) algorithm has proven to be most popular as it utilizes the first order derivative of the objective function in an iterative manner.

Similar approaches are also investigated in Miehle (1958), Kuhn and Kuenne (1962) and Cooper (1963, 1964). The convergence of this algorithm was proven by Kuhn (1973), among others. However, a limitation of this approach is that the objective function is not differentiable at the demand points, so it could fail if the optimal solution coincides with one of the demand points (Kuhn 1973, Chandrasekaran and Tamir 1989, Wesolowsky 1993, Church and Murray 2009). Accounting for this is therefore necessary, and can be readily done (Love and Yeong 1981, Church and Murray 2009). Isodapanes, which are lines representing equal transportation cost, can also be used to solve the Weber problem (Weber 1909, Hoover 1937). The solution procedure usually involves composite overlay of multiple distance layers, one layer for the distance to each demand point. Though this approach can be readily integrated in GIS, challenges with the computational requirements remain as this is excessive if the number of demand points is large, particularly compared to Weiszfeld algorithm.

When demand is continuously distributed, the objective function remains to minimize the average distance from the facility location, but it is more complicated to formalize. Consider the following additional notation:

$g(x, y)$ = function of demand at point (x, y)

$d(x, y, X, Y) = \sqrt{(X-x)^2 + (Y-y)^2}$

R = region of demand

The Weber problem with continuous demand can therefore be expressed as follows given a single demand area or region (Church and Murray 2009):

$$Minimize \quad \iint_{(x,y)\in R} g(x,y)d(x,y,X,Y)dxdy \tag{3-2}$$

The sum in the discrete version of the Weber problem, (3-1), is replaced in (3-2) by a double integral over the demand area, making the model with continuously distributed demand a stochastic optimization problem that turns out to be quite difficult to solve in practice.

The difference between (3-1) and (3-2) is evident in Figure 3.1, where the points shown in Figure 3.1a reflect the input necessary for the discrete Weber problem in (3-1) and the actual surface in Figure 3.1b shows the continuously varying demand across the area modeled in (3-2). As noted previously, a number of approaches have been explored for solving the Weber problem with continuously distributed demand, (3-2). Unfortunately all rely on fairly restrictive simplification assumptions, making their practical use and applicability rather limited.

3.4 Spatial Representation

In a digital environment, representation of geographic space is challenging, with many options for carrying this out. At issue is whether demand is most appropriately viewed as a field or an object. Physical features that can be considered as discrete objects include roads, trees, houses, etc. Phenomena like rainfall amount, air quality, elevation, temperature and land use are often thought of as continuous fields. As a result, there are basically two spatial representation methods relied on in commercial GIS to reflect these perspectives (Goodchild 1992): vector and raster. In a vector representation, geographical features are described by geometry: point, line and polygon. Each vector object may have one or more attributes, like expected demand. Figure 3.1a is an example of a point representation of demand where each point has a unique demand value or weight. The raster based method represents geographical space using a continuous tessellation consisting of square cells and each raster cell has an associated attribute(s)[‡]. Figure 3.1b displays continuous demand as a raster surface.

Vector and raster representations have advantages and disadvantages (Longley *et al*. 2011). A raster is a simple data structure, making it less computationally expensive for certain types of spatial processing. A limitation, however, is that the attribute within a cell is approximated as a single value, so accuracy is dependent on the resolution of the cell. A vector representation usually requires less data storage overall, but for some spatial operations computational processing can be high. Of course, the choice of representation depends on the spatial phenomena under study and the analysis context, but also data availability, data accuracy, computational efficiency, etc. among which geographic scale is a key element to be considered. For example, Miller (1996) discusses geometric representation related to different spatial scales in facility location problems and Murray *et al*. (2008) explored the uncertainty in coverage solutions obtained using different spatial representations of continuous demand space.

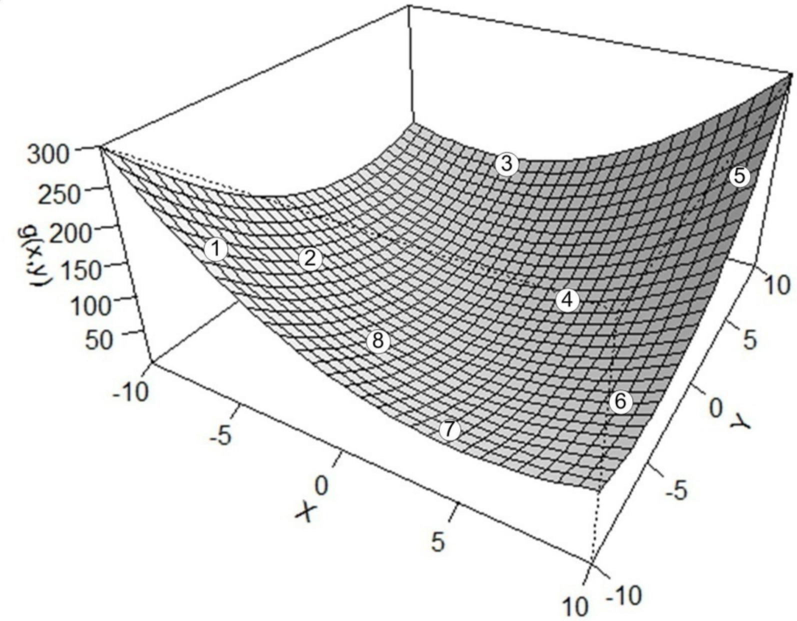
[‡] In general, a tessellation may be regular or irregular, and consists of squares (cells) or rectangles as well as triangles and hexagons (Goodchild 1992, Church and Murray 2009, Longley *et al*. 2011). Discussion is limited here to a regular raster for simplification, but any representation can be considered.

Discrete demand points in location models reflect an object perspective, whereas a continuous surface representing demand varying over space is a field view. The benefit of the field view is that it accounts for specific demand at any location (x, y) in continuous space. It is theoretically possible to represent a continuous surface as a mathematical function, $g(x, y)$. While an infinite number of functional forms are possible, three continuous demand surfaces given by first, second and third order mathematical functions are shown in Figures 3.1b and 3.2, where $g(x, y) = x + y$ in Figure 3.1b, $g(x, y) = x^2 + xy + y^2$ in Figure 3.2a and $g(x, y) = x^3 + x^2 y + xy^2 + y^3$ in Figure 3.2b. The issue that arises with these or any other functions, however, is that $g(x, y)$ must either be known or be fitted. Of course, the demand function is never simply known or given with certainty, so in practice it is usually fitted through spatial sampling in order to characterize the underlying geographic distribution.

There are a number of techniques available in GIS that can be used to fit or approximate continuous surfaces, usually referred to as spatial interpolation approaches. Spatial interpolation is to estimate unknown attribute values in geographic space using the known attribute values at measured locations, generating a tessellation across the study region (Lam 1983, Cressie 1993). Suppose $g(x, y)$ is the true attribute value at a location (x, y), the goal of spatial interpolation then can be considered as the estimation of $g(x, y)$ for all locations (x, y) with as little error as possible. Consider the following notation:

k = index of sample locations

(a)



(b)

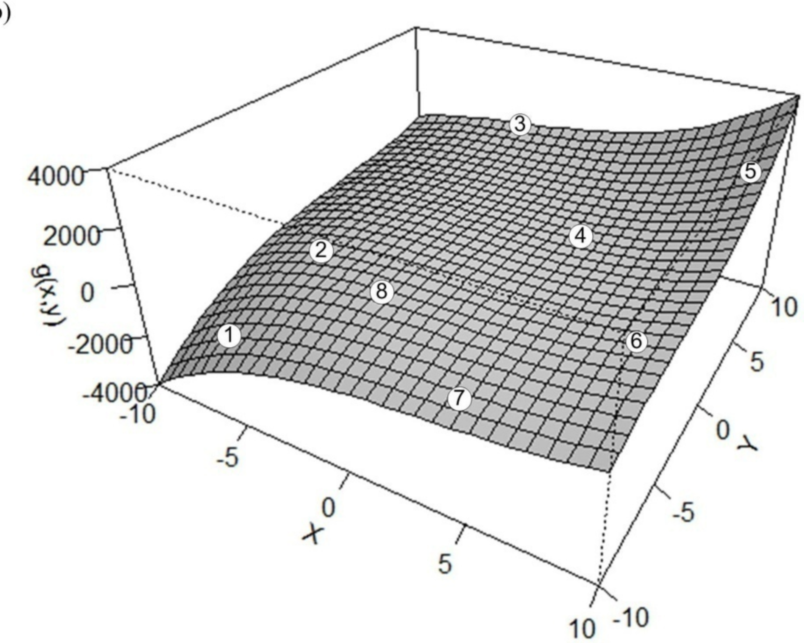


Figure 3.2: Alternative continuous demand representations:
(a) $g(x, y) = x^2 + xy + y^2$ and (b) $g(x, y) = x^3 + x^2 y + xy^2 + y^3$

(x_k, y_k) = location of sample k

α_k = observed attribute value of sample k

m = number of samples

$g(x, y)$ = true attribute value at location (x, y)

$\hat{g}(x, y)$ = estimated attribute value at location (x, y)

$\varepsilon(x, y)$ = estimation error at location (x, y)

Thus, the estimated attribute value, $\hat{g}(x, y)$, can be expressed as a function of the observed sample data as follows:

$$\hat{g}(x, y) = f(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_m) \tag{3-3}$$

where f is some function. Generally, the goal is minimization of the estimation errors which can be formally written as:

$$Minimize \quad E = \iint_{(x,y)\in R} \varepsilon(x, y)\,dxdy \tag{3-4}$$

where $\varepsilon(x, y) = \left| g(x, y) - \hat{g}(x, y) \right|$.

There are generally two types of interpolation techniques to support and fit a continuous surface, point and areal. Examples of point interpolation methods include inverse distance weighting (IDW), nearest-neighbor, splines, polynomial regression, Kriging and others. For example, Weber and Englund (1992) employed IDW and Kriging approaches to estimate the contaminant concentrations using multiple sample datasets. Janssen *et al*. (2008) used a Kriging-based model for the approximation of air quality surfaces. Areal interpolation, in contrast, estimates attribute values for areas within a regions, typically based on another set of reporting zones for which boundaries do not coincide (Goodchild *et al*. 1993). Areal weighting and dasymetric mapping are two common areal interpolation approaches. The latter is more sophisticated in terms of the adoption of ancillary data and has been widely applied in population density estimation (Langford and Unwin 1994, Eicher and Brewer 2001, Mennis 2003).

An issue, however, is that estimation is uncertain, with significant error and variance (Lam 1983, Goodchild *et al*. 1993). This is due to the quality of interpolation, influenced by many factors, such as the actual distribution, spatial sample and measurement accuracy as well as assumptions employed. Many efforts have been devoted to improve spatial interpolation methods with the aid of ancillary information or supplementary datasets (e.g. Wu and Murray 2007). However, errors and uncertainty are inevitable. For instance, density surfaces derived from dasymetric mapping are subject to satellite image misclassification. Therefore, it is always impossible to get a known $g(x, y)$ with certainty, and thus any estimation, $\hat{g}(x, y)$, will no doubt be subject to error, which indicates $\varepsilon(x, y) > 0$ for most locations (x, y). This means that cumulative error and uncertainty is significant (Tomczak 1998), and subsequent facility location siting based on continuous demand surfaces represented by either assumed $g(x, y)$ or approximated $\hat{g}(x, y)$ surfaces would likely be uncertain, erroneous or biased in many ways.

3.5 Solution Approach

It is obvious that in practice surfaces representing continuous demand are necessarily approximated by spatial interpolation methods using sample data. Thus, it is often impossible to obtain accurate estimation because the complete knowledge about the true underlying demand distribution is usually unavailable. Therefore, asymptotic theories are generally employed to acquire approximate results rather than precise estimation. In terms of continuous representation, the most straightforward way is to increase the sample size to infinity since more information from larger sample would increase the accuracy in the estimation of unknown attribute values (Stein 1999). Infill asymptotics (Cressie 1993), or fixed-domain asymptotics (Stein 1999) is such a theory that is widely applied in spatial data analysis, particularly in spatial interpolation for a fixed and bounded region. According to infill asymptotics, the estimated attribute value approaches the true attribute value when the sample size increases to infinity, as defined by (3-5):

$$\lim_{m \to \infty} \hat{g}(x, y) = g(x, y) \tag{3-5}$$

Based on the expression in (3-4), (3-5) is equivalent to:

$$\lim_{m \to \infty} E = 0 \tag{3-6}$$

That is, the estimation error approximates zero when the number of sample points goes to infinity. As the volume of sample data increases, point based spatial interpolation methods will theoretically provide better estimations of attribute values at unsampled locations.

In the context of location modeling, although a discrete representation of continuous demand as a finite set of points can introduce errors into model solutions (Murray 2003, Francis *et al*. 2009), an improved approximation of continuous demand occurs as $n \to \infty$. The assumption is that with an increase in the density of demand points, the set better approximates the actual continuous distribution. The implication is that a solution to the discrete Weber problem with greater demand point density will approach the theoretical optimum for the continuous Weber problem. Computationally, however, the hope would be that some finite number of demand points is sufficient. Therefore, in this study a finite set of fixed points is used to represent continuously distributed demand to explore this theoretical convergence. Simplification of continuous demand as a discrete sample set is anticipated to be less restrictive and less problem altering compared to existing approaches. To explore the relationship between point density and the accuracy of facility locations derived from discrete representations, point density is systematically increased to better approximate continuous demand.

Based on a point-based demand representation, the proposed process for solving the continuous Weber problem is outlined in Figure 3.3, incorporating the following parameters:

INIT_D: initial point density

NUM_R: number of layers for each density

STEP_D: step size of density

Γ : number of layers to consider in examining convergence of objective value

τ : tolerance used to determine convergence of the objective value

l : point layer

t : index tracking number of layers of same density

D : density of point representation

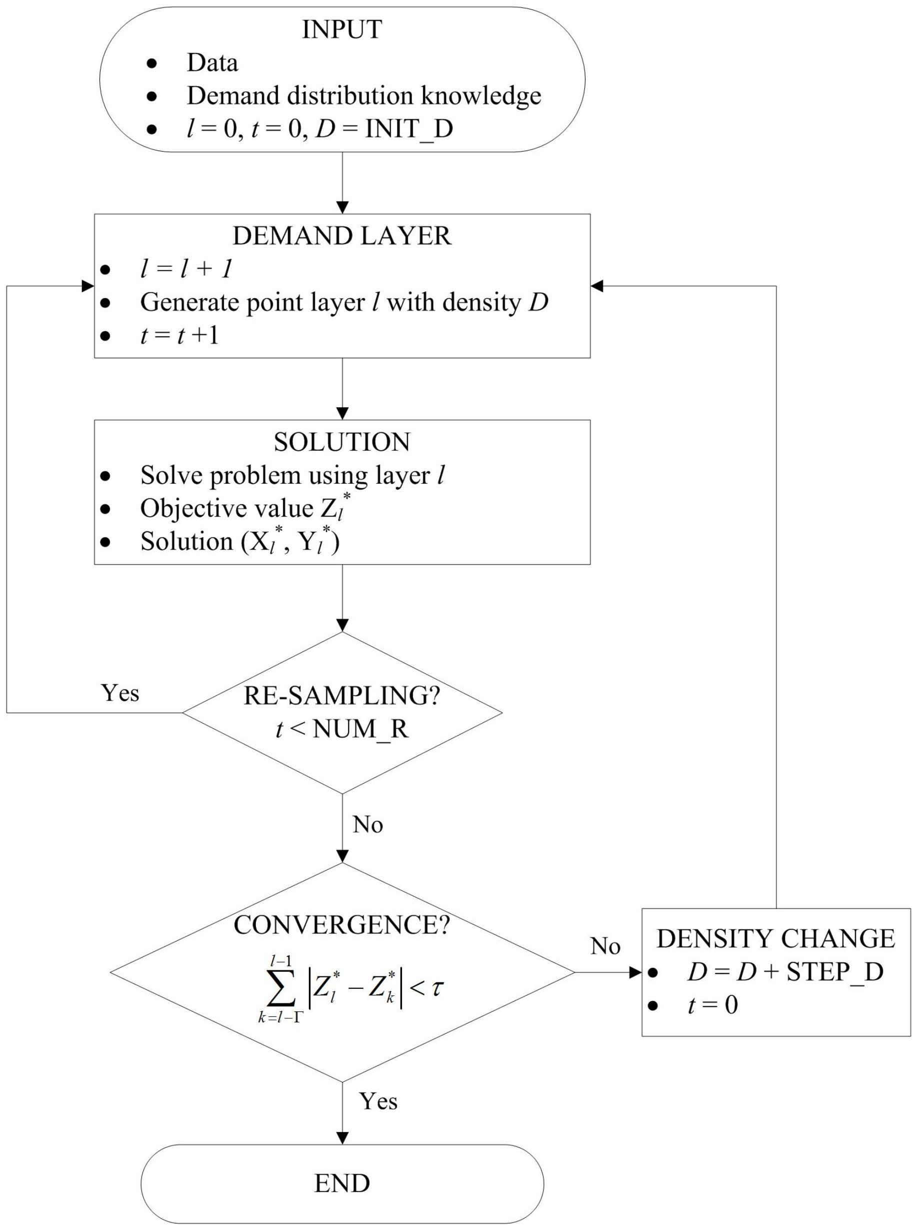


Figure 3.3: Solution approach for the continuous Weber problem

Following input of data and initialization (INPUT), Figure 3.3 then proceeds to generating a discrete demand layer that reflects continuous demand density for a region (DEMAND LAYER). This represents a demand approximation, $\hat{g}_l(x, y)$, for the actual continuously distributed demand. The infill sampling process can be carried out with readily available GIS software such as ArcGIS (ESRI, Redlands, California). Spatial sampling schemes employed can be random, stratified, clustered or contour sampling (Longley *et al*. 2011), depending on knowledge of the underlying demand distribution, if any. Figure 3.4 gives an example of demand layer generation, where initially the two sub-regions comprising the entire demand area have 4 and 2 points, respectively. With iteration proceeding, the number of points is systematically increased, keeping the proportion of demand densities among the two sub-regions constant.

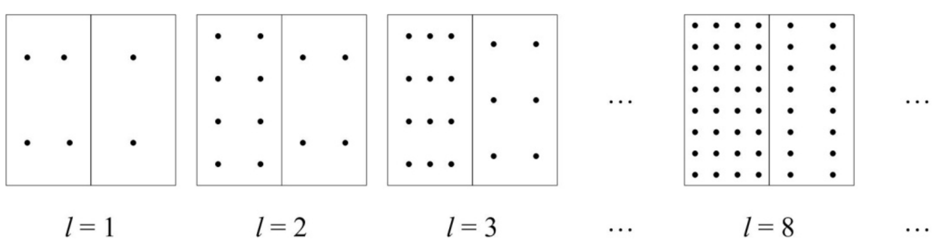


Figure 3.4: Process of demand layer generation

Given a demand layer l, the associated Weber problem is solved (SOLUTION). In this study, solution is obtained using the Weiszfeld algorithm[§] which relies on an iterative process to improve a current solution. Let (X^k, Y^k)

[§] Any other solution approaches could be used to solve the Weber problem with discrete point demand. As well, other optimization models involving continuous demand could be considered, too. Discussion is limited to the continuous Weber problem solved by Weiszfeld algorithm.

and (X^{k+1}, Y^{k+1}) be the facility location in the k th and $(k+1)$ th iteration. The Weiszfeld procedure can be formalized as:

$$X^{k+1} = \frac{\sum_{i=1}^{n} \frac{w_i x_i}{d_i}}{\sum_{i=1}^{n} \frac{w_i}{d_i}} \qquad Y^{k+1} = \frac{\sum_{i=1}^{n} \frac{w_i y_i}{d_i}}{\sum_{i=1}^{n} \frac{w_i}{d_i}} \tag{3-7}$$

where $d_i = \sqrt{(X^k - x_i)^2 + (Y^k - y_i)^2}$ using Euclidean norm**. If the set Ω is all of the discrete representations generated, then any representation $l \in \Omega$ has an optimal facility location, (X_l^*, Y_l^*), and a corresponding objective value, Z_l^*, identified by this solution technique. In this case, Z_l^* is calculated as the average travel distance, that is, $Z_l^* = \sum_{i=1}^{n} w_i d_i / \sum_{i=1}^{n} w_i$.

Figure 3.3 suggests that once a demand layer solution has been obtained, re-sampling can take place (RE-SAMPLING?), provided that t<NUM_R. Assuming so, the new demand layer is created, then solved. Otherwise, a test for convergence is undertaken (CONVERGENCE?). The test for convergence is based on the last Γ layers solved. The difference between the current solution Z_l^* and a previous solution Z_k^* is calculated for each layer $k \in \Omega$ where $(l - \Gamma) \le k \le (l - 1)$. If the sum of those differences is smaller than the predefined tolerance τ , the procedure terminates with an approximated optimal solution

---

** Here we use $d_i = \sqrt{(X^k - x_i)^2 + (Y^k - y_i)^2} + \sigma$, where σ is a small value, in order to avoid issues when a current location is at a demand point (see Wesolowsky and Love 1972, Church and Murray 2009).

(END). Otherwise, demand density is increased (DENSITY CHANGE) and demand layer generation and solution are repeated.

Again, the goal is to find the true optimal location (X^*, Y^*) for $g(x, y)$. Essential here is the proper representation of continuous demand, which is achieved in this case by a series of approximated demands surfaces, $\hat{g}_l(x, y)$, based on which the Weber problem is solved, giving (X_l^*, Y_l^*) as an approximation of the optimal location. According to infill asymptotics, the accuracy of (X_l^*, Y_l^*) will improve with increased demand point density. Thus, as $n \to \infty$, $Z_l^* \to Z^*$ and $(X_l^*, Y_l^*) \to (X^*, Y^*)$, if the solution space is convex. In other words, convergence to the optimal solution is achieved as point density increases. In practice, however, it is hoped that only a finite number of demand points are needed, and the iterative process will terminate in a reasonable amount of processing time.

3.6 Empirical Results

To solve the continuous Weber problem, two variants of continuous demand are explored using the proposed approach. The first case considers unweighted demand distributed uniformly in an irregularly shaped region. The second case examines an application where demand varies across the region.

3.6.1 Uniformly distributed demand

Figure 3.5 shows three Census tract regions: “rectangular”, “circular” and “concave”. The three regions vary in shape and there is no ancillary information about the underlying demand distribution. In such a situation, it is necessary to assume demand to be uniformly distributed as there is no information to indicate otherwise. In this case, demand is represented by equally spaced points and can be easily implemented using discretization functionality in GIS. To solve the continuous Weber problem for these regions, the parameters of proposed solution procedure are as follows: INIT_D = STEP_D = 100 points/ km^2, NUM_R = 1, Γ = 5 and τ = 0.5. The process is carried out on a Mac OS X system with 3.06 GHz intel Core 2 Duo processor and 4GB, 800 MHz memory.

Convergence is achieved rather quickly for the “rectangular” and “circular” regions ($|\Omega| = 14$ and 5 seconds, and $|\Omega| = 88$ and 16.2 minutes, respectively), and a little longer for the “concave” region ($|\Omega| = 144$ and 1.8 hours). Figure 3.5 summarizes the optimal facility location, (X_l^*, Y_l^*), for each layer $l \in \Omega$. As the number of points increases, they appear to converge to a single point – the theoretical optimal location. For the “circular” region, Figure 3.6 displays the solutions in greater detail associated with Figure 3.5b, where each solution is labeled by specific representation l and a line connects consecutive layers. By visual inspection, it is not difficult to observe convergence behavior in the optimal facility locations as demand density increases. Figure 3.7 summarizes the distance between optimal locations of successive representations for the “circular” region.

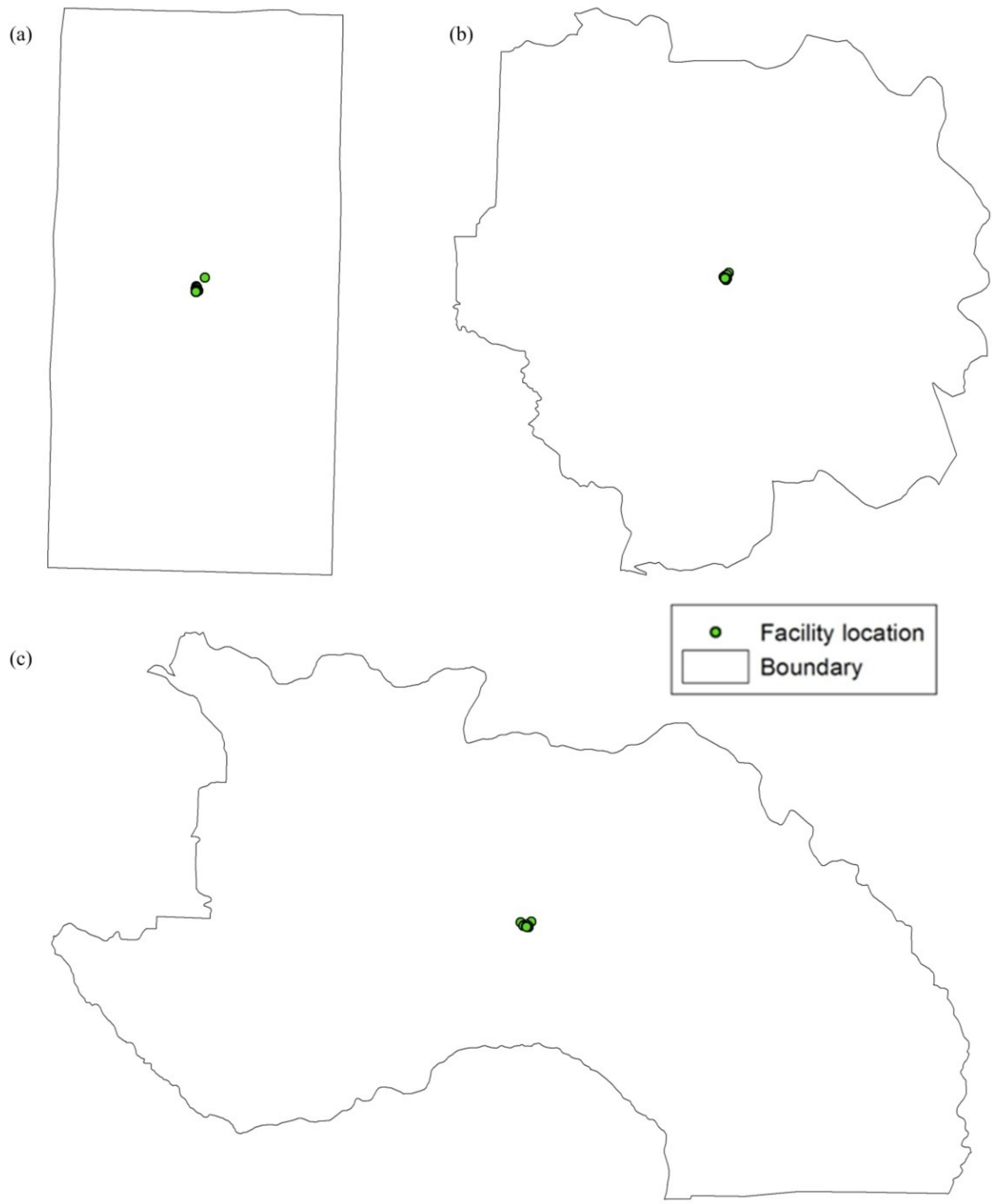


Figure 3.5: Solutions for the different shaped regions (homogeneous demand): (a) Rectangular, (b) Circular and (c) Concave

Though the curve in Figure 3.7 is not monotonically decreasing, in general, we have $\lim_{l\to\infty}\sqrt{(X_l^* - X_{l-1}^*)^2 + (Y_l^* - Y_{l-1}^*)^2} = 0$, supporting that (X_{l-1}^*, Y_{l-1}^*) and (X_l^*, Y_l^*) become closer and closer, approaching the theoretical optimal facility location

(X^*, Y^*) for the continuous Weber problem. In particular, most of the distance variation in Figure 3.7 is less than $10\,m$ after the first a few iterations. For most facilities, siting to within $\pm 10m$ would be considered extremely precise.

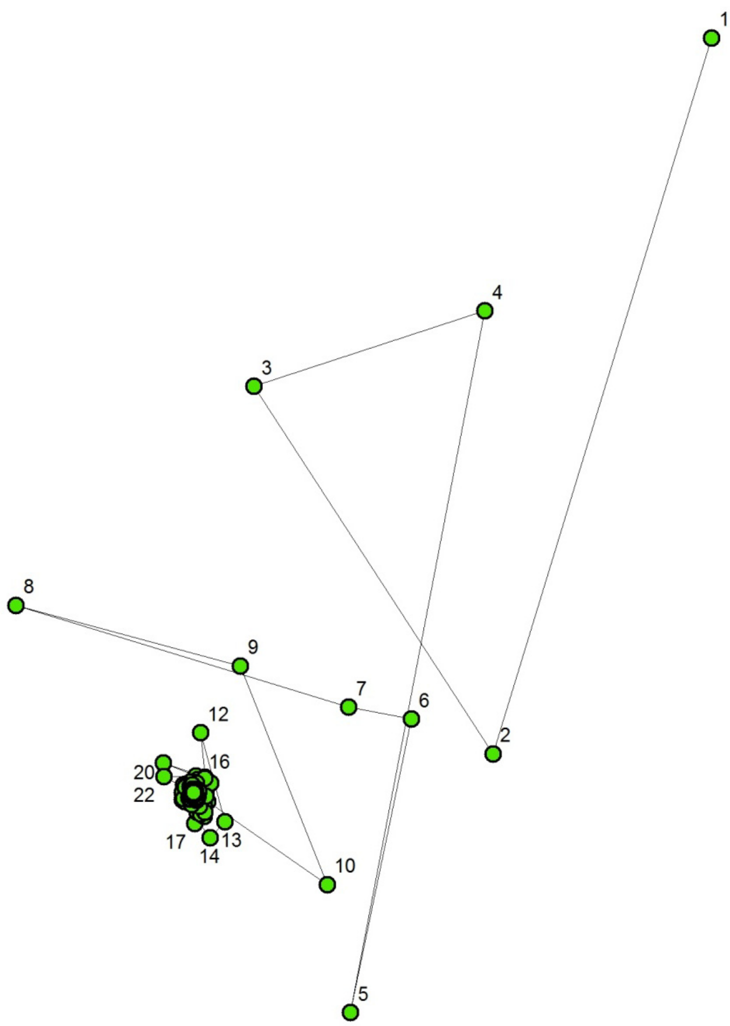


Figure 3.6: Solution location variability based on demand density change (for Figure 3.5b)

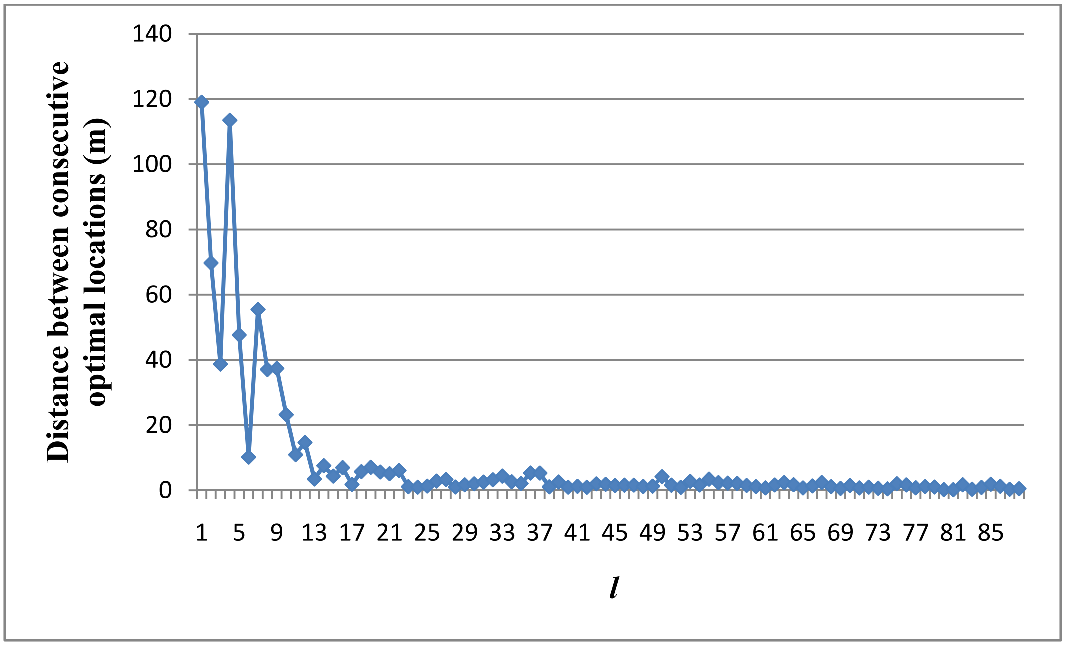


Figure 3.7: Distance between facility locations for successive representations (for Figure 3.5b)

3.6.2 Region with varying demand

Now our attention is turned to a region with varying demand. Specifically, the region contains 55 Census tracts. The area of each sub-region ranges from 0.325 km^2 to 23.495 km^2, and the total area is 209.968 km^2. The continuous demand density is estimated based on the sampling points, one for each sub-region, with density varying from 224unit/km^2 to 9493unit/km^2. A demand surface layer fitted from these points is shown in Figure 3.8a. The actual demand density in each sub-region is depicted by Figure 3.8b.

An alternative for representing the continuous demand in this region could be discrete points, as done previously. Though the demand changes across the entire area, it must be assumed uniformly distributed over each sub-region.

Specifically, the demand is represented using equally spaced points but with different spatial density in each sub-region. This is illustrated in Figure 3.8c. Thus, the proposed solution approach can be applied in this situation as well, with processing to ensure appropriate spatial density proportions. The continuous Weber problem is solved with following parameters: INIT_D = STEP_D = 2 points/km^2, NUM_R = 1, Γ = 5 and τ = 0.5.

When convergence is achieved, the total number of layers generated is 106. That is $|\Omega| = 106$ with processing time of 57.5 minutes. To obtain the best approximate optimal solution, the maximum number of points created for each sub-region varies from 70 to 5,028, and the total number of points for the entire region is 247,253. Shown in Figure 3.9 are the optimal facility locations for each representation. On the whole, the solution set has a very small geographic extent. The detailed map on the right side of Figure 3.9 depicts a similar convergence trend observed previously. The approach quickly converges to the optimal location.

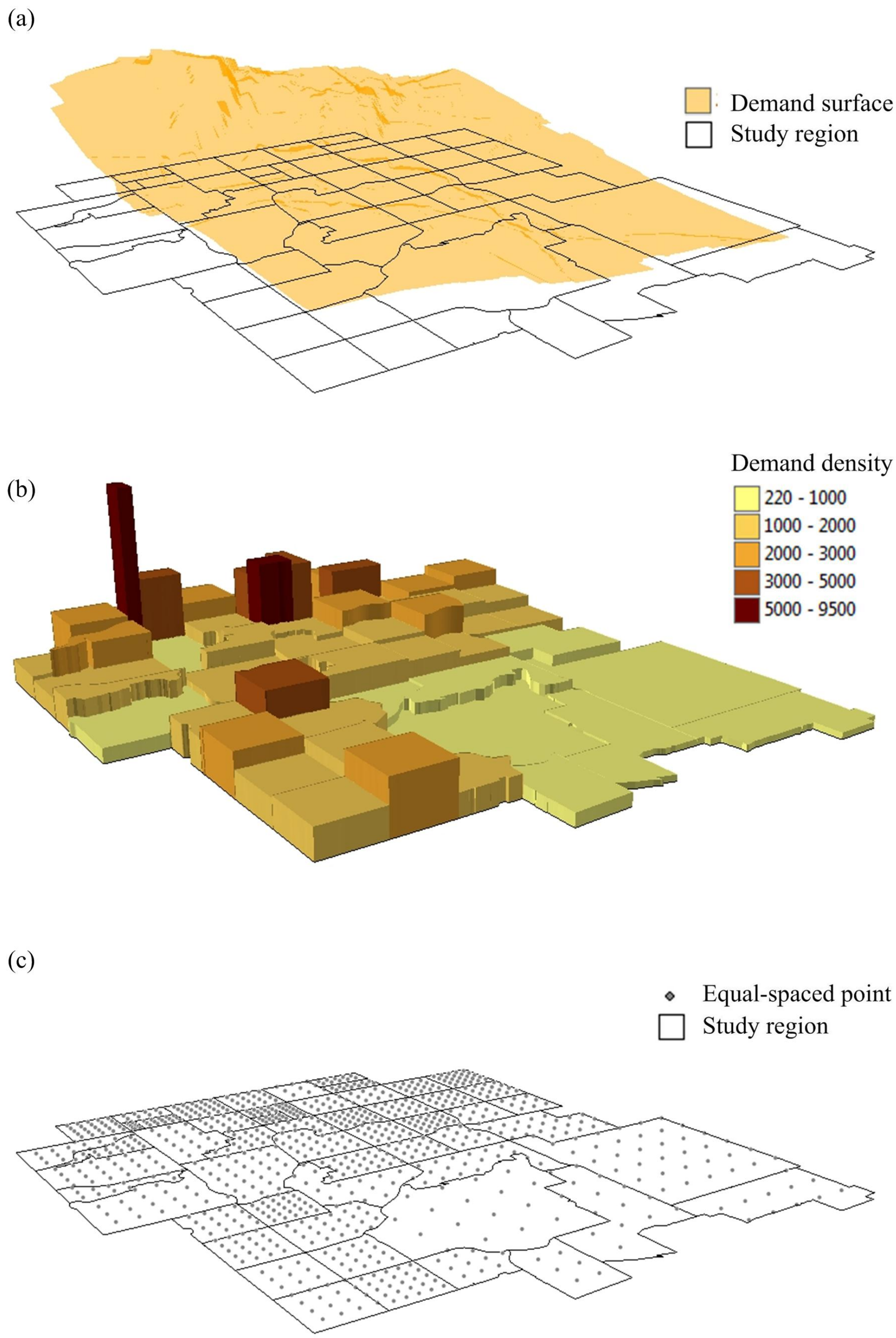


Figure 3.8: A region with varying demand: (a) demand surface, (b) actual demand and (c) discrete demand point

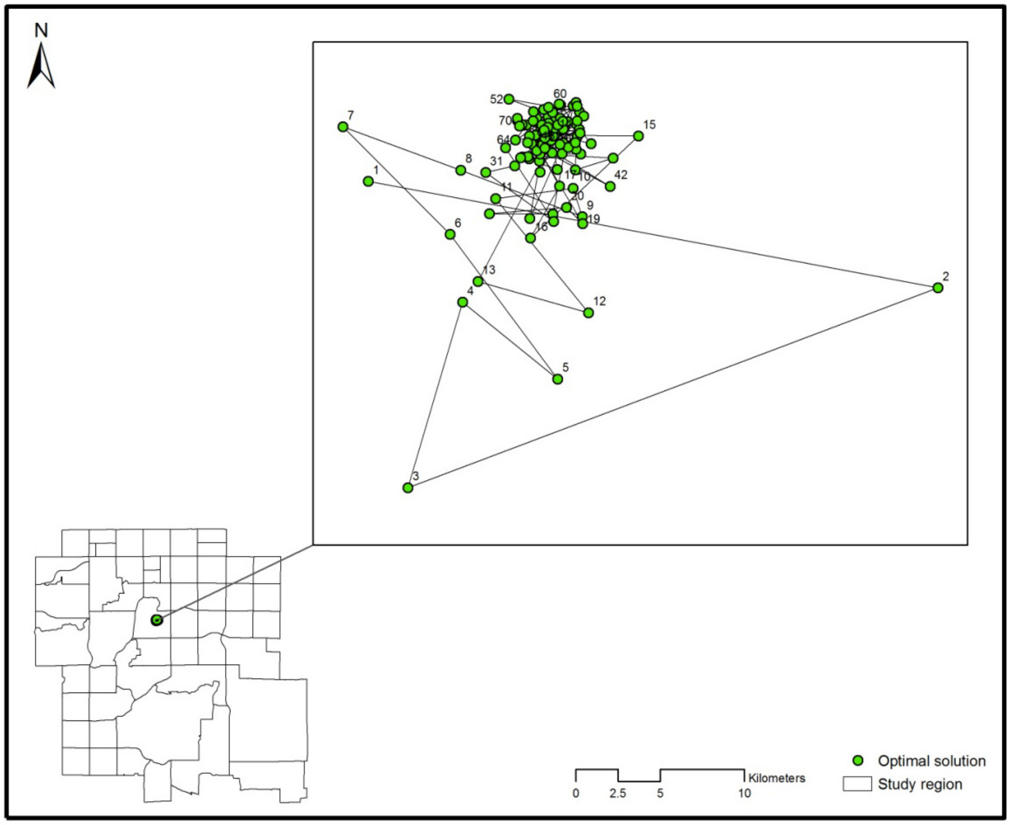


Figure 3.9: Optimal facility locations for varying regional demand

3.7 Discussion and Conclusions

In location science, the Weber problem has proven to be an important planning model with numerous extensions. Though continuously distributed demand reflects important realities of the real world, solving the Weber problem in this case presents challenges. One approach is to introduce a double integral in the objective function, complicating the problem as well as making solution extremely difficult. Further, assuming that a demand distribution $g(x, y)$ is known and accurate is highly problematic in any planning context. This chapter thus proposed an effective approach to solve the continuous Weber problem.

A key issue in the solution procedure is how to appropriately represent continuous demand. As noted previously, it is possible to describe demand in a GIS environment as an object or a field. On one hand, though field views can depict a continuous surface given by an exact demand density function, the problem is that such theoretical mathematical functions are not known in practice. Further, a fitted surface derived from spatial interpolation is subject to errors and uncertainty. On the other hand, according to infill asymptotics, the object view appears to be a better alternative as demand point density increases, which is therefore adopted in the proposed approach.

In other words, the developed solution method for the continuous Weber problem uses discrete points of different demand densities to approximate continuous demand. Empirical applications explored the relationship between solutions and point demand density. Optimal facility locations and objective values were found to converge quickly in all cases. However, regional shape appears to influence performance behavior. For example, the distance between (X_1^*, Y_1^*) and (X_2^*, Y_2^*) in Figure 3.7 is about $50\,m$, much larger than the values on the tail of the curve, most of which are less than $1\,m$. However, as point density increased, the differences became smaller and smaller until finally convergence is achieved.

Though the convergence criterion in the proposed approach relies on the objective value assessment, other methods of convergence could be adopted as well. For example, since convergence is also observed in the optimal facility locations from different representations with increasing demand density,

$\sum_{k=l-\Gamma}^{l-1} \left| Z_l^* - Z_k^* \right| < \tau$ in Figure 3.3 could be substituted by

$\sum_{k=l-\Gamma}^{l-1} \sqrt{(X_l^* - X_k^*)^2 + (Y_l^* - Y_k^*)^2} < \tau$. That is, the solution procedure terminates only when the optimal facility location changes little. Also, convergence could be examined by statistical methods such as root and ratio tests. For instance, objective functions values or distances between optimal facility locations from successive representations can be considered as a series of values, where a convergence test could be carried out to find the best representation and associated solution.

As discussed above, the proposed method recognizes important limitations in accurately representing continuous demand and employed discrete representations as a good approximation for continuous demand in facility location models provided that demand point density was sufficient. When convergence is achieved, an approximate optimal solution based on infill asymptotics is possible. GIS facilitates this process, making it straightforward to vary density in an informed and justified manner. Results from the empirical studies demonstrated that the proposed approach is effective in solving the continuous Weber problem.

Chapter 4

THE CONTINUOUS MULTI-WEBER PROBLEM††

4.1 Introduction

Location theory is concerned with analysis and placement of socio-economic activities, including land use, industrial production, central places and spatial competition (Murray 2008). Founded on such theory, location modeling has long been recognized as playing an important role in regional and urban planning, as well as other contexts, as “all human activities involve the choice, either explicit or implicit, of location” (Church and Sorensen 1996). Location related decisions are fundamental in many aspects of human activities. Examples include, but are not limited to, locating retail stores, deploying switching centers in communication networks, selecting nature reserves to preserve threatened species, etc. In fact, whenever a question about where to place goods and services is posed, a location problem arises.

Facility location problems usually concern determining where to site one or more facilities subject to certain constraints in order to optimize objectives (Brandeau and Chiu 1989). Of particular interest here is a minimization problem – the multi-facility Weber problem with continuously distributed demand, or simply the continuous multi-Weber problem. The Weber problem is a classic facility location model, with the task of identifying a site for a single facility in continuous space in order to minimize the total transportation costs from the

†† A modified version of this chapter has been submitted for publication, co-authored with Alan Murray.

facility to a set of fixed demand points (Weber 1909). The multi-facility extension of the Weber problem requires multiple facilities to be sited (Cooper 1963), and remains a challenge to solve (see Rosing 1992, Righini and Zaniboni 2007). Add to this the need to deal with continuously distributed demand, the continuous multi-Weber problem is arguably one of the most difficult facility location problems to solve optimally or heuristically.

Continuous demand is often encountered in planning and analysis, reflecting the distribution of people, risk, danger, vegetation, businesses, retail markets, and customers more generally. Typically, continuous demand is abstracted as a finite set of discrete points in order to facilitate model formulation as well as reduce computational expense (Miller 1996, Ouyang and Daganzo 2006, Francis *et al*. 2009). Though attractive, point-based simplifications can lead to significant errors and uncertainty in analyses due to the loss of spatial detail (Murray and O'Kelly 2002, Murray *et al*. 2008, Francis *et al*. 2009, Alexandris and Giannikos 2010, Cromley *et al*. 2012). An alternative is to describe continuous demand as a surface defined by a mathematical function (Drezner 1995, Gastner and Newman 2006, Brimberg *et al*. 2008, Murat *et al*. 2010). The issue, however, is that the exact demand distribution/function is never known or given with certainty. Therefore, how to effectively deal with continuously distributed demand in facility location remains a great challenge. To this end, the intent of this chapter is to address continuous demand representation in the continuous multi-Weber problem.

When facilities can be located anywhere and demand is continuously distributed, model formulation as well as solution becomes mathematically complicated. Traditionally, facility location problems have been solved by operations research techniques like linear and integer programming. In recent years, solution capabilities have been significantly enhanced by spatial optimization approaches incorporating geographic information systems (GIS) functions and methodologies (Church 2002, Wei *et al*. 2006, Church and Murray 2009, Matisziw and Murray 2009, Murray 2010, Cromley *et al*. 2012). GIS is used to facilitate data input and visualization, but arguably is more valuable and meaningful as part of modeling and solution processes (Murray 2010).

The aim of this chapter is to develop an approach that combines GIS and optimization methods for solving the continuous multi-Weber problem. The next section reviews relevant research in this area. The problem specification is given in section 4.3. Spatial representation in GIS is covered in section 4.4. This is followed by a detailed description of the proposed method in section 4.5 and empirical results in section 4.6. Finally, the chapter ends with discussion and conclusions.

4.2 Background

As one of the first location problems formally posed, the Weber problem involves placing a single facility anywhere in space (continuous space) to serve a finite set of demand points so that total transportation costs are minimized. It has been extensively investigated and continues to be of interest since first proposed in the

17th century (Wesolowsky 1993, Drezner *et al*. 2002). The classic context is siting a factory in order to minimize the transportation costs to acquire raw materials and distribute products, but could as well involve locating a fire station or hospital to minimize average response time. This seemingly simple problem has attracted so much interest because of its overall applicability as well as potential for extension. Many location models can, in fact, be formally connected to the Weber problem (Drezner *et al*. 2002).

One extension of interest here is replacing discrete demand by continuous demand, usually referred to as the continuous Weber problem (Drezner 1995, Fekete *et al*. 2005). In this context, the Weber problem becomes a stochastic optimization problem with the goal of minimizing average distance to the demand area from the facility. The most straightforward method for dealing with a demand area is to represent it by a single point, such as a centroid (Bennett and Mirakhor 1974). Given the well-known concerns and limitations of data aggregation, research has employed knowledge from computational geometry to evaluate average travel distance (Drezner 1995, Carrizosa *et al*. 1998, Fekete *et al*. 2005) or its bounds (Carmi *et al*. 2005, Abu-Affash and Katz 2008, Puerto and Rodríguez-Chía 2011). Such approximations of average distance are relatively straightforward for regular shapes but can be quite complex in other cases. Finally, some efforts have focused on numerical methods for problem solution. For example, the Weiszfeld algorithm, popular for solving the Weber problem, has been adapted for dealing with continuous extensions (Drezner and Wesolowsky 1980, Franco *et al*. 2008).

Another relevant extension considers multiple facilities instead of a single facility, known as the multi-Weber problem (Cooper 1963). The distinct feature of this extension is that it concerns siting multiple facilities simultaneously in continuous space with discrete demand assigning to its closest facility. The multi-Weber problem has challenged generations of researchers because the objective function is neither concave nor convex, making it difficult to find the global minima. Since the early ALTERNATE heuristic developed by Cooper (1964), many heuristics have been proposed, such as Tabu search (Brimberg and Mladenović 1996), genetic algorithms (Houck *et al*. 1996, Salhi and Gamal 2003) and variable neighborhood decomposition (Brimberg et al. 2006). The benefit of such approaches is that they can solve large problems quickly as well as provide a good initial solution for exact approaches. In contrast, exact or optimal methods are constrained by application problem size (Rosing 1992, Righini and Zaniboni 2007). A detailed survey of exact methods and heuristic approaches can be found in Brimberg et al. (2008).

When extension involves both continuous demand and multiple facilities, the problem becomes extremely complicated as facilities may be sited anywhere and demand varies across space. Algorithms have been proposed for the simplest assumption of a uniform demand distribution (Marucheck and Aly 1981, Drezner 1986). Work dealing with other probabilistic distributions of demand can be found in Rao and Varma (1985) and Altınel et al. (2009). To facilitate the allocation process, some researchers have utilized Voronoi partitions of the demand area (see Suzuki and Okabe 1995, Gastner and Newman 2006). More recently, Murat *et al*. (2010) applied an approximate line search method to find a single facility location for each Voronoi polygon, but it can be computationally expensive due to a focus on the allocation process.

Compared to the continuous Weber problem and the multi-Weber problem, the extension of both multiple facilities and continuous demand has drawn less attention. It is worth noting that current solution approaches for the continuous multi-Weber problem are problem specific, inexact or not computationally efficient. Therefore, addressing multiple facilities and continuous demand represents an unresolved issue, with much potential for use and application. Given the inherent spatial nature of the problem and advanced spatial techniques linked to GIS, spatial optimization represents potential for improved problem solution that exploits knowledge of distributed demand. This avoids problematic assumptions characteristic of existing solution approaches for this problem.

4.3 Problem Specification

The continuous multi-Weber problem involves continuously distributed demand served by several facilities that may be located anywhere in continuous space. To begin, the mathematical formulation of the classic Weber problem is detailed. Then, two related extensions, the continuous Weber problem and the multi-Weber problem, are given. The continuous multi-Weber problem is then presented.

4.3.1 Weber Problem

Given demand represented by discrete, aggregate points, the objective function of the Weber problem is to minimize the sum of weighted distance to the sited facility over the finite set of demand points. This is equivalent to minimizing average distance to demand. The decision variables are the location of the facility. Consider the following notation:

i = index of demand points

(x_i, y_i) = location of demand point *i*

n = number of demand points

w_i = weight associated with demand point *i*

(X, Y) = facility location decision

The Weber problem is as follows (Wesolowsky 1993):

$$\textit{Minimize} \qquad \sum_{i=1}^{n} w_i \sqrt{(X - x_i)^2 + (Y - y_i)^2} \tag{4-1}$$

The goal is to find the best location in continuous space defined by (X, Y) so that the total weighted distances from the demand points to the sited facility is minimized. Notice that the distance function in (4-1) is defined by the Euclidean metric, and involves (X, Y) as the facility location site to be determined.

4.3.2 Continuous Weber Problem

As mentioned above, the Weber problem, (4-1), can be extended in many ways. Previous extensions include models considering other distance measures, negative weights associated with demand, or multiple objective functions, among others (Drezner *et al*. 2002).

An important extension is accounting explicitly for demand varying continuously over space, not assuming aggregate demand points. This extension obviously complicates things as representation of a continuous surface is not trivial. Consider the following additional notation:

$g(x,y)$ = function of demand at point (x,y)

$d(x,y,X,Y) = \sqrt{(X-x)^2+(Y-y)^2}$

R = region of demand

Discrete demand at a priori defined locations, w_{ij}, are replaced by a function corresponding to demand at any location (x,y). Further, the distance function reflects this change as well. Formal specification of the continuous Weber problem follows (Church and Murray 2009):

$$Minimize \quad \iint_{(x,y)\in R} g(x,y)d(x,y,X,Y)dxdy \qquad (4\text{-}2)$$

The objective, (4-2), involves a double integral over the demand area instead of the sum in the Weber problem, (4-1). The distance, $d(x,y,X,Y)$, remains the Euclidean metric, but the objective function now accounts for continuously distributed demand using $g(x,y)$, making this far more complex. The reason it is more complex is that in order to apply this model, $g(x,y)$ must first be mathematically defined.

Though more complicated in formulation, the continuous Weber problem, (4-2), remains equivalent to (4-1), minimizing the average travel distance to the sited facility (Carrizosa *et al*. 1998).

4.3.3 Multi-Weber Problem

Another extension to the Weber problem considers multiple facilities. Thus, decisions regarding where each facility should be located must also consider what demand is assigned to which facility. This makes the problem a location and allocation problem, or simply a location-allocation problem. Additional notation:

j = index of facilities

p = number of facilities to be located

(X_j, Y_j) = location of facility j

$$\Psi_{ij} = \begin{cases} 1 & \text{demand point } i \text{ is served by facility } j \\ 0 & \text{otherwise} \end{cases}$$

Now there are decision variables for each of the p facilities to be sited. Further, allocation variables to assign demand to each facility are introduced. The formulation of the multi-Weber problem can be expressed as follows (Cooper 1963, 1964):

$$\textit{Minimize} \quad \sum_{i=1}^{n}\sum_{j=1}^{p} w_i \Psi_{ij} \sqrt{(X_j - x_i)^2 + (Y_j - y_i)^2} \qquad (4\text{-}3)$$

$$\textit{Subject to}: \quad \sum_{j=1}^{p} \Psi_{ij} = 1 \quad \forall i \qquad (4\text{-}4)$$

The objective, (4-3), remains to minimize total weighted distances. Constraints (4-4) specify that each demand point is served exactly by one facility.

This model is more complicated with multiple facilities than the Weber problem, (4-1), since it not only considers facility location, $\left(X_j, Y_j\right)$, but now also accounts for allocation of demand points among several facilities, indicated by Ψ_{ij}.

4.3.4 Continuous Multi-Weber Problem

Combining the above extensions, the continuous Weber problem and the multi-Weber problem, we get a very challenging facility location problem. The variation of the Weber problem in this case concerns both continuous demand and multiple facilities, referred to as the continuous multi-Weber problem. Consider the following additional notation:

$d(x, y, X_j, Y_j) = \sqrt{(X_j - x)^2 + (Y_j - y)^2}$

R_j = sub-region of demand served by facility *j*

The discrete allocation variables, Ψ_{ij}, are now replaced by the variables R_j that define a portion of the demand region to be served. More specifically, this is the area served by facility *j*, $\left(X_j, Y_j\right)$, and includes $\left(x, y\right) \in R_j$ reflecting demand assignment to the facility. The continuous Multi-Weber problem follows:

$$\text{Minimize} \quad \sum_{j=1}^{p} \iint_{(x,y)\in R_j} g(x,y) d(x,y,X_j,Y_j)\, dx\, dy \qquad (4\text{-}5)$$

$$\textit{Subject to:} \qquad \bigcup_j R_j = R \tag{4-6}$$

$$R_j \cap R_{j'} = \varnothing \quad \forall j, j' \text{ where } j' \neq j \tag{4-7}$$

The objective, (4-5), includes both a double integral over each sub-region R_j served by facility j and a sum over the entire demand area characterized by a demand distribution $g(x, y)$. Similar to the previous models, this is equivalent to minimizing average weighted distance. In addition to seeking the best facility locations, the model must simultaneously determine the best allocation as well, reflected in constraints (4-6) and (4-7).

Essential here is making the distinction between a discrete and continuous representation of demand across space. To this end, this chapter will investigate how GIS can accommodate the continuous distribution of demand, as well as how this might facilitate structuring and solving the multiple facility problem associated with such a demand representation.

4.4 Spatial Representation

The appropriate representation of geographic space has long been a concern in the field of GIScience and spatial analysis (Miller and Wentz 2003, Goodchild and Haining 2004). In GIS, discrete-object and continuous-field are two common ways to conceptualize geographic space (Worboys and Duckham 2004, Longley *et al*. 2011). Discrete objects usually refer to features with distinct boundaries, such as houses, trees and roads, while continuous fields are often utilized to describe phenomena dispersed over space, such as pollution, elevation and precipitation. Though many more advanced concepts have been proposed in recent years to describe complex spatial phenomena (Cova and Goodchild 2002, Yuan 2001), the object/field views remain the basis for geographic representation (Goodchild *et al*. 2007).

Given the geographic nature of facility location problems, spatial representation is no doubt crucial to consider. A necessity in any facility location model is to appropriately represent facilities and demand in geographic space. Of particular interest here is continuous demand, which can be conceived of as an object or a field view. Discrete points based on the object perspective have been popular in location modeling for the last several decades, largely attributed to model and computation simplification (Miller 1996, Francis *et al*. 2009). However, significant errors can result from data aggregation if the underlying demand is continuous in nature (Murray and O'Kelly 2002, Murray *et al*. 2008, Francis *et al*. 2009, Alexandris and Giannikos 2010, Cromley *et al*. 2012).

Because of theoretical as well as practical limitations associated with discrete demand, recent interest in location modeling has shifted to addressing continuous representations of demand. There are primarily two approaches that have been considered in location models. One option is to address continuous demand using an exact surface defined by a mathematical function (Carrizosa *et al.* 1998, Murat *et al*. 2010). One advantage is that demand is readily known for every location (x, y) in continuous space. In addition, solution approaches may be possible that exploit and benefit from the specific mathematical properties of these theoretical functions. The issue, however, is that such functions are never known with certainty due to the lack of complete information about the underlying demand distribution. This is clearly a problem.

As an alternative, the continuous surface may be approximated then used in a model. This is typically done in practice based on sample data. There are a number of techniques in GIS that can be used to fit a continuous surface, including inverse distance weighting (IDW), natural-neighbor, trend surface, Kriging and so on. They are generally known as spatial interpolation methods (Longley *et al*. 2011). The goal of spatial interpolation is to estimate attribute values at unobserved locations using the known values collected/measured at a finite set of locations (Cressie 1993). Consider the following notation:

$g(x, y)$ = true attribute value at location (x, y)

$\hat{g}(x, y)$ = estimated attribute value at location (x, y)

$\varepsilon(x, y)$ = $\left|g(x, y) - \hat{g}(x, y)\right|$

The intent of spatial interpolation is to estimate $\hat{g}(x, y)$, necessarily making it an approximation of $g(x, y)$. Clearly it is desirable to minimize total estimation errors associated with an approximated function. This may be stated formally as follows:

$$\textit{Minimize} \qquad E = \iint_{(x,y)\in R} \varepsilon(x, y)dxdy \tag{4-8}$$

It is well recognized that estimation error $\varepsilon(x, y)$ is unavoidable, and results from inaccurate sample data, assumptions inherent in interpolation approaches, as well as others (Lam 1983, Cressie 1993). Thus, errors and uncertainty exist in any derived approximation surfaces. That is, $\varepsilon(x, y) > 0$ regardless of the interpolation techniques employed.

If such fitted surfaces are further used in spatial analysis, cumulative errors can be introduced in obtained results and findings (Wood and Fisher 1993, Miller and Wentz 2003, Oksanen and Sarjakoski 2005). In the case of continuous demand in facility location problems, though errors caused by point-based abstraction have been well recognized, issues associated with the use of continuous surfaces are not well understood. Chapter 2 explicitly demonstrated error in the obtained facility location when the continuous Weber problem is considered. Addressing continuous representation in the multi-facility case remains a major challenge.

4.5 Solution Approach

As discussed above, it is impossible to represent continuous demand without any associated error using a mathematical function or spatial interpolation. Rather than pursuing exact representation, asymptotic theories are often employed in practice to obtain results. Infill asymptotics (Cressie 1993), or fixed-domain asymptotics (Stein 1999) is such a theory widely applied in spatial analysis related fields. The underlying principle is that the estimation error in spatial interpolation tends to zero when the sample size increases to infinity. In other word, when the number of sample points, m , goes to infinity, the difference between the true attribute value and the estimated attribute value can be negligible, and can be formally expressed as follows:

$$\lim_{m\to\infty} \hat{g}(x,y) = g(x,y) \tag{4-9}$$

From equation (4-8) this implies:

$$\lim_{m\to\infty} E = 0 \tag{4-10}$$

The infill asymptotic theory implications for facility location problems reliant on continuous demand are, though the point-based simplification can lead to errors in the solution of optimal facility locations, that it is possible to improve continuous demand approximation by increasing the number or density of sample points for a bounded region. Similar conclusions are also proven by Francis and Lowe (2011).

The proposed heuristic approach for solving the continuous multi-Weber problem is therefore based on the discrete approximation for continuous demand, exploiting the above asymptotic theory in an intelligent way. The assumption is that better approximation of continuous demand can be obtained if $m \to \infty$, which indicates that a solution to the multi-Weber problem with higher demand point densities would approach the optimum of the continuous multi-Weber problem. In practice, however, a finite set of demand points is hopefully sufficient to achieve a good approximation of the actual continuous distribution. In one respect, the discrete point approximation is less restrictive compared to other alternatives as the demand points can be identified/collected based on known characteristics of the underlying continuous distribution. In another respect, demand point density can be systematically increased to improve accuracy as a discrete representation.

The proposed heuristic solution procedure for the continuous multi-Weber problem is given in Figure 4.1. Consider the following parameters:

INIT_D = initial demand point density

l = demand point layer

p = number of facilities

t = index of sub-areas in the demand region

s = number of sub-areas in the demand region

δ_t^l = demand point density in sub-area t

Γ = number of layers to consider in examining convergence of objective

τ = tolerance used to determine convergence of the objective value

Δ_j = change in demand point density for facility j service area

The process starts in Figure 4.1 with INPUT, where initialization occurs. With the data and knowledge of the underlying continuous distribution, an approximation layer l is generated. The layer approximation is systematically enhanced as l increases in Figure 4.1. Considering the spatial variation of demand over space, the study region is divided into a set of sub-areas with demand density δ_t^l. Thus, each sub-area has a unique density. Conversion of the region into discrete points can be carried out with GIS software like ArcGIS, a very popular commercial GIS package.

Once a demand layer l is approximated, the multi-Weber problem is solved in Figure 4.1. The algorithm applied to solve the multi-Weber problem is ALTERNATE (Cooper 1964) because of its wide application and ease of

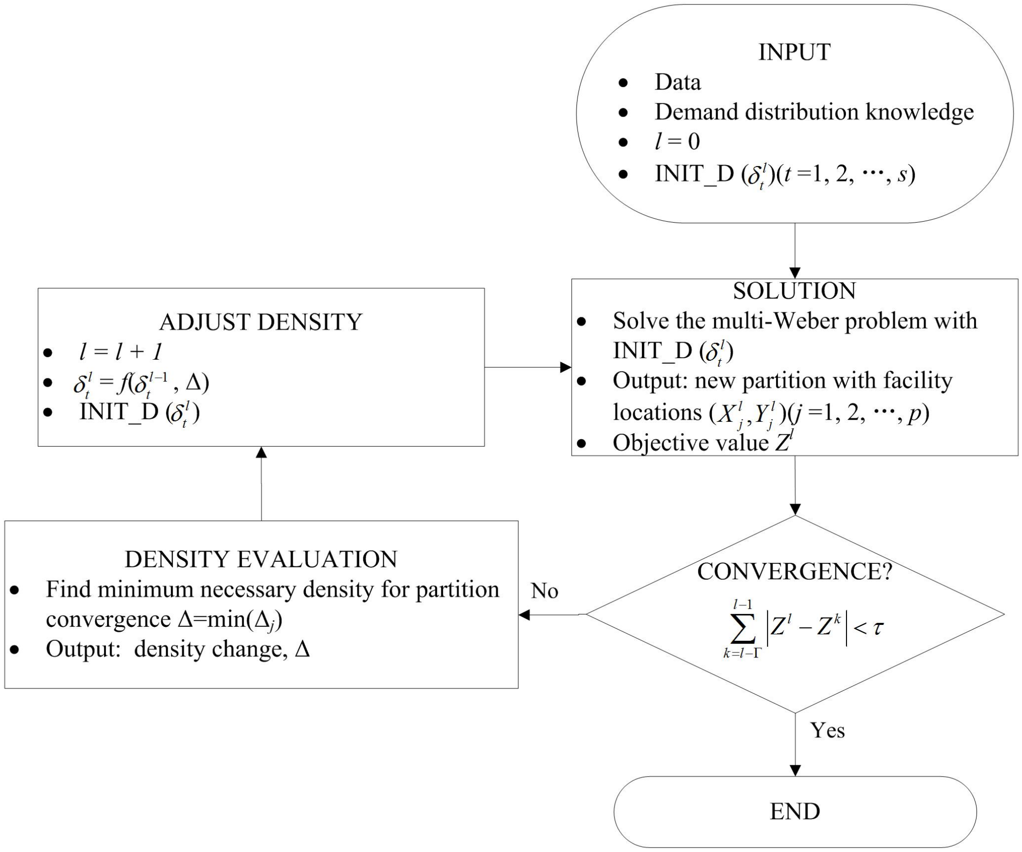


Figure 4.1: Solution procedure for the continuous multi-Weber problem

implementation. The general description of ALTERNATE is given in Figure 4.2. First, the study area is divided into several adjacent partitions, one for each facility, so that all demand points are served. Next, the Weber problem is solved to find the optimal facility location for each partition. The two steps are repeated until no further improvement in the objective can be achieved. This process usually repeats several times and the best result is selected. The final solution is the facility locations (X_j^l, Y_j^l) for layer l as well as the allocation of each demand point to its nearest facility. Also, an objective value representing the minimized average distance to the closest facility, Z^l, is obtained, calculated as

$Z^l = \sum_{i=1}^{n}\sum_{j=1}^{p} w_i \Psi_{ij} d_{ij} / \sum_{i=1}^{n} w_i$, where $d_{ij} = \sqrt{(X_j - x_i)^2 + (Y_j - y_i)^2}$.

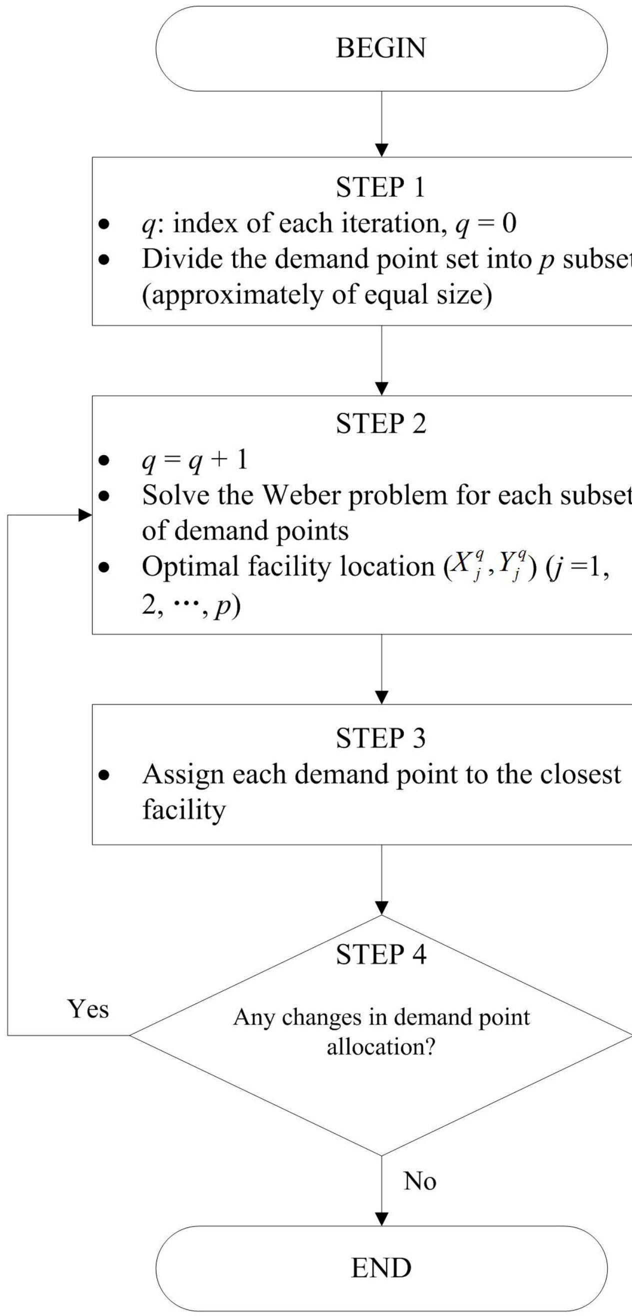


Figure 4.2: ALTERNATE heuristic

The next step in Figure 4.1 is assessment of convergence (CONVERGENCE?). The evaluation is based on the last Γ layers, calculated as the sum of the absolute difference between the objective value of current layer Z^l and previous layers, Z^k where 0<*k*<*l*. Specifically, the convergence measure is $\sum_{k=l-\Gamma}^{l-1} \left| Z^l - Z^k \right|$, and is then compared to a given tolerance threshold, τ. If τ is larger, convergence is assumed and the solution procedure terminates (END). As a result, the solution derived from layer l is the approximate solution for the continuous multi-Weber problem with average distance to the nearest facility Z^l.

Otherwise, the process in Figure 4.1 proceeds to DENSITY EVALUATION. In this stage, the Weber problem is solved for each partition to seek a good point density for the discrete representation. The intent of this step is to better approximate the continuous surface. Suppose ($\Delta_1,\ \Delta_2,\ \cdots,\ \Delta_p$) are the changes in the point densities that are necessary to obtain solution convergence for different partitions, the minimum value $\Delta = \min(\Delta_1,\ \Delta_2,\ \cdots,\ \Delta_p)$ is used as the density change for the entire layer in ADJUST DENSITY.

Based on change density Δ and the point density of the current layer, Figure 4.1 indicates that a new demand layer is generated by adjusting the demand density for each sub-area, keeping the proportion of point densities among all sub-areas fixed. The procedure of demand density adjustment is similar to that given in Figure 3.4. Demand point layer is then used as input and re-solved. This process is repeated until convergence is achieved.

In summary, the proposed approach uses a systematic discrete point-based approximation of demand, each time moving closer toward asymptotic convergence. At each iteration, the best solution is found and comparison is made to previous representations. This continues until the no improvement is possible. Upon convergence, the best facility locations (X_j, Y_j) are found that minimize average travel distance *Z* for the continuous multi-Weber problem. The approximation of $g(x, y)$ using a series of demand point layers results in a series of solutions, where $Z^l \rightarrow Z$ as $n \rightarrow \infty$ (recall *n* is the number of demand points). Again, this is based on asymptotic theory, but in practice convergence is achieved for a reasonable *n*. Finding this reasonable *n* is accomplished through the proposed heuristic process given in Figure 4.1.

4.6 Empirical Results

The approach outlined in Figure 4.1 to solve the continuous multi-Weber problem was implemented using the Python programming language integrating an Open Source GIS library, Shapelib, to process geographic data. Analysis was carried out on a personal computer (Mac OS X system) running a 3.06 GHz, intel Core 2 Duo processor and 4GB, 800 MHz memory.

To assess the performance of the proposed heuristic, two types of demand distributions are considered: uniform and non-uniform. The former represents the situation where continuous demand is uniformly distributed over a region, and the latter reflects the situation where demand varies across space.

4.6.1 Uniform Demand

The data used here are the Census tracts shown in Figure 4.3, including two demand regions. The areas covered are 19.373 km^2 and 2.574 km^2 for Region A and Region B, respectively. It is assumed that the continuous demand is uniformly distributed over both regions, with the same initial point density, INI_D = 2*point*/ km^2. Specifically, equal spaced points are used in this case and can be obtained using available functions in GIS. Four facilities (p =4) are sited. Termination criteria are based on the last five consecutive data layers. Specifically, the parameters are $\Gamma = 5$ and $\tau = 2$.

Using the proposed solution procedure (Figure 4.1) with the above parameters, the best facility locations are illustrated in Figure 4.3 as green dots. Suppose Ω is the set of total data layers generated in reaching convergence, then for solution here we have $|\Omega| = 94$. Thus, every four points in different groups in Figure 4.3 correspond to the facility locations derived from a data layer l, where $l \in \Omega$. It can be seen that the four groups of facility locations are naturally clustered, with a few points littered around the group centers. A more detailed description of spatial distribution, as well as the distance between consecutive approximate locations of one facility group in Figure 4.3 is provided in Figure 4.4. The numbers in the top portion of Figure 4.4 correspond to the indices of data layers, l. It can be observed that the approximated facility locations for the first several demand layer approximations are far away from the others. The facilities converge as the point densities increase.

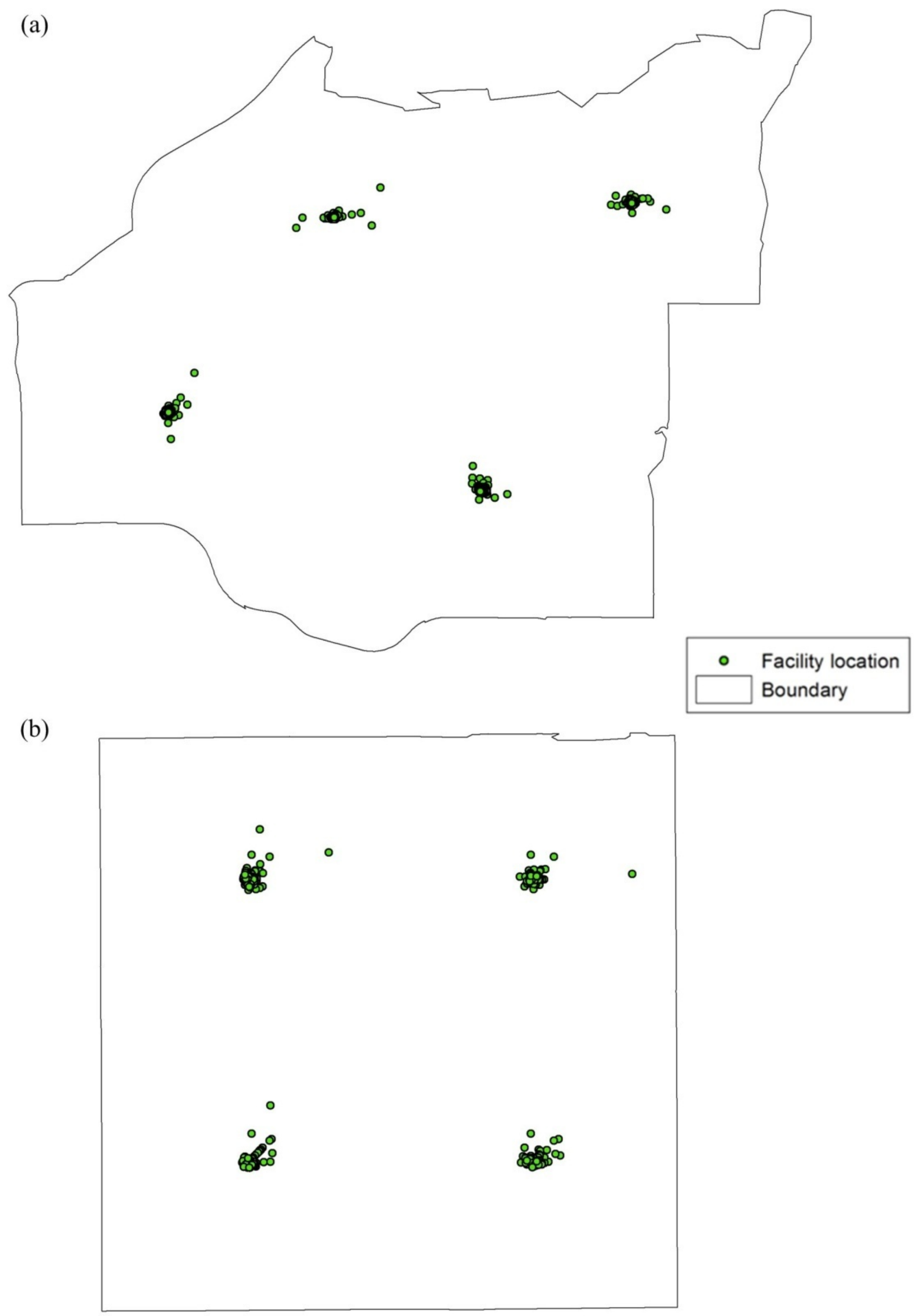


Figure 4.3: Solution for the study region with uniform continuous demand:
(a) Region A and (b) Region B

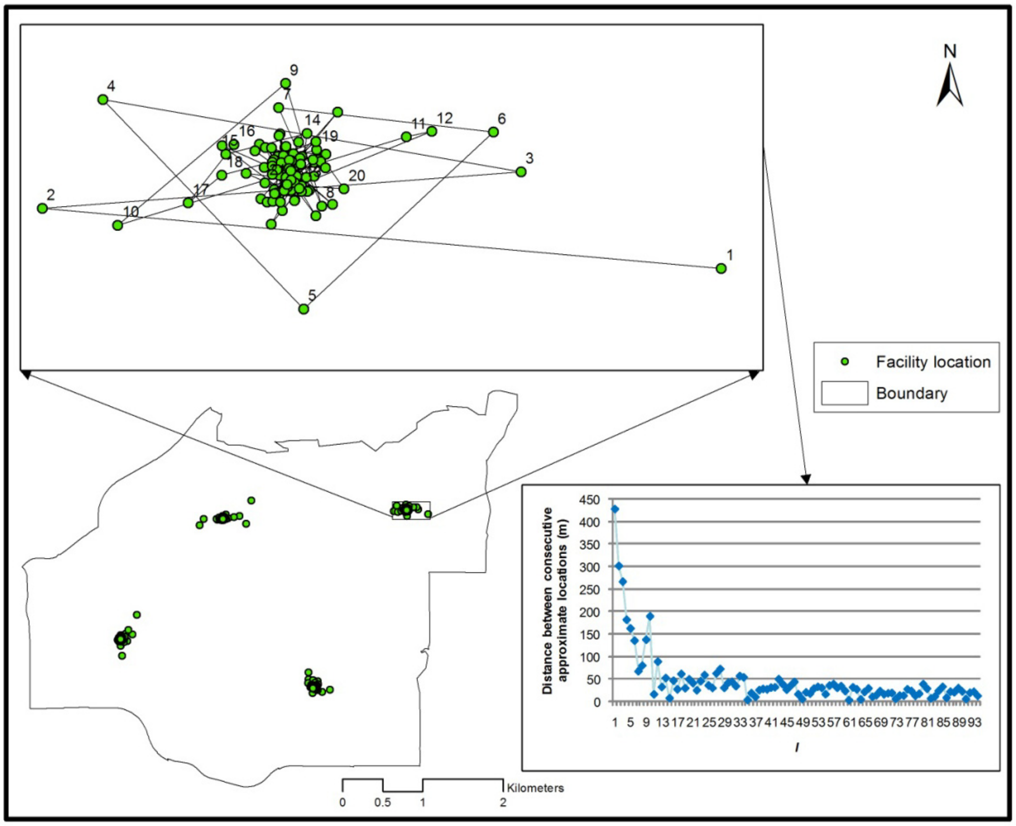


Figure 4.4: Detailed solution information for Figure 4.3(a)

This pattern is more apparent in the lower right portion of Figure 4.4 showing the distance between the two facility locations obtained from layer l and $l-1$, given $l \in \Omega$ and $l > 1$. The longest distance between the facilities is as high as 430*km*. This value quickly declines with the increase of point densities, and there is much less variation among the distances when $l > 12$, most of which are within 50*m*. The implication is that $\lim_{l \to \infty} \sqrt{(X_l - X_{l-1})^2 + (Y_l - Y_{l-1})^2} = 0$. The solution from the multi-Weber problem tends to be changing little and appears to approach the theoretical optimal solution of the continuous multi-Weber problem.

The distance variations for the last several layers are less than 30*m*, no doubt reasonable accuracy in practice.

4.6.2 Varying Demand Across Space

The second application concerns non-uniformly distributed continuous demand over space. This involves 55 Census tracts with varying demand. The smallest tract has an area 0.325 km^2 and the largest area is 23.495 km^2, with the study region covering a total area of 209.968 km^2. The demand density among the tracts varies 224unit/ km^2 to 9493unit/ km^2 as shown in Figure 4.5(b). Figure 4.5(a) describes the demand using a continuous surface fitted from spatial interpolation. An alternative representation is using discrete points. The assumption is that though the demand density varies across the study region, it can be considered uniform within each tract as depicted in Figure 4.5(c). The reason for this is that no further information exists to define the demand distribution at a finer level. Again, the developed method in Figure 4.1 is applied after the discretization of continuous demand in this context. The specific parameters used are: p = 4, INI_D = 1 points/ km^2, Γ = 5 and τ = 8.

The solution is shown in Figure 4.6. Similar to the results in Figure 4.4, the best facility locations are clustered in four groups, representing the sites derived for each facility from all the data layers, Ω. In total 20 layers are generated ($|\Omega|$ = 20) before the convergence criterion is satisfied. The spatial distribution of the facility locations in each group is also investigated, with distance between consecutive locations shown in Figure 4.6. A clear convergence

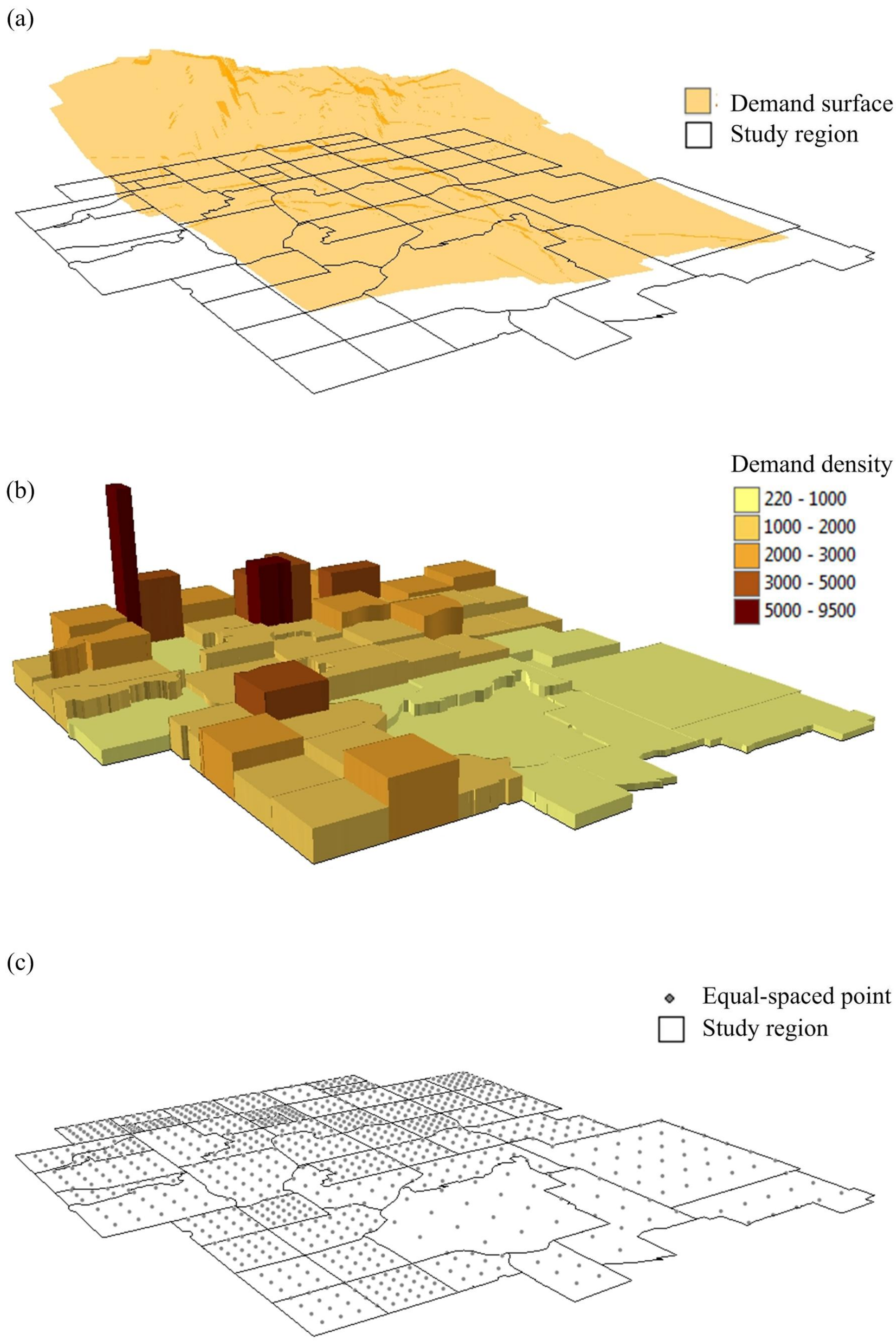


Figure 4.5: A region with varying demand: (a) demand surface,
(b) actual demand and (c) discrete demand point

pattern can be observed. The location differences for all the four location groups quickly decreases to within 50*m* after only a few iterations, then get much smaller when convergence is achieved. For example, in the lower left portion of Figure 4.6, the final distance is less than 10*m*, which is very accurate for actual facility planning.

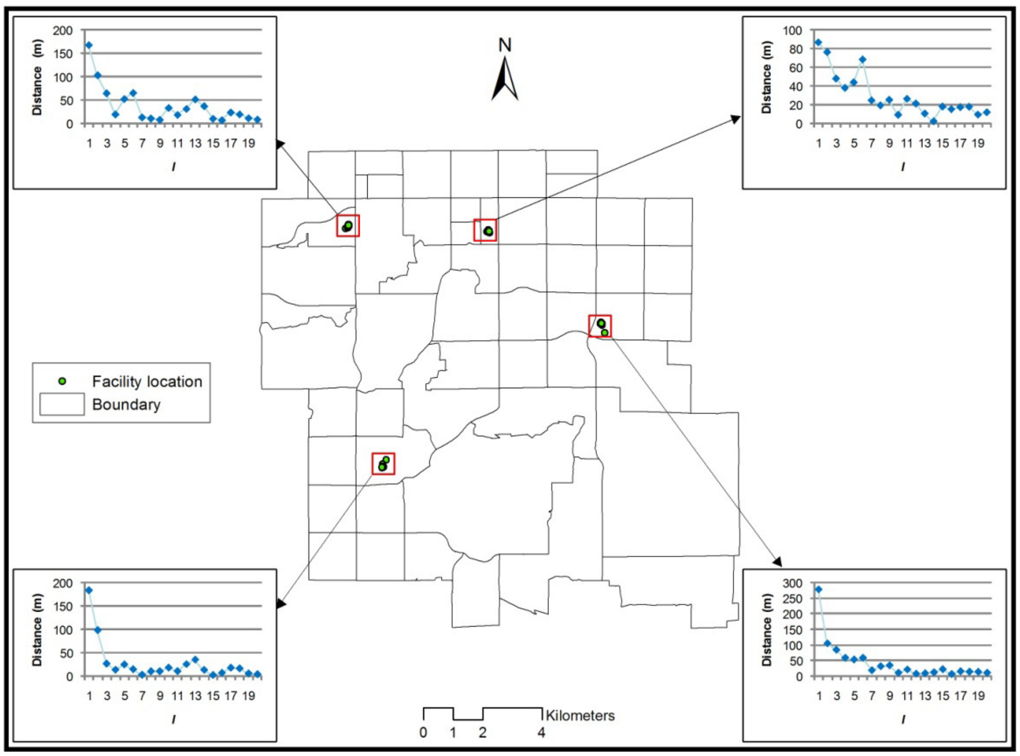


Figure 4.6: Solution for the study region with uniform continuous demand

4.7 Discussion and Conclusions

Continuous demand is common in regional planning that concerns siting facilities to provide social services for underlying demand. However, how to appropriately represent continuous demand in a digital environment is challenging and remains a key issue in facility location modeling. Traditional mathematical functions and fitted surfaces through spatial interpolation would inevitably introduce errors and uncertainties to the analysis results. Based on infill asymptotic theory, this research developed a spatial optimization approach integrating GIS functionalities and optimization techniques to solve the continuous multi-Weber problem in a way that reduces representation error and conforms to what is actually known about the demand region.

Essential here is how to represent continuous demand. It is well known that discrete-object and continuous-field are two primary models that are widely applied in GIS to represent geographic space. Though the surfaces based on the field view can reflect the continuous nature of the underlying demand, surfaces defined by mathematical functions or fitted by spatial interpolations are subject to significant errors. The discrete point approximation employed in this research can be a better option, both in terms of error minimization as well as no unrealistic assumptions about functional form. According to infill asymptotic theory, the discrete representation can be improved by increased demand point densities in order to approach the actual continuous distribution, which implies that the results derived will get closer to the theoretical optimal solution.

The empirical applications demonstrated the consistent convergence patterns in both facility locations and average travel distance (Figures 4.4 and 4.6). Though the solution changes among different representations, this variation becomes negligible as demand density increases sufficiently. For example, the distance between the first two solution layers was as large as $430\,m$ but became smaller as the demand density increased, ultimately less than $10\,m$ at solution convergence.

Given the spatial nature of facility location problems, it is necessary and crucial to incorporate GIS into the solution process. Based on GIS functionalities, this chapter developed a spatial optimization approach to solve the continuous multi-Weber problem, addressing continuous demand representation. Results from empirical applications showed the effectiveness of the proposed method, and its general applicability to support planning and decision making processes.

Chapter 5

CONCLUSIONS

Facility location models rely on spatial representations of facilities and demand in geographic space, either discretely or continuously. In one respect, abstraction using discrete points is often employed in facility location modeling as it can greatly facilitate model formulation and reduce computational complexity. However, errors and uncertainty introduced by such simplifications are well recognized and could lead to significant impacts on analysis results. In another respect, the continuous assumption is more reasonable in many practical situations where facilities can be located anywhere in continuous space and/or demand is continuously distributed over space. The issue is that continuous representation presents challenges to model specification and solution.

Of interest in this dissertation are facility location problems involving continuous representation, including the continuous Weber problem and the continuous multi-Weber problem. The former considers a single facility and the latter concerns siting multiple facilities simultaneously. Spatial optimization approaches integrating optimization techniques and GIS functionality were proposed for both problems. Application results were provided to demonstrate the effectiveness of developed methods and the significance of incorporating GIS into solution procedures.

5.1 Summary

This dissertation investigated the extensions of the Weber problem involving continuous demand and multiple facilities. First in Chapter 2, the implications of continuous surface approximation were addressed. Approaches for representing geographic space in a GIS environment were reviewed. In particular, spatial interpolation techniques are typically used to fit continuous surfaces in GIS. In addition, it was presented that errors in fitted surfaces through spatial interpolation were inevitable regardless of the interpolation approach employed. The empirical results showed that such errors can impact facility location analysis solutions, leading to cumulative errors and uncertainty in optimal facility locations and objective values.

Chapter 3 then focused on solving the continuous Weber problem. Given various simplified assumptions regarding the continuous demand distribution, a spatial representation method using discrete points was proposed built upon asymptotic theory, easily operationalized using GIS functionality. Discrete demand can provide a better approximation of underlying continuous distribution with increased point density, and this relaxes the demand distribution assumptions adopted by existing methods. Based on the proposed continuous representation approach, a spatial optimization heuristic was developed to solve the continuous Weber problem. It was applied to both uniformly and non-uniformly distributed demand in empirical studies. The application results showed that the facility locations as well as the objective values converged as the demand density increased, indicating the effectiveness of the developed approach.

Chapter 4 explored solution approaches for the continuous multi-Weber problem. It is more complex and difficult to solve than the continuous Weber problem because of the consideration of both location and allocation processes. Similar to the solution techniques for the continuous Weber problem, existing approaches addressing the multi-Weber problem largely rely on assumed theoretical functions describing continuous demand distribution. Mathematical properties of such functions then can be utilized in the solution procedure. The issue is that these functions are never simply known or given with certainty. Again, based on the spatial representation approach proposed in Chapter 3, a spatial optimization method was developed for solving the continuous multi-Weber problem. That is less restrictive in terms of an assumed demand distribution in existing approaches. Once the demand region is partitioned into several sub-regions, the problem is equivalent to solving the continuous Weber problem for each sub-region served by a single facility. The proposed method in Chapter 3 can then be applied. The results from the empirical applications involving both uniform and non-uniform continuous demand demonstrated that solutions obtained using the developed approach exhibited favorable characteristics. Futher, the solution procedure was greatly improved by incorporating GIS functions.

5.2 Future Research

Spatial representation of geographic space is a fundamental issue in facility location problems. By incorporating GIS functionality, this dissertation proposed spatial optimization approaches for solving the continuous Weber and multi-Weber problems, allowing for the representation of a continuous distribution based on known knowledge of underlying demand. While obtained analysis results were satisfactory, there is still room for potential improvement.

The spatial optimization approaches proposed in Chapter 3 and Chapter 4 can likely be improved in three respects. First, the objectives of the continuous Weber and multi-Weber problems are to minimize the average distance from demand to facilities providing service. The distance measure used in this dissertation is the Euclidean norm, as employed in the Weber problem. However, many other distance measures also can be used in practice. For example, travel distance or time along the road network is often used to represent geographic proximity to service providers. It is apparent that future research necessarily account for other distance measures reflecting the spatial context under study. Also, in both developed heuristics, the convergence criterion is based on the last several objective values, calculated as the sum of differences between the last objective value and the preceding ones. Alternatively, convergence tests using statistical methods can also be potential options to assess the convergence of a series of objective values. Further, varying demand densities for different sub-regions are used to reflect the spatial variation of the underlying continuous distribution. Such spatial variation also can be demonstrated by uniformly distributed points with varying weights. This situation is worth exploring in the future to examine the spatial bias for the analysis results of continuous location models.

The proposed spatial optimization heuristics incorporated existing approaches to solve the problems with discrete demand likely can also be enhanced. Once the continuous demand is discretized using sample points, the continuous Weber and multiple Weber problems are equivalent to the Weber problem and the multi-Weber problem, solved using Weiszfeld algorithm and ALTERNATE heuristic, respectively. Of course, since those two discrete counterparts have been extensively studied, other available solution approaches could be applied.

REFERENCES

Abu-Affash, A.K. and Katz, M. J., 2009. Improved bounds on the average distance to the Fermat-Weber center of a convex object. *Information Processing Letters*, 109 (6), 329–333.

Alexandris, G. and Giannikos, I., 2010. A new model for maximal coverage exploiting GIS capabilities. *European Journal of Operational Research*, 202 (2), 328–338.

Altınel, I.K., Durmaz, E., Aras, N., and ÖzkIsacIk, K.C., 2009. A location-allocation heuristic for the capacitated multi-facility Weber problem with probabilistic customer locations. *European Journal of Operational Research*, 198 (3), 790–799.

Anselin, L., 1989. *What is special about spatial data? Alternative perspectives spatial data analysis*. Santa Barbara, CA: National Center for Geographic Information and Analysis, Technical Report 89-4.

Bender, T., Hennes, H., Kalcsics, J., Melo, M.T., and Nickel, S., 2002. Location software and interface with GIS and supply chain management. *In*: Z. Drezner, H. Hamacher, eds. *Facility Location: Applications and Theory*. New York: Springer, 233–274.

Bennett, C.D. and Mirakhor, A., 1974. Optimal facility location with respect to several regions. *Journal of Regional Science*, 14 (1), 131–136.

Berry, B.J.L. and Baker, A.M., 1968. Geographic sampling. *In*: B.J.L. Berry and D.F. Marble, eds. *Spatial analysis: a reader in statistical geography*. Englewood, NJ: Prentice-Hall, 91–100.

Bhattacharya, R. and Bandyopadhyay, S., 2010. Solving conflicting bi-objective facility location problem by NSGA II evolutionary algorithm. *International Journal of Advanced Manufacturing Technology*, 51, 397–414.

Brandeau, M.L. and Chiu, S.S., 1989. An overview of representative problems in location research. *Management Science*, 25, 645–674.

Brimberg, J., Hansen, P., and Mladenović, N., 2006. Decomposition strategies for large-scale continuous location-allocation problems. *IMA Journal of Management Mathematics*, 17, 307–316.

Brimberg J., Hansen P., Mladenović, N., and Salhi, S., 2008. A survey of solution methods for the continuous location - allocation problem. *International Journal of Operations Research*, 5, 1–12.

Brimberg, J. and Mladenović, N., 1996. Solving the continuous location-allocation problem with Tabu search. *Studies in Locational Analysis*, 8, 23–32.

Burrough, P.A., 1996. Natural objects with indeterminate boundaries. *In*: P.A. Burrough and A.U. Frank, eds. *Geographic objects with indeterminate boundaries*. London: Taylor and Francis, 2–28.

Burrough, P.A. and McDonnell, R.A., 1998. *Principles of geographical information systems*. Oxford: Oxford University Press.

Burt, J.E., Barber, G.M., and Rigby, D.L., 2009. *Elementary statistics for geographers*. 3rd ed. New York: Guilford Press.

Carlisle, B., 2005. Modelling the spatial distribution of DEM error. *Transactions in GIS*, 9 (4), 521–540.

Carmi, P., Har-Peled, S., and Katz, M.J., 2005. On the Fermat-Weber center of a convex object. *Computational Geometry - Theory and Applications*, 32 (3), 188–195.

Carrizosa, E., Muñoz-Márquez, M., and Puerto, J., 1998. The Weber problem with regional demand. *European Journal of Operational Research*, 104, 358–365.

Chandrasekaran, R. and Tamir, A. 1989. Open questions concerning Weiszfeld's algorithm for the Fermat–Weber location problem. *Mathematical Programming*, 44 (3), 293–295.

Chen, R. 2001. Optimal location of a single facility with circular demand areas. *Computers and Mathematics with Applications*, 41, 1049–1061.

Church, R.L., 1999. Location modelling and GIS. *In*: P.A. Longley, M.F. Goodchild, D.J. Maguire, D.W. Rhind, eds. *Geographical information systems*. New York: Wiley.

Church, R.L., 2001. Spatial optimization models. *In* N.J. Smelser and P.B. Baltes, eds. *International Encyclopedia of the Social & Behavioral Sciences*. Oxford, UK: Elsevier Science Ltd, 14,811–18.

Church, R.L., 2002. Geographic information systems and location science. *Computers & Operations Research*, 29, 541–562.

Church, R.L. and Murray, A.T., 2009. *Business site selection, location analysis and GIS*. New York: Wiley.

Church, R.L., and ReVelle, C., 1974. The maximal covering location problem. *Papers of the Regional Science Association*, 32, 101–118.

Church, R.L. and Sorensen, P., 1996. Integrating normative location models into GIS. *In*: P. Longley, M. Batty, eds. *Spatial analysis: Modeling in a GIS environment*. Cambridge, UK: GeoInformation International.

Cooper, L., 1963. Location-allocation problems. *Operations Research*, 11, 331–343.

Cooper, L., 1964. Heuristic methods for location allocation problems. *SIAM Review*, 6, 37–52.

Couclelis, H., 1992. People manipulate objects (but cultivate fields): beyond the raster vector debate in GIS. *In*: A.U. Frank, I. Campari, and U. Formentini, eds. *Theory and methods of spatio-temporal reasoning in geographic space*. Berlin: Springer, 65–77.

Cova, T.J. and Church, R.L., 2000. Contiguity constraints for single-region site search problems. *Geographical Analysis*, 32, 306–329.

Cova, T.J. and Goodchild, M.F., 2002. Extending geographical representation to include fields of spatial objects. *International Journal of Geographical Information Science*, 16, 509–532.

Cressie, N., 1993. *Statistics for spatial data*. Revised Ed. New York: Wiley.

Cressie, N. and Wikle, C.K., 2011. *Statistics for spatio-temporal data*. New York: Wiley.

Cromley, R.G., Lin, J., and Merwin, D.A., 2012. Evaluating representation and scale error in the maximal covering location problem using GIS and intelligent areal interpolation. *International Journal of Geographical Information Science*, 26 (3), 495–517.

Dasci, A. and Verter, V., 2001. A continuous model for production-distribution system design. *European Journal of Operational Research*, 129, 287–298.

Dasci, A and Verter, V., 2005. Evaluation of plant focus strategies: A continuous approximation framework. *Annals of Operations Research*, 136, 303–327.

de Smith, M.J., Goodchild, M.F., and Longley, P.A., 2009. *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*. 3rd ed. Leicester: Troubador Publishing Ltd.

Densham, P.J., 1994. Integrating GIS and spatial modelling: visual interactive modelling and location selection. *Geographical Systems*, 1, 203–219.

Downs, J.A., Gates, R.J., and Murray, A.T., 2008. Estimating carrying capacity for sandhill cranes using habitat suitability and spatial optimization models. *Ecological Modelling*, 214, 284–292.

Drezner, Z., 1986. Location of Regional Facilities. *Naval Research Logistics Quarterly*, 33, 523–529.

Drezner, Z., 1995. Replacing discrete demand with continuous demand. *In*: Z. Drezner, eds. *Facility Location: A Survey of Applications and Methods*. New York: Springer–Verlag, 33–42.

Drezner, T. and Drezner, Z., 1997. Replacing continuous demand with discrete demand in a competitive location model. *Naval Research Logistics*, 44, 81–95.

Drezner, Z., Klamroth, K., Schöbel, A., and Wesolowsky, G.O., 2002. The Weber problem. *In*: Z. Drezner, H. Hamacher, eds. *Facility location: applications and theory*. Berlin: Springer, 1–36.

Drezner, Z. and Suzuki, A., 2010. Covering continuous demand in the plane. *Journal of the Operational Research Society*, 61, 878–881.

Drezner, Z. and Weslowsky, G.O., 1978. A note on optimal facility location with respect to several regions. *Journal of Regional Science*, 18 (2), 303.

Drezner, Z. and Wesolowsky, G.O., 1980. Optimal location of a facility relative to area demands. *Naval Research Logistics*, 27, 199–206.

Eicher, C.L. and Brewer, C.A., 2001. Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science*, 28 (2): 125–138.

Egenhofer, M.J., Glasgow, J., Gunther, O., Herring, J.R., and Peuquet, D.J., 1999. Progress in computational methods for representing geographical objects. *International Journal of Geographical Information Science*, 13, 775–796.

Fekete, S.P., Mitchell, J.S.B., and Beurer, K., 2005. On the continuous Fermat–Weber problem. *Operations Research*, 53, 61–76.

Fisher, P.F. and Langford, M., 1995. Modelling the errors in areal interpolation between zonal systems by Monte Carlo simulation. *Environment and Planning A*, 27, 211–224.

Francis, R.L. and Lowe, T.J., 2011. Comparative error bound theory for three location models: continuous demand versus discrete demand. *Top*, DOI 10.1007/s11750-011-0244-2.

Francis, R.L., Lowe, T.J., Rayco, M.B., and Tamir, A., 2009. Aggregation error for location models: survey and analysis. *Annals of Operations Research*, 167, 171–208.

Franco, C.V., Rodríguez-Chía, A.M., and Espejo-Miranda, I., 2008. The single facility location problem with average-distances, *Top*, 16, 164–194.

Frank, A.U., 1992. Spatial concepts, geometric data models and geometric data structures. *Computers & Geosciences*, 18 (4), 409–417.

Gastner, M.T. and Newman, M.E.J., 2006. Optimal design of spatial distribution networks. *Physical Review E*, 74 (1), 016117.

Ghosh, A. and Craig, C.S., 1984. A location-allocation model for facility planning in a competitive environment. *Geographic Analysis*, 16, 39–51.

Gong, J., Li, Z., Zhu, Q., Shu, H., and Zhou, Y., 2000. Effects of various factors on the accuracy of DEMs: an intensive experimental investigation. *Photogrammetric Engineering and Remote Sensing*, 66 (9), 1113–1117.

Goodchild, M.F., 1992. Geographical data modeling. *Computers and Geosciences*, 18, 400–408.

Goodchild, M.F., 2004. GIScience: geography, form, and process. *Annals of the Association of American Geographers*, 94, 709–714.

Goodchild, M.F., 2010. Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science*, 1, 3–20.

Goodchild, M.F., Anselin, L., and Deichmann, U., 1993. A framework for the areal interpolation of socioeconomic data. *Environment and Planning A*, 25, 383–397.

Goodchild, M.F. and Gopal, S., eds., 1989. *Accuracy of spatial databases*. London: Taylor and Francis.

Goodchild, M.F. and Haining, R.P., 2004. GIS and spatial data analysis: converging perspectives. *Papers in Regional Science*, 83, 363–385.

Goodchild, M.F., Yuan, M., and Cova, T.J., 2007. Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21, 239–260.

Goodchild, M., Haining, R., Wise, S., and 12 others., 1992. Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems*, 6, 407–423.

Gugat, M. and Pfeiffer, B., 2007. Weber problems with mixed distances and regional demand. *Mathematical Methods of Operations Research*, 66, 419–449.

Hakimi, S.L., 1964. Optimum location of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12, 450–459.

Hakimi, S.L., 1965. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, 13, 462–475.

Heuvelink, G.B.M., 1998. *Error propagation in environmental modelling with GIS*. London: Taylor and Francis.

Heuvelink, G.B.M. and Burrough, P.A., 1993. Error propagation in cartographic modelling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems*, 7 (3), 231–246.

Hillsman, E.L. and Rhoda, R., 1978. Errors in Measuring Distances from Populations to Service Centers. *Annals of the Regional Science Association*, 12, 74–88.

Hoover, E.M., 1937. *Location Theory and the Shoe and Leather Industries*. Cambridge, Mass.: Harvard University Press.

Host, G., Omre, H., and Switzer, P., 1995. Spatial interpolation errors for monitoring data. *Journal of the American Statistical Association*, 90 (431), 853–861.

Hotelling, H., 1929. Stability in Competition. *Economic Journal*, 39, 41–57.

Houck, C.R., Joines, J.A., and Kay. M.G., 1996. Comparison of genetic algorithms, random restart and two-opt switching for solving large location-allocation problems. *European Journal of Operational Research*, 20, 387–396.

Janssen, S., Dumont, G., Fierens, F., and Mensink, C., 2008. Spatial interpolation of air pollution measurements using CORINE land cover data. *Atmospheric Environment*, 42, 4884–4903.

Kemp, K.K., 1997a. Fields as a framework for integrating GIS and environmental process models. Part I: representing spatial continuity. *Transactions in GIS*, 1, 219–234.

Kemp, K.K., 1997b. Fields as a framework for integrating GIS and environmental process models. Part 2: specifying field variables. *Transactions in GIS*, 1, 235–246.

Kim, H. and O'Kelly, M.E., 2009. Reliable p-hub location problems in telecommunication networks. *Geographical Analysis*, 41 (3), 283–306.

Kjenstad, K., 2006. On the integration of object-based models and field-based models in GIS. *International Journal of Geographical Information Science*, 20, 491–509.

Koshizuka, T. and Kurita, O., 1991. Approximate formulas of average distances associated with regions and their applications to location-problems. *Mathematical Programming*, 52, 99–123.

Kuhn, H.W. and Kuenne, R.E., 1962. An efficient algorithm for the numerical solution of the generalized Weber problem. *Journal of Regional Science*, 4, 21–34.

Kuhn, H.W. 1973. A note on Fermat's problem. *Mathematical Programming*, 4, 98–107.

Kyriakidis, P.C. and Goodchild, M.F., 2006. On the prediction error variance of three common spatial interpolation schemes. *International Journal of Geographical Information Science*, 20 (8), 823–855.

Lam, N.S., 1983. Spatial interpolation methods: a review. *The American Cartographer*, 10 (2), 129–149.

Langford, M. and Unwin, D.J., 1994. Generating and mapping population density surfaces within a geographical information system. *Cartographic Journal*, 31 (1), 21–26.

Longley, P.A., Goodchild, M.F., Maquire, D.J., and Rhind, D.W., 2011. *Geographic information systems and science*. 3rd ed. New York: Wiley.

Longley, P.A. and Batty, M., (eds.) 1996. *Spatial Analysis: Modelling in a GIS Environment*. Cambridge: GeoInformation International.

Love, R.F., 1972. A computational procedure for optimally locating a facility with respect to several rectangular regions. *Journal of Regional Science*, 12, 233–242.

Love, R.F. and Yeong, W., 1981. A stopping rule for facilities location algorithms. *AIIE Transactions*, 13, 357–362.

Maruchek, A.S. and Aly, A.A., 1981. An efficient algorithm for the location-allocation problem with rectangular regions. *Naval Research Logistics*, 28, 309–323.

Matisziw, T.C. and Murray, A.T., 2009. Siting a facility in continuous space to maximize coverage of continuously distributed demand. *Socio-Economic Planning Sciences*, 43, 131–139.

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

Mennis, J., 2003. Generating surface models of population using dasymetric mapping. *The Professional Geographer*, 55 (1), 31–42.

Miehle, W., 1958. Link length minimization in networks. *Operations Research*, 6, 232–243.

Miller, H.J., 1996. GIS and geometric representation in facility location problems. *International Journal of Geographical Information Systems*, 10, 791–816.

Miller, H.J. and Wentz, E.A., 2003. Representation and spatial analysis in geographic information systems. *Annals of the Association of American Geographers*, 93 (3), 574–594.

Mitas, L. and Mitasova, H., 1999. Spatial interpolation. *In*: P.A. Longley, M.F. Goodchild, D.J. Maguire, and D.W. Rhind, eds. *Geographical information systems: principles, techniques, management and applications*. New York: John Wiley & Sons, 481–492.

Morrison, J.L., 1971. *Method-produced error in isarithmic mapping*. Washington, DC: American Congress on Surveying and Mapping, 76 p.

Murat, A., Verter, V., and Laporte, G., 2010. A continuous analysis framework for the solution of location-allocation problems with dense demand. *Computers and Operations Research*, 37, 123–136.

Murat, A., Verter, V., and Laporte, G., 2011. A multi-dimensional shooting algorithm for the two-facility location-allocation problem with dense demand. *Computers & Operations Research*, 38, 450–463.

Murray, A.T., 2003. Site placement uncertainty in location analysis. *Computers, Environment and Urban Systems*, 27, 205–221.

Murray, A.T., 2005. Geography in coverage modeling: exploiting spatial structure to address complementary partial service of areas. *Annals of the Association of American Geographers*, 95, 761–772.

Murray, A.T., 2008. Economic geography: location theory. *In*: R. Kitchin, N. Thrift, eds. *International encyclopedia of human geography*. Oxford: Elsevier.

Murray, A.T., 2010. Advances in location modeling: GIS linkages and contributions. *Journal of Geographical Systems*, 12, 335–354.

Murray, A.T. and Church, R.L., 1996. Analyzing cliques for imposing adjacency restrictions in forest models. *Forest Science*, 42, 166–175.

Murray, A.T. and O'Kelly, M.E., 2002. Assessing representation error in point-based coverage modeling. *Journal of Geographical Systems*, 4, 171–191.

Murray, A.T., O'Kelly, M.E., and Church, R.L., 2008. Regional service coverage modeling. *Computers and Operations Research*, 35, 339–355.

Myers, D.E., 1994. Spatial interpolation: an overview. *Geoderma*, 62, 17–28.

Oksanen, J. and Sarjakoski, T., 2005. Error propagation of DEM-based surface derivatives. *Computers & Geosciences*, 31, 1015–1027.

Openshaw, S., 1983. *The modifiable areal unit problem*. Concepts and techniques in modern geography (CATMOG) No. 38. Norwich: Geo Books.

Ouyang, Y. and Daganzo, C.F., 2006. Discretization and validation of the continuum approximation scheme for terminal system design. *Transportation Science*, 40, 89–98.

Peuquet, D., 1988. Representation of geographic space: toward a conceptual synthesis. *Annals of the Association of American Geographers*, 78, 375–394.

Plastria, F., 1995. Continuous location problems. *In*: Z. Drezner, eds. *Facility Location. A Survey of Applications and Methods*. Berlin: Springer, 85–127.

Puerto, J. and Rodríguez-Chía, A.M., 2011. On the structure of the solution set for the single facility location problem with average distances. *Mathematical Programming*, 128 (1-2), 373–401.

Rao, J.R. and Varma, N.H., 1985. Stochastic multi-facility minisum location problem involving Euclidean distances. *Operations Research*, 22, 232–240.

ReVelle, C. and Swain, R., 1970. Central facilities location. *Geographical Analysis*, 2, 30–34.

Righini, G. and Zaniboni, L., 2007. A branch-and-price algorithm for the multi-source Weber problem. *International Journal of Operational Research*, 2 (2): 188–207.

Rosing, K.E., 1992. An Optimal Method for Solving the (Generalized) Multi-Weber Problem. *European Journal of Operational Research*, 58, 414–426.

Salhi, S. and Gamal, M.D.H., 2003. A genetic algorithm based approach for the uncapacitated continuous location-allocation problem. *Annals of Operations Research*, 123, 203–222.

Stein, M., 1999. *Interpolation of spatial data: some theory for kriging*. Springer Series in Statistics. New York: Springer Verlag.

Suzuki, A. and Okabe, A., 1995. Using Voronoi diagrams. *In*: Z. Drezner, eds. *Facility location: a survey of applications and methods*. New York: Springer.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234–240.

Tomczak, M., 1998. Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW) – cross-validation/Jackknife approach. *Journal of Geographic Information and Decision Analysis*, 2 (2), 18–30.

Tsao, Y.C. and Lu, J.C., 2012. A supply chain network design considering transportation cost discounts. *Transportation Research Part E*, 48, 401–414.

Vaughan, R., 1984. Approximate formulas for average distances associated with zones. *Transportation Science*, 18 (3), 231–244.

Wang, C.Y., Gao, C.Y., and Shi, Z.J., 1997. An algorithm for continuous type optimal location problem. *Computational Optimization and Applications*, 7, 239–53.

Weber, A., 1909. Uber den standort der Industrien. Tubingen. (Fredrich, C.J., translation, 1929, *Theory of the Location of Industries*, Chicago: University of Chicago Press).

Weber, D.D. and Englund, E.J., 1992. Evaluation and comparison of spatial interpolators. *Mathematical Geology*, 24 (4), 381–391.

Wei, H., Murray, A.T., and Xiao, N., 2006. Solving the continuous space p-center problem: planning application issues. *IMA Journal of Management Mathematics*, 17, 413–425.

Weiszfeld, E., 1936. Sur le point pour lequel la somme des distances de n points donnés est minimum. *The Tohoku Mathematical Journal*, 43, 355–386.

Wesolowsky, G.O., 1977. The Weber problem with rectangular distances and randomly distributed destinations. *Journal of Regional Science*, 17, 53–60.

Wesolowsky, G.O., 1993. The Weber problem: history and perspective. *Location Science*, 1, 5–23.

Wesolowsky, G.O. and Love, R.F., 1972. A nonlinear approximation method for solving a generalized rectangular distance weber problem. *Management Science*, 18 (11), 656–663.

Winter, S., 1998. Bridging vector and raster representation in GIS. *In*: Laurini, R., Makki, K., Pissinou, N. (Eds.): *Proceedings of the 6th international symposium on Advances in Geographic Information Systems (ACM-GIS '98),* November 6-7, 1998, Washington, DC, USA. New York: ACM Press, 57–62.

Wood, J.D. and Fisher, P.F., 1993. Assessing interpolation accuracy in elevation models. *IEEE Computer Graphics and Applications*, 13 (2), 48–56.

Worboys, M.F. and Duckham, M., 2004. *GIS: a computing perspective*. 2nd ed. Boca Raton, FL: CRC Press.

Wu, C. and Murray, A.T., 2007. Population estimation using Landsat Enhanced Thematic Mapper imagery. *Geographical Analysis*, 39, 26–43.

Wu, X. and Murray, A.T., 2007. Spatial contiguity optimization in land acquisition. *Journal of Land Use Science*, 2, 243–256.

Yuan, M., 2001. Representing complex geographic phenomena with both object- and field-like properties. *Cartography and Geographic Information Science*, 28, 83–96.

Zimmerman, D., Pavlik C., Ruggles, A., and Armstrong, M.P., 1999. An experimental comparison of ordinary and universal kriging and inverse distance weighting. *Mathematical Geology*, 31 (4), 375–390.