Addressing Geographic Uncertainty In

Spatial Optimization

by

Ran Wei

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

Approved January 2013 by the
Graduate Supervisory Committee:

Alan T. Murray, Chair
Luc Anselin
Sergio J. Rey
Elizabeth A. Mack

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

There exist many facets of error and uncertainty in digital spatial information. As error or uncertainty will not likely ever be completely eliminated, a better understanding of its impacts is necessary. Spatial analytical approaches, in particular, must somehow address data quality issues. This can range from evaluating impacts of potential data uncertainty in planning processes that make use of methods to devising methods that explicitly account for error/uncertainty. To date, little has been done to structure methods accounting for error. This research focuses on developing methods to address geographic data uncertainty in spatial optimization. An integrated approach that characterizes uncertainty impacts by constructing and solving a new multi-objective model that explicitly incorporates facets of data uncertainty is developed. Empirical findings illustrate that the proposed approaches can be applied to evaluate the impacts of data uncertainty with statistical confidence, which moves beyond popular practices of simulating errors in data. Spatial uncertainty impacts are evaluated in two contexts: harvest scheduling and sex offender residency. Owing to the integration of spatial uncertainty, the detailed multi-objective models are more complex and computationally challenging to solve. As a result, a new multi-objective evolutionary algorithm is developed to address the computational challenges posed. The proposed algorithm incorporates problem-specific spatial knowledge to significantly enhance the capability of the evolutionary algorithm for solving the model.

Dedicated to my parents

ACKNOWLEDGEMENTS

Over the past three years I have received support and encouragement from a great number of people. Dr. Alan T. Murray, my advisor, is the first person I have to thank. This dissertation would not be possible without his patient guidance and consistent support. He is always there when I need help. He showed me how to perform solid research, taught me how to write journal articles, and gave me invaluable advice on my career development. I am forever grateful for his mentorship.

I also want to express my gratitude to my committee members, Dr. Luc Anselin, Dr. Sergio J. Rey and Dr. Elizabeth A. Mack, for their instructions to my research and career. It is my great fortune to study in the GeoDa Center and School of Geographical Science and Urban Planning at the ASU.

The funding provided by the National Science Foundation under grants 0924001 and 0922737, the Completion Fellowship from the Arizona State University Graduate College, and the Benjamin H. Stevens Fellowship from the Regional Science Association International is appreciated for financially supporting this dissertation.

I would like to give my deepest thanks to my parents and sisters for always standing by me and supporting my graduate studies. Finally, I want to thank my husband for his help in my study and accompanying me during the process.

TABLE OF CENTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Background

Geographic information systems (GISs) provide the capacity to digitally create, store, manipulate, analyze and display various types of geographic information (Longley et al. 2011). While these functionalities enable the handling of spatial data in a much more rapid and precise way than traditional paper-based approaches, uncertainties still remain in geographic information and will not likely ever be completely eliminated.

There are various sources of uncertainty in digital geographic information. Uncertainty can arise from the inaccuracy of source documents and processing. As an example, the original paper maps may be distorted because of folding, stretching and humidity; when digitizing the paper maps, additional errors could be introduced because of an operator's control of the cursor (Goodchild 1989); if satellite images are employed to generate a digital vector map, the resolution of images and the raster-to-vector transformation process could all result in errors in the final vector map; and geocoding errors are considerable (see Cayo and Talbot 2003 and Goldberg 2011). Uncertainty can also be due to the vagueness or ambiguity in definitions of classes of objects (Fisher 1999). For instance, a forest region is usually delineated into a series of “stands” where species composition, density and size are assumed to be homogenous. However, there are always transition zones between stands and stand boundaries that are actually vague in reality (Goodchild 1989). Uncertainty can be attributed to missing or inadequate information as well. Finally it is common to estimate attributes by interpolation techniques, which could lead to more errors.

When digital spatial information is relied upon in spatial analysis or for decision making, errors or uncertainties contained in the data will propagate through analysis and affect decisions (Longley et al. 2011). A significant literature has demonstrated error propagation effects, ranging from simple analysis, like polygon area calculation (Goodchild et al. 1999) and map algebra operations (Arbia et al. 1999; Griffith et al. 1999; Abbapour et al. 2003; Leung et al. 2004b), to complex statistical analysis and modeling, such as regression modeling (Heuvelink et al. 1989; Das et al. 2002; Griffith et al. 2007) , clustering methods (Goovaerts 2006; Malizia 2012), and spatial optimization models (Goodchild 1984; Murray 2003; Aerts et al. 2003; Bonneu and Thomas-Agnan 2009). In fact, any spatial analytical method or decision making involving geographic information could be impacted by spatial data uncertainty. To ensure the appropriate use of spatial information, it is essential to evaluate whether the accuracy of information used is sufficient for the intended application. If it is not, this may result in inaccurate analysis results and biased/incorrect decisions (Heuvelink 1998). As an example, Goodchild et al. (1999) showed that a parcel area could range from 6,000 square meter to 14,000 square meter given 10 meter positional uncertainty; Aerts et al. (2003) estimated that construction costs for a ski run might increase approximately 32 percent when accounting for errors in a digital elevation model; and, Goovaerts (2006) found that ignoring spatial uncertainty in cancer risk estimates could lead to misallocation of medical resources.

Given error propagation and its impacts on decision making, considerable research effort have been made in developing methods to account for spatial data uncertainty in applications. These can be generally categorized into analytically based methods and simulation-based methods (Heuvelink et al. 2002; Shi et al. 2004). The analytically based methods, such as the Taylor series approximation and Rosenblueth's method (see Heuvelink 1998), link input spatial data to output results by deriving an operation function, $y = f(x)$, where x represents input data and y is the analysis results (Leung et al. 2004a). The error or uncertainty in input data can then be transformed into impacts on analysis results using the operation function. Analytical methods can therefore theoretically generate statistical descriptions of the implications of spatial uncertainty if the operation function is known. However, it is usually difficult to derive operation functions, especially when complicated analyses are involved. So far, analytically based methods are mainly applied to straightforward problems, such as length and area measurement (Leung et al. 2004d) and overlay analysis (Arbia et al. 1998; Griffith et al. 1999; Leung et al. 2004c; Shi et al. 2004). In contrast to analytically based method, simulation-based methods mimic errors in spatial data and then compare analysis scenarios given simulated error. This is relatively easy to implement and has been applied to many sophisticated decision making contexts (see Fisher 1991; Aerts et al. 2003; Brown and Heuvelink 2007; Bruin et al. 2008; Hevelink et al. 2010; Bonneu et al. 2011).

Of interest in this research is error propagation in spatial optimization, a spatial analytical method that contributes to a wide range of environmental and urban planning problems, including transportation, districting, natural resource management, and land-use planning, among others. Overviews of spatial optimization may be found in Church (2001), Xiao (2008), and Tong and Murray (2012). A spatial optimization problem involves determining the best location/assignment of people, goods or activities interacting across space while some constraints or conditions are maintained. A problem is therefore represented as a mathematical model, with decision variables, an objective(s), and constraints (Murray 2010; Tong and Murray 2012). Given the multiple components in the model, the operation function, $y = f(x)$, is typically challenging to derive. As a result, the simulation-based method is popular in evaluating data uncertainty impacts in spatial optimization. Examples can be found in Goodchild (1984), Hodgson (1991), Aerts et al. (2003), Murray (2003), Heuvelink et al. (2010) and Bonneu et al. (2011). Nevertheless, simulation relies on a limited sample size and makes it impossible to assess uncertainty impacts with true statistical confidence (Lilburne and Tarantola 2009). In addition, the computational load associated with the simulation-based method is generally high, especially for complex models (Hevelink et al. 2010).

1.2 Research Objectives

Given the challenges in understanding the impacts of geographic uncertainty in spatial optimization, this research has two core objectives. The first objective is to develop methods that enable the impacts of spatial data uncertainty on spatial optimization models to be evaluated with a degree of statistical certainty. More specifically, new multi-objective models that explicitly account for geographic data uncertainty will be proposed. The second objective is to develop a new heuristic for the proposed multi-objective optimization models. The heuristic can ensure the identification of high-quality solutions within a reasonable amount of time.

1.3 Organization of the research

This research is organized as follows. Chapter 2 starts with a review of existing work associated with uncertainty in spatial optimization, and then discusses the implications of spatial data uncertainty in a dispersion model. Next, a new multi-objective model is proposed to incorporate spatial data uncertainty. This is followed by an application of the developed model to evaluate the implications of offender residency restriction laws in the Phoenix metropolitan area.

Chapter 3 evaluates the impacts of spatial data uncertainty in dispersion modeling in the context of harvest scheduling, where separation requirements are determined by contiguity instead of distance as studied in Chapter 2. After reviewing existing approaches to address uncertainty issues in harvest scheduling, an algorithm to assess contiguity uncertainty is developed and an alternative modeling approach that explicitly account for spatial uncertainty is presented and compared with the model detailed in Chapter 2. The implications of spatial uncertainty in harvest scheduling are then examined in a forest region located in northern California.

Chapter 4 develops a new multi-objective evolutionary algorithm for the models presented in Chapters 2 and 3. The chapter begins by reviewing existing solution techniques for a dispersion model, then a genetic algorithm and its advantages and design issues in multi-objective optimization problems are introduced. Next, a multi-objective genetic algorithm is proposed. Finally, computational results are presented and discussed.

The final chapter, Chapter 5, summarizes the research results of this dissertation and provides concluding comments. In addition, future research directions are discussed.

CHAPTER 2

INCORPORATING GEOGRAPHIC UNCERTIANTY IN SPATIAL OPTIMIZATION*

As presented in Chapter 1, this dissertation seeks to better understand the impacts of geographic uncertainty on spatial optimization. This chapter develops an approach to incorporate spatial data uncertainty into a dispersion model, a particular type of spatial optimization problem. The integrated approach characterizes uncertainty impacts by constructing and solving a new multi-objective model that explicitly accounts for data uncertainty.

2.1 Introduction

The growth and popularity of geographic information systems (GIS) and associated digital spatial information is remarkable, fundamentally changing planning and management processes as well as spawning the development of advanced spatial analytical methods. While there has been much done to identify and ameliorate data error and uncertainty issues, and more generally improve overall data quality, imperfections in spatial information remain. This ultimately may be attributed to the fact that the three dimensional real world is abstracted as a simplified digital representation in modeling and analyses (Longley et al. 2011).

---

* This chapter represents a slightly revised version of a paper published in *International Journal of Geographical Information Science*, co-authored with Dr. Alan T. Murray.

When spatial data is relied upon in applications, error or uncertainty in data will not disappear but rather will be propagated through subsequent manipulation, processing and analysis (Longley et al. 2011). There are many examples in the literature that demonstrate this, including Arbia et al. (1998) and Abbaspour et al. (2003), where error propagation is tracked in simple map algebra operations like overlay. More sophisticated analytical approaches have been subjected to data uncertainty impacts as well, including Fisher (1991), Heuvelink (1998), Bruin et al. (2001), Aerts et al. (2003) and Heuvelink et al. (2007), to name but a few. In particular, Aerts et al. (2003) found a significant difference in construction costs for a ski run when accounting for potential errors in derived slopes using in a land use optimization model. What the above work highlights is that all spatial analysis is likely impacted by data error/uncertainty in some way. Better understanding impacts and implications of data uncertainty in spatial analysis using various methods therefore requires techniques to explicitly account for uncertainty, lest the analysis be biased and unreliable.

While uncertainty impacts associated with any and all spatial analytical methods are a concern, this research focuses on spatial optimization, an analytical method that contributes to transportation, retail, natural resource management, location modeling, medical geography, land use planning and districting, among others. The importance of optimization in GIScience is well recognized, spanning database structure and access, algorithm design, cartographic display and spatial analysis (see Church 1999; Murray 2007). Overviews of spatial optimization may be found in Church (2001), Xiao (2008) and Tong and Murray (2012). A spatial optimization problem is one where there are decisions to be made regarding the placement and/or allocation of a good or service, and the best decisions are sought subject to maintaining any constraints or conditions. A problem or issue is therefore represented (implicitly or explicitly) as a mathematical model, with decision variables, an objective(s) and constraints (Tong and Murray 2012). Spatial optimization research has in fact long recognized data uncertainty issues. For example, Cooper (1974) observed that demand locations may not be known with certainty in location models and Goodchild (1984) discussed that potential demand locations in a location-allocation model for retail site selection may be somewhat inaccurate. Reviews of work on one class of spatial optimization problem, location models, associated with data uncertainty can be found in Murray (2003) and Snyder (2006).

Even though data uncertainty in spatial optimization models has been widely acknowledged, a generally applicable method to account for uncertainty and evaluate its impacts on modeling results remains elusive (Church 1999; Murray 2003). One commonly used method is simulating error in spatial data. This has involved intentional perturbation of input or output data to mimic error, then comparing analyses obtained with and without simulated error (see Goodchild 1984; Hodgson 1991; Aerts et al. 2003; Murray 2003; Beech et al. 2008). This makes intuitive sense, and provides some capacity for evaluation and sensitivity assessment. However, simulation along these lines is computationally intensive, effectively necessitating an infinite number of combinatorial possibilities to be considered in practice. It is impossible to consider all combinations of variability, which is why samples have been relied upon. Unfortunately, uncertainty necessarily remains, often absent any type of certainty bounds on derived findings or results (Salema et al. 2007; Ascough et al. 2008).

This chapter aims to develop an explicit approach for considering error/uncertainty in spatial optimization. To illustrate this, a particular type of spatial optimization problem known as a dispersion model is considered. Uncertainty in geographic proximity is expressly represented and incorporated into the model, enabling evaluation and assessment with a degree of statistical certainty. The next section reviews existing literature on data uncertainty in spatial optimization. This is followed by the introduction of the optimization model being considered here. A discussion highlighting spatial uncertainty associated with the use and application of this dispersion model is then presented. A new model is then introduced that enables data uncertainty to be explicitly considered. Application results are presented to illustrate the effectiveness of the new model. Finally, a discussion and concluding comments are given.

2.2 Background

Spatial optimization models are implicit or explicit mathematical representations of planning problems, necessarily abstractions of reality, and could be affected by uncertainties or errors in various ways. Such uncertainty may arise in model specification, solution sub-optimality, attribute variation (estimates/classification/precision), spatial representation, locational accuracy and proximity interpretation, among others. Many studies have sought to better understand elements of uncertainty along these lines. Table 2.1 lists representative work associated with uncertainty in spatial optimization grouped in six categories. The last three, spatial representation, locational accuracy and proximity interpretation, are particularly relevant to issues of spatial uncertainty. As a result, the review that follows focuses on these three categories. While uncertainty in model specification, sub-optimality and attributes are no doubt important, they are generally not associated directly with error in spatial position.

Table 2.1: Literature addressing one or more aspects of uncertainty in spatial optimization models

Model specification	Solution optimality	Attribute variation	Spatial representation	Locational accuracy	Proximity interpretation
Ratick and White (1988)	Karp (1977)	Hodder and Dincer (1986)	Goodchild (1979)	Cooper (1974)	Bach (1981)
Eiselt and Laporte (1995)	Johnson et al. (1989)	Carson and Batta (1990)	Daskin et al. (1989)	Mirchandani and Odoni (1979)	Mirchandani and Oudjit (1980)
Kuby et al. (2011)	Rardin and Uzsoy (2001)	Zhao and Kockelman (2002)	Current and Schilling (1990)	Weaver and Church (1983)	Berman and Odoni (1982)
	Brookes (2001)	Daskin et al. (2002)	Fotheringham et al. (1995)	Goodchild (1984)	Hodgson (1991)
	Aerts and Heuvelink (2002)	Salema et al. (2007)	Miller (1996)	Hodgson (1991)	Brimberg, J. and R. F. Love (1995).
	Li and Yeh (2005)	Wagner et al. (2009)	Drezner and Drezner (1997)	Murray (2003)	Andersson et al. (1998)
		Klibi et al. (2010)	Murray and Gottsegen (1997)	Aerts et al. (2003)	Plastria, F. (2001)
			Francis et al. (1999)	Bonneu and Thomas-Agnan (2009)	AltInel et al. (2009)
			Murray and Weintraub (2002)	AltInel et al. (2009)	
			Murray and O’Kelly (2002)		
			Emir-Farinas and Francis (2005)		
			Murray (2005)		
			Francis et al. (2009)		

There has been much work on spatial representation associated with uncertainty and error in spatial optimization modeling, as suggested by the representative work noted in Table 2.1. There are two basic streams associated with spatial representation that relate to uncertainty. One is aggregation and another is simplification. Aggregation has involved the intentional merging of spatial units, either to reduce computational effort or to accommodate the modeling approach. Work by Goodchild (1979), Daskin et al. (1989), Current and Schilling (1990), Fotheringham et al. (1995), Murray and Gottsegen (1997), Francis et al. (1999) and Francis et al. (2009) has sought to examine some aspect of unit aggregation in spatial optimization. In many cases it has been demonstrated that errors associated with such aggregation can be minimal in certain circumstances. Another way that aggregation has been used is to re-structure spatial units in a particular planning problem so that they conform to more easily implementable mathematical characterizations. For example, Murray (1999) demonstrates how aggregation of smaller management units into larger blocks then allows binary restrictions between blocks to be imposed in a constraint, avoiding the combinatorial complexity of enumerating possible blocks. Murray and Weintraub (2002) provide empirical evidence of error introduced as a result. The second aspect of spatial representation uncertainty is due to simplification. Miller (1996), Drezner and Drezner (1997) and Church (1999) detail that spatial simplification of demand/facility objects is fairly common, typically involving the conversion of an area to a single representative point. Evidence of error in such simplification along these lines is detailed in Murray and O’Kelly (2002).

Locational uncertainty noted in Table 2.1 is another important area of work in spatial optimization modeling, encompassing inaccuracy in given demand locations and sited facility locations. Demand location inaccuracy could be caused by systematic or random errors in data. Hodgson (1991) and Aerts et al. (2003) simulate demands with such kind of errors to examine its impacts on solutions. While Hodgson (1991) finds little sensitivity in model results given input data errors, Aerts et al. (2003) demonstrate substantial impacts of locational errors in practice. Demand locations could also be stochastic for some planning situations, either uncertain in location or changing over time (see Cooper 1974; Mirchandani and Odoni 1979; Weaver and Church 1983; AltInel et al. 2009; Bonneu and Thomas-Agnan 2009). Another form of locational uncertainty, solution variation, is that identified sites are not available due to land use and accessibility issues. Goodchild (1984) and Murray (2003) evaluate optimality loss by simulating possible locational offsets for sited facilities.

A final aspect of spatial uncertainty identified in Table 2.1 is associated with proximity interpretation. Distances between facilities/demands are commonly employed to measure proximity in spatial optimization. Many distance measures are possible, such as Euclidean, rectilinear and network. Uncertainty associated with possible distance measures could have significant impacts on modeling results. For example, Bach (1981) and AltInel et al. (2009) imply that model solutions and computational effort can be quite dissimilar when using different distance measures. Other related studies include Andersson et al. (1998) and

Plastria (2001). The impacts of distance measurement errors are discussed in Hodgson (1991) and Brimberg and Love (1995). Mirchandani and Oudjit (1980) and Berman and Odoni (1982) consider stochastic distances, simulating several scenarios to represent expected distances.

Since data uncertainty could have a substantial impact on spatial optimization model results, as suggested above, there is always a need to evaluate potential impacts. Previous work in this area has approached the assessment of uncertainty through simulation-based approaches, requiring repeated solution of many modified problems where data inputs have been systematically altered in some manner. What is lacking is the capacity to qualify or statistically infer something about the impacts of data uncertainty in definitive terms.

2.3 Uncertainty implications

The intent of this chapter is to demonstrate how spatial uncertainty can be addressed in an integrated and explicit fashion. A particular spatial optimization model, a dispersion model called the anti-covering location problem, is used here given it wide-spread application. It has been applied to a broad range of urban and environmental contexts, including forest planning (Barahona et al. 1992; Hochbaum and Pathria 1997; Murray 1999; Goycoolea et al. 2005), market saturation (Zeller et al. 1980), examination of habitat carrying capacity (Downs et al. 2008) and undesirable service provision (Grubesic and Murray 2008).

However, the implications of spatial uncertainty for this dispersion model are not understood.

The anti-covering location problem (ACLP), also referred to as node packing, vertex packing, maximal independent set and stable set problems, aims to determine the maximum number of service locations that can be sited while maintaining a minimum separation between locations. It has been formulated as integer program (see Padberg 1973; Nemhauser and Trotter 1975; Moon and Chaudhry 1984; Nemhauser and Sigismondi 1992; Murray and Church 1997; Murray and Kim 2008), with much attention on efficient solution. Consider the following notation:

$$i = index\ of\ areas$$

$$\Gamma = separation\ distance$$

$$\Phi_i = set\ of\ areas\ whose\ distances\ to\ area\ i\ are\ within\ \Gamma$$

$$x_i = \begin{cases} 1, & if\ area\ i\ is\ selected, \\ 0, & otherwise. \end{cases}$$

The variables, x_i, therefore represent the decisions associated with whether or not a service facility is sited at location i. All areas that are within the specified distance Γ from i, denoted by the set Φ_i, cannot simultaneously site a service facility if area i is already selected. With this notation, the formulation follows:

Anti-Covering Location Problem (ACLP)

$$\mathit{Maximize} \qquad \sum_i x_i \tag{2.1}$$

$$\mathit{Subject\ to} \qquad x_i + x_j \leq 1 \quad \forall i, j \in \Phi_i \tag{2.2}$$

$$x_i = \{0,1\} \quad \forall i \tag{2.3}$$

The objective of ACLP, equation (2. 1), is to maximize the number of selected areas to locate facilities/services. That is, the goal is to site as many as possible in order to provide the greatest level of service possible (or the most accessible). Constraints (2.2) are spatial proximity constraints ensuring that no two pairs of conflicting areas could both site facilities. Constraints (2.3) impose binary integer restrictions on decision variables.

The ACLP, or equivalently the node packing problem, is a deterministic optimization model that assumes the input data to be precise and accurate. In this case, the locations i and the distance between them are assumed to be known precisely. However, there actually exist various sources of uncertainty associated with any spatial data relied upon, and most certainly this should be understood and likely accounted for when used to support any substantive planning and analysis. For example, potential facility locations are typically inaccurate in some way. The reason is that data/map layers are usually produced from field survey, remote sensing imagery or paper-map reproduction, each of which involves systematic and random errors. Imprecision also occurs in any transformation process where raw data is somehow converted to a desirable data format, like digitization and geo-referencing. The metadata information for potential sites, such as commonly used parcels, has a specific spatial accuracy associated with it. For example, parcel data in Ramsey, Minnesota has a horizontal accuracy of 7 to 30 feet (RCSO 2005), while the accuracy of parcel data in Santa Barbara, California is 3.28 to 328 feet (SBSD 2011). Parcel level data is generally one of the most spatially accurate data sources, but even so a boundary of a parcel might be a few hundred feet or more off.

If such data is utilized in an optimization model, such as the ACLP, there is potential for error or bias in any analysis. In this case, a boundary being within 300 feet of the reported location would likely alter what is potentially conflicting. Murray and Grubesic (2011) discuss other sources of error and uncertainty, but suffice it to say that there is some error, either due to the data or other spatial considerations. Such error may be defined as a composite level of spatial uncertainty, ε, as follows:

$$\varepsilon = f\left(\varepsilon_b, \varepsilon_d, \varepsilon_p\right) \tag{2.4}$$

where,

$$\varepsilon_b = uncertainty\ associated\ with\ boundry\ delineation$$

$$\varepsilon_d = uncertainty\ associated\ with\ distance\ measurement$$

$$\varepsilon_p = uncertainty\ associated\ with\ proximity\ interpretation$$

Murray and Grubesic (2011) conclude that spatial location, distance measurement and proximity evaluation are most critical for the ACLP, but certainly other forms of error could be considered as well in equation (2.4). The challenge therefore is generating a valid inference based on the co-mingling of uncertainty. An approach for doing this is now detailed.

2.4 Accounting for error/uncertainty

A spatial optimization model is now derived to explicitly address locational error/uncertainty characterized by ε. This is in contrast to the generation of multiple input/output scenarios using simulated errors, an approach commonly relied upon in the literature. The developed model enables the possible impacts of uncertainty to be thoroughly evaluated, including worst case to best case scenarios. Given composite error in equation (2.4), the strict proximity restrictions in the ACLP become somewhat ambiguous. Specifically, some adjacency constraints may or may not become necessary, depending on the actual spatial proximity between two units i and j. For instance, if the pre-specified separation distance is 1,320 feet and the estimated error resulting from spatial uncertainty is 50 feet under some confidence level, the dispersion conditions, constraints (2.2), might need to impose a restriction when they are measured to be as close as 1,220 feet (1320-2*50=1220) or as far away as 1,420 feet (1320+2*50=1420). Alternatively, if the units are actually outside of the separation standard, then they should not be enforced as there is no problem if both are selected. Unfortunately, all proximity constraints in the ACLP are required to be imposed.

In order to address spatial error/uncertainty, consider the following additional notation:

$$p_{ij} = probability\ that\ area\ i\ is\ in\ conflict\ with\ j$$

$$y_{ij} = \begin{cases} 1, & if\ \ the\ proximity\ restriction\ \ between\ i\ and\ j\ is\ relaxed, \\ 0, & otherwise. \end{cases}$$

$$\Omega_i = set\ of\ areas\ whose\ distances\ to\ i\ are\ within\ (\Gamma - 2\varepsilon)$$

$$\Psi_i = set\ of\ areas\ whose\ distances\ to\ i\ are\ between\ (\Gamma - 2\varepsilon)\ and\ (\Gamma + 2\varepsilon)$$

The idea here is to track proximity restrictions based on what is certain and what is uncertain. Those restrictions that are uncertain will be accounted for in the model using decision variables y_{ij}. In this sense, the intent is to possibly relax uncertain restrictions, possibly impose uncertain restrictions, or possibly impose some but relax others. To accomplish this, a conflict probability, p_{ij}, is introduced, reflecting a potential violation of a dispersion requirement. This allows for differentiation between different types of restrictions based on ancillary criteria. As an example, when two units are considered further away, the conflict probability might be less than if they were closer to each other. Given the potential errors involved in ACLP, Ω_i represents a conservative conflict set that is certain based on the characterization of error. Restrictions between unit i and members of this set should always be imposed. Alternatively, the uncertain conflicts are represented by the set Ψ_i. It may or may not be necessary to impose restrictions between unit i and members of this set, depending on actual proximity. These two sets, therefore, account for the error scope ($\Gamma \pm 2\varepsilon$). Figure 2.1 depicts the two sets surrounding a residential parcel i.

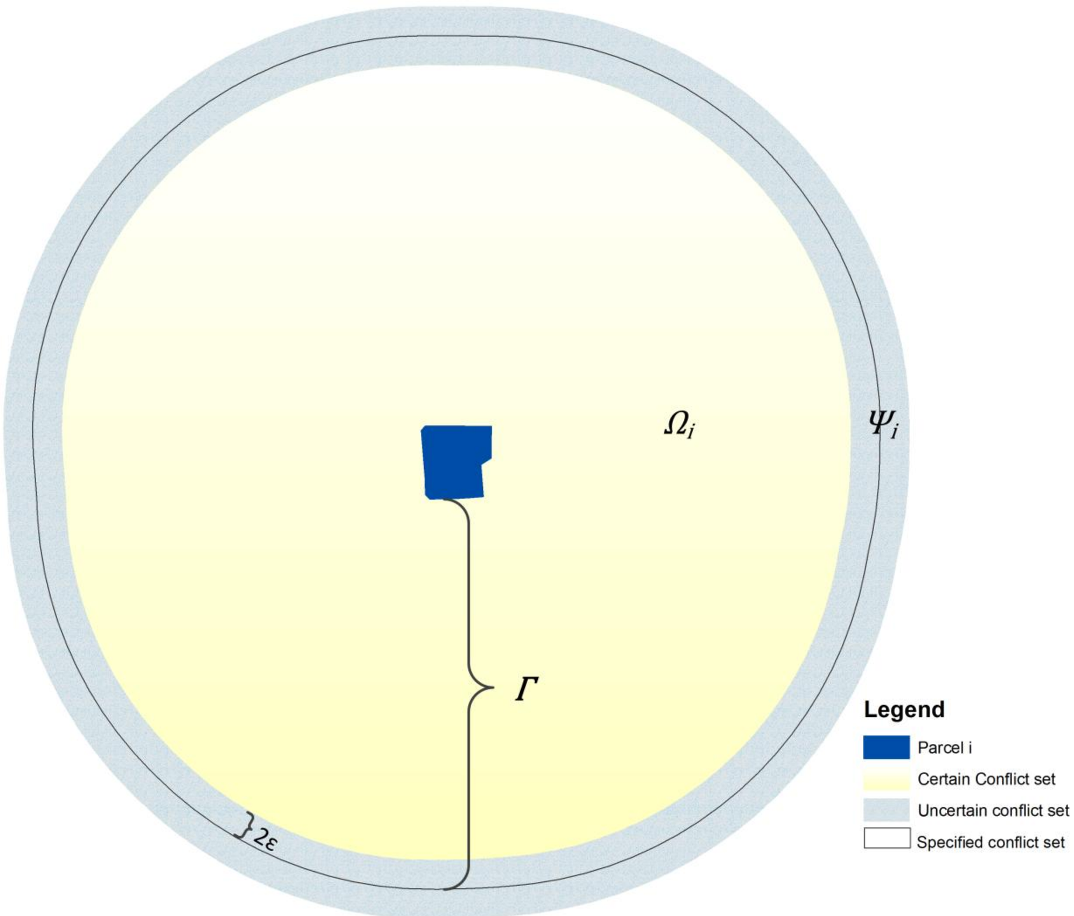


Figure 2.1 Sets Ω_i and Ψ_i

Structuring a new model that incorporates certainty and uncertainty will make use of the two sets, Ω_i and Ψ_i.

Error - Anti-Covering Location Problem (E-ACLP)

$$Maximize \qquad \sum_i x_i \tag{2.5}$$

$$Minimize \qquad \sum_i \sum_{j \in \Psi_i} p_{ij} y_{ij} \tag{2.6}$$

$$Subject\ to \qquad x_i + x_j \leq 1 \quad \forall i, j \in \Omega_i \tag{2.7}$$

$$x_i + x_j - y_{ij} \le 1 \quad \forall i, j \in \Psi_i \tag{2.8}$$

$$x_i = \{0,1\} \quad \forall i \tag{2.9}$$

$$y_{ij} = \{0,1\} \quad \forall i, j \in \Psi_i$$

This new model is structured with multiple objectives. The first objective, (2.5), is to maximize the number of selected areas, exactly as structured in the ACLP. The second objective, (2.6), is to minimize the total conflict probabilities of relaxing separation constraints that are spatially uncertain. Constraints (2.7) ensure that no two selected sites conflict among the certain restrictions. Constraints (2.8) track separation of those that might be in conflict. These are the constraints that could be okay to relax or ignore, depending on error. When y_{ij} equals one, both x_i and x_j could be one in constraint (2.8), indicating both areas could be selected; otherwise, the proximity constraint is imposed, and at most one of them can be selected. Constraints (2.9) impose binary integer restrictions on decision variables.

The E-ACLP is a multi-objective extension of the ACLP. Further, it is related to the work of Hochbaum and Pathria (1997) who proposed the generalized independent set problem, where all proximity constraints can be violated with some penalty cost. Thus, if the set Ω_i is empty and Ψ_i is not empty, then the E-ACLP would be equivalent to the generalized independent set problem. Thus, the E-ACLP can be considered as an extension or generalization of the generalized independent set problem in that it imposes restrictions between conflicts that are certain, but allows those that are uncertain to possibly be relaxed, similar to the spirit of the generalized independent set problem.

Solution of the E-ACLP is challenging given the multiple objectives and binary integer variables. Depending on the significance of uncertainty and the penalty costs, there will no doubt be tradeoffs between the two objectives. The model can be solved to identify trade off solutions using multi-objective techniques, such as the weighting or constraint methods (see Cohen 1978).

2.5 Application

The ACLP and the E-ACLP detailed above are applied in the context of evaluating the implications of offender residency restriction laws. In order to mitigate risk exposure to communities by convicted sex offenders that reintegrate into society, local governments have considered/proposed regulations that limit where convicted sex offenders can live relative to other offenders. Grubesic and Murray (2008) detailed how the ACLP could be used to assess this issue, providing insights regarding the maximum number of offenders that could reside in a region under such conditions as well as the geographic implications of this residence restriction law. The study area is a community in the Phoenix metropolitan area containing 1,583 parcels of which 1,295 are residential parcels. The separation distance Γ has been established as 1,320 feet, identical to proposed legislation detailed in Grubesic and Murray (2008). Based on parcel data accuracy and other sources of spatial uncertainty, as well as imputed quality, ε is estimated to be ± 50 feet.

The spatial optimization models are structured using Python, an open source object-oriented programming language, and subsequently solved using a commercial optimization package, Gurobi. Computational processing was carried out on a MS Windows-based, Intel Xeon (2.53 GHz) computer with 6 GB of RAM.

Application of the ACLP indicates that the maximum number of offenders that could reside in this community would be 16 in order to ensure a dispersion distance 1,320 feet between each pair of offenders. This possible spatial configuration is shown in Figure 2.2. The analysis using the ACLP, as noted previously, assumes that the data is free of error or uncertainty. The issue then is what are the implications of data error/uncertainty in this case, both at a regional level in terms of total number of offenders as well as more locally in terms of spatial patterns of residency.

To assess the impacts of data error/uncertainty, the E-ACLP is applied. For the sake of simplicity, the conflict probabilities are assumed to all equal one. The constraint method is used here to identify all trade off solutions associated with these two objectives.

Figure 2.2 Offender distribution identified using the ACLP

Spatial configuration of the 15 residences identified using the E-ACLP

The computational results for each of the five different tradeoff solutions are displayed in Table 2.2. The “maximal residences” column in Table 2.2 indicates the number of offenders that can reside in the region, specified in objective (2.5). The “minimal conflict probability” column in Table 2.2 effectively corresponds to the number of restriction constraints relaxed since all the conflict probabilities are equal. This is objective (2.6) in the E-ACLP. Table 2.2 highlights that accounting for error in the model means that as few as 15 residences could be established, but up to 19 are possible. This is in contrast to the 16 identified using the ACLP, assuming certainty in spatial location. Table 2.2 shows that increasing the number of residences means that the conflict probability would increase, or rather separation constraints would need to be relaxed. This trade off is depicted in Figure 2.3, showing the Pareto optimal curve and the associated non-dominated solutions attained by changing the relative importance of objective (2.5) with respect to objective (2.6). For the 15 residences in Figure 2.3, all uncertain restrictions are actually imposed, explaining the zero penalty cost in Table 2.2 for this solution. On the other extreme, it is possible that 19 residences could be established, and doing so would mean that 22 uncertain restrictions would not be imposed. The other solutions in Figure 2.3 (and Table 2.2) therefore reflect a trade off ranging between these extremes of 15 to 19 for number of residences and 0 to 22 for conflict probability.

Table 2.2: Computational results for the E-ACLP

Maximal residences (Objective 5)	Minimal penalty cost (Objective 6)	Solution Time
15	0	147.05
16	2	19,287.59
17	7	825.57
18	12	2,878.05
19	22	1,223.61

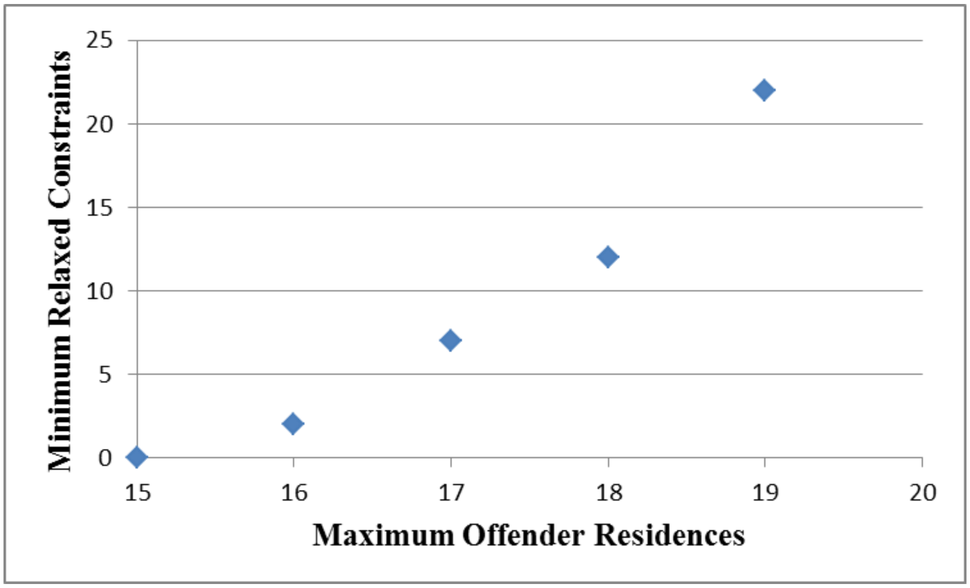


Figure 2.3 Non-dominated solutions found using the E-ACLP

Figure 2.4 Spatial configuration of the 15 residences identified using the E-ACLP

Beyond the implications for total number of residences possible, each possible solution reflects a different spatial pattern that could result. In the case of 15, this corresponds to the situation that the most expansive separation distance $\Gamma + 2\varepsilon$ (1320+100 feet) is effectively imposed. The resulting spatial configuration is shown in Figure 2.4. All of the other solutions reflect a relaxation of this strict interpretation of error. Thus, Figure 2.5 corresponding to 16, Figure 2.6 corresponding to 17, Figure 2.7 corresponding to 18 and Figure 2.8 corresponding to 19 each move more towards the other end of separation, $\Gamma - 2\varepsilon$ (1320-100 feet).

Figure 2.5 Spatial configuration of the 16 residences identified using the E-ACLP

Figure 2.6 Spatial configuration of the 17 residences identified using the E-ACLP

Figure 2.7 Spatial configuration of the 18 residences identified using the E-ACLP

Figure 2.8 Spatial configuration of the 19 residences identified using the E-ACLP

It is most evident that the E-ACLP provides the capacity to identify a range of implications for spatial uncertainty in the geographic position of land units. However, the specific comparative nuances are particularly interesting. The contrast between the ACLP and E-ACLP is very evident when examining Figure 2.2 (ACLP) and Figure 2.5 (E-ACLP), where 16 residences are identified by each approach. What can be observed is a change in the spatial distributions of identified residences. The reason for this is that the separations between selected residences in Figure 2.2 are mostly less than 1,420 feet. When error is taken into account, this changes things, which means that proximity restrictions must be relaxed in the E-ACLP if 16 residences are to be selected.

2.6 Discussion

Addressing uncertainty in spatial data no doubt complicates any analysis endeavor. Further discussion on a number of points is provided here associated with analysis complexity, model solution and alternative approaches.

The application results presented in the previous section illustrate that the ACLP is in fact sensitive to spatial data uncertainty. If the potential errors in spatial data are ignored, the analysis and any conclusions are incomplete. The Error-Anti-Covering Location Problem (E-ACLP) was introduced to illuminate how spatial uncertainty would affect analysis findings. In this case, it is possible that 15 to 19 offenders could reside in the region, and in each case there would be a different associated level of neighborhood danger or risk as a result. Thus, there is considerable complexity in the analysis of associated impacts of uncertainty. Addressing spatial error/uncertainty using the E-ACLP has necessitated the use of multiple objectives. Multi-objective models mean that there are trade-off, reflecting the complexities of planning situation and requiring subjective interpretation of the results. The application results suggest a range of possible impacts associated with data error/uncertainty on the modeling results; moreover, the trade-off curve provides a statistical confidence that can be related to data uncertainty. This rationale may also be applied to other spatial optimization models.

A final point in need of discussion is model solution. Because the uncertainty associated with spatial data impacts dispersion requirements, the E-ACLP allows for the relaxation of uncertain proximity constraints with some conflict probability. At a basic level, this means that a difficult problem, the ACLP or node packing problem, has been extended through the use of an additional objective and new binary integer variables. This point is particularly important when one considers that the ACLP/node packing problem is itself a challenging optimization problem to solve, and subject to considerable research on better solution approaches. In particular, Murray and Church (1997), Goycoolea et al. (2005) and Murray and Kim (2008), among others, have focused on methods for identifying better facets or alternative formulations of the problem. Given that the application reported in this chapter considered 1,295 potential residential locations and required less than 2 seconds to solve the ACLP, it is not terribly surprising to see in Table 2.2 that the E-ACLP was considerably more difficult to solve, requiring 19,287 seconds in the worst case. As problem size grows, it is unlikely that the E-ACLP would be optimally solved using commercial software. Therefore, it is clear that future research is needed for developing alternative approaches, both heuristic and exact, to solve the E-ACLP.

2.7 Conclusion

In this chapter we have highlighted the error and uncertainty common in virtually all spatial information. The implications of error/uncertainty in spatial data are that any methods making use of the data must understand their effects. Spatial optimization work has approached geographic uncertainty through the use of simulated error propagation, but such an approach is computationally intensive and fails to provide capabilities for establishing statistical significance in derived findings. To address this deficiency, we have proposed a new integrated approach to explicitly account for data uncertainty in a spatial optimization model. This entailed the formulation and solution of a new multi-objective model. The application results illustrated the effectiveness of this approach, providing the capacity to establish bounds on spatial uncertainty. It is hoped that more research will follow along these lines as addressing uncertainly will no doubt mean that more complex and difficult models will arise.

CHAPTER 3

SPATIAL UNCERTAINTY IN HARVEST SCHEDULING*

Chapter 2 addressed spatial data uncertainty issues in a dispersion model where geographic proximity is measured by Euclidean distance. However, adjacency in many planning problems is evaluated by contiguity, such as sharing a common boundary or vertex. This chapter deals with spatial uncertainty in the context of harvest scheduling, where adjacency between harvest units is determined by sharing a common boundary. An algorithm to assess uncertainty in contiguity-based adjacency is developed and an alternative modeling approach is proposed. Comparison to the model presented in Chapter 2 is also undertaken.

3.1 Introduction

Harvest scheduling involves important resource management decisions, with significant implications for economic and environmental well being. The diverse and competing uses of forest resources, such as economic productivity, recreation and flora and fauna sustainability, make the scheduling of harvest units challenging. To assist in this difficult task, optimization models have been widely relied upon to develop harvest schedules, where constraints are structured and imposed to limit spatial disturbance with an objective to maximize harvesting

---

* This chapter represents a slightly revised version of a paper published in *Annals of Operations Research*, co-authored with Dr. Alan T. Murray.

benefits (Thompson et al. 1973; Murray and Weintraub 2002; Goycoolea et al. 2005; Constantino et al. 2008).

Murray (1999) discusses that there are two general approaches to represent and impose spatial disturbance restrictions in harvest optimization models. One assumes that the timber units are delineated so that harvesting any two adjacent units would exceed a maximum area of disturbance. The other approach anticipates that the harvest units are much smaller than a stipulated maximal clear-cut area, making it possible to harvest several neighboring units simultaneously. Murray (1999) refers to these approaches as the unit restriction model (URM) and the area restriction model (ARM), respectively. The URM and ARM are deterministic and assume model input to be precise and accurate. Unfortunately spatial information is typically uncertain in many ways, particularly spatial location and harvest unit boundaries. De Groeve and Lowell (2001) highlight the significance of this issue, showing that the width around forest unit boundaries within which the actual location could reside ranges from 24.7 to 44.4 meters. This is not surprising considering that management units have historically been delineated by using automated paper map conversion approaches and aerial photographs, but even GPS and other survey based data are limited in spatial accuracy (Edwards and Lowell 1996; Brown 1998; De Groeve and Lowell 2001; Radoux and Defourny 2007).

It is well recognized that deterministic optimization models could be sensitive to spatial uncertainty (Cooper 1974; Drezner and Drezner 1997; Murray and Weintraub 2002; Murray 2003; Aerts et al. 2003; Altinel et al. 2009). Unfortunately, there is no universal understanding of spatial uncertainty impacts or biases. This means that evaluation must be designed for individual models before employing them to assist decision-making processes (Church 1999; Murray 2003). Even though a substantial amount of work has investigated uncertainty issues in harvest scheduling (Hoganson and Rose 1987; Weintraub and Abramovich 1995; Klenner et al. 2000; Boyland et al. 2005; Peter and Nelson 2005; Palma and Nelson 2009), spatial uncertainty has not been explicitly examined in harvest scheduling optimization models.

This chapter details new approaches to address spatial uncertainty in harvest scheduling subject to spatial disturbance restrictions. We structure two multi-objective approaches to assess how spatial uncertainty could impact harvest schedules, focusing on the URM. The next section reviews previous research related to data uncertainty in forest planning. This is followed by the formulation of the URM and a discussion of spatial uncertainty associated with its application. This chapter then structures two new modeling approaches that explicitly account for spatial data uncertainty. Application results are presented that illustrate the range of potential solutions possible when spatial uncertainty is considered. Finally, a discussion and concluding comments are given.

3.2 Background

There are many facets of uncertainty in forest management planning. Management unit delineation, yield, market value, production demands, sustainability requirements and other inputs for harvest scheduling optimization models may be highly uncertain. Failure to account for data uncertainty in models could lead to suboptimal, infeasible or biased solutions (Pickens and Dress 1988). A number of studies have examined aspects of data uncertainty in harvest scheduling models.

One widely employed approach is simulating possible scenarios and solving the resulting deterministic scheduling model for each scenario. For example, natural disturbance scenarios, like fires or avalanches, have been simulated to evaluate the disturbance impacts on harvest scheduling in Klenner et al. (2000), Von Gadow (2000), and Peter and Nelson (2005), where average harvest profit could increase $1.8 million per year when fire disturbance is incorporated. Scenarios with uncertain inventory data are generated in Pkukkala (1998) and Eid (2000), while uncertainties in timber yield are simulated in Hoganson and Rose (1987) and Eriksson (2006). Eid (2000) also illustrate that an error level of 15 % in inventory data could result in expected net present value losses between 64 NOK per ha and 1471 NOK per ha. Boyland et al. (2005) simulate the deviations of harvest schedules to assess robustness. Scenario-based approaches are straightforward and have been applied in various forest planning contexts.

However, the identification of scenarios is necessary, which might be difficult and likely requires substantial additional data. Secondly, since enumeration of all possible scenarios is not feasible due to computational practicalities, sampling becomes necessary, leaving much remaining uncertainty in any derived findings (Snyder 2006).

Stochastic programming methods have also been applied to harvest scheduling. As an example, Boychuk et al. (1996) construct a stochastic programming model to account for the impacts of fire loss. Reeves and Haight (2000) incorporate the means and covariances of stumpage prices into a harvest scheduling model. Chance-constrained programming has been used to assess uncertainty in timber yields (Weintraub and Vera 1991; Weintraub and Abramovich 1995; Hof et al. 1996) and production demands (Hof and Pickens 1991). Wind damage and spatial structure of stands are taken into account by employing probabilistic models (Meilby 2001) and Markov decision process (Forsell et al 2011).They also demonstrate that the expected net present value of stands would significantly increase if accounting for the risk of wind damage (Forsell et al. 2011). A limitation is the assumption that the probability distributions of uncertain parameters are known. Moreover, the computational burden of these approaches requires the development of problem specific exact or heuristic solution techniques, limiting their general application. Robust optimization models have also been structured to integrate data uncertainty in harvest scheduling (Palma and Nelson 2009; Bohle et al. 2010).

There is little doubt that the evaluation of data uncertainty impacts in harvest scheduling is essential. Previous forest research has investigated various uncertainties, but spatial uncertainty has not been explored relative to adjacency based restrictions, leaving important questions about expected impacts and potential biases that may exist in derived harvesting plans.

3.3 URM and spatial uncertainty

The URM has been widely applied to support harvest scheduling (see Thompson et al. 1973; Murray 1999; Murray and Weintraub 2002; Goycoolea et al. 2005). The primary feature of the URM is constraints that prohibit any two neighboring management units from being simultaneously harvested. Beyond this, numerous extensions are possible, including the addition of volume flows, green-up requirements, road building/maintenance, etc. Considerable research has focused on developing efficient exact and heuristic solution techniques for the URM (Murray and Church 1995; Murray and Church1996; Snyder and Revelle 1997; Weintraub et al. 2000). Consider the following notation:

$i = index\ of\ planning\ units$

$\beta_i = benefit\ associated\ with\ harvesting\ unit\ i$

$\Phi_i = set\ of\ planning\ units\ adjacent\ to\ unit\ i$

$$x_i = \begin{cases} 1, & if\ unit\ i\ is\ harvested, \\ 0, & otherwise. \end{cases}$$

Without loss of generality, only a single planning period is considered. The URM formulation is:

Unit Restriction Model (URM)

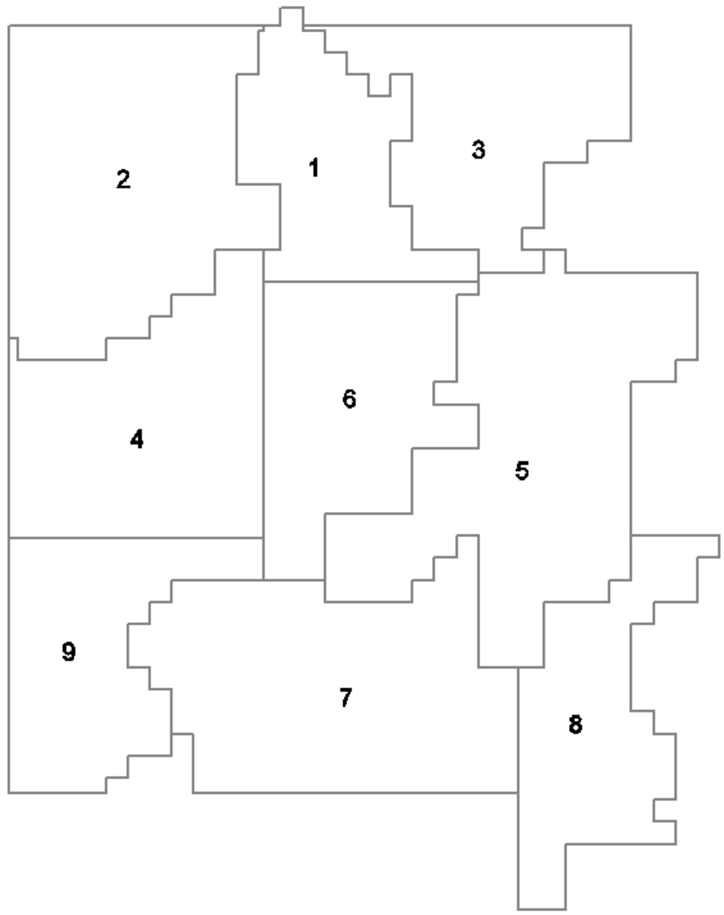
$$Maximize \qquad \sum_i \beta_i x_i \tag{3.1}$$

$$Subject\ to \qquad x_i + x_j \leq 1 \quad \forall i, j \in \Phi_i \tag{3.2}$$

$$x_i = \{0,1\} \quad \forall i \tag{3.3}$$

The objective of URM, (3.1), is to maximize the benefits associated with harvest scheduling, which could be economic or environmental benefits. Constraints (3.2) ensure that no two adjacent units are simultaneously harvested. Constraints (3.3) impose binary restrictions on decision variables.

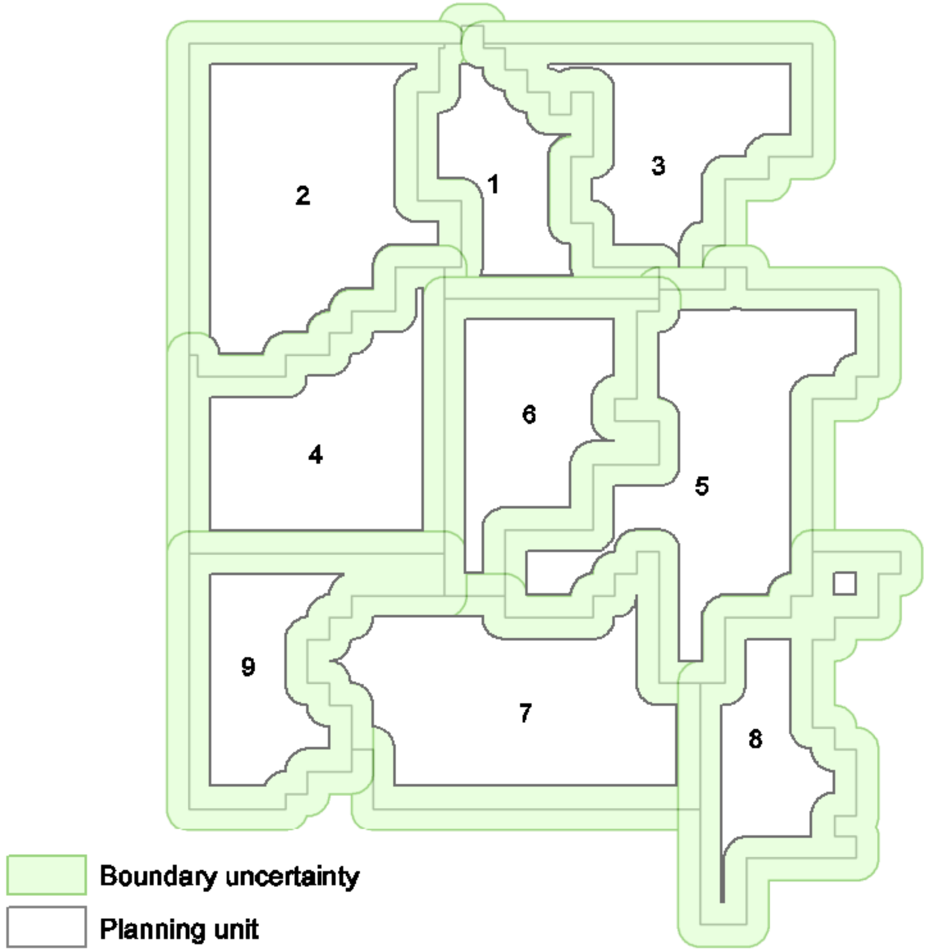
Again, various extensions to this basic formulation are possible. Common concerns include road building/maintenance, as well as even flows in the case of temporal models. In addition, much effort has focused on exact and heuristic solution techniques (Murray and Church 1995; Murray and Church 1996; Snyder and Revelle 1997; Weintraub et al. 2000). An assumption in the URM is that adjacency relationships are certain. That is, the input adjacency set Φ_i is assumed to be precise and accurately defined. However, there actually exist various sources of uncertainty associated with spatial information and its use in data processing and manipulation. Given the potential for erroneous or biased results, any uncertainty should be understood and accounted for if possible when data is used to support any substantive planning and analysis. For example, boundary uncertainty of forest units has been discussed in Edwards and Lowell (1996), Naesset (1998), De Groeve and Lowell (2001) and Orzanco et al. (2004), and concludes that forest unit boundaries are inaccurate in various ways. Figure 3.1 demonstrates that boundary imprecision could have a significant influence on identified components of adjacency sets. Depicted in Figure 3.1a are planning units assumed to spatially be precise and accurate. Figure 3.1b shows the geographic extent of where the actual unit boundary could be when taking spatial uncertainty into account. Specifically, units 6 and 9 are considered neighbors in Figure 3.1a but may not actually be neighbors if accounting for boundary uncertainty (Figure 3.1b). Alternatively, unit pairs 6 and 3 and 6 and 2 are not considered adjacent but could actually be adjacent if accounting for boundary uncertainty (Figure 3.1b). In addition to boundary delineation, uncertainty could also arise in adjacency interpretation. For example, various adjacency interpretations are considered in the literature, such as weak adjacency in Goycoolea et al. (2005) and strong adjacency in Constantino et al. (2008). Other forms of uncertainty could be considered as well (see Murray and Grubesic 2011). The challenge now is evaluating the impacts of spatial uncertainty in the URM given such uncertainty.



(a)

Figure 3.1 Harvest unit boundary imprecision: (a) planning units assumed to spatially be precise and accurate, (b) planning units accounting for boundary uncertainty

Figure 3.1 continued



(b)

3.4 Accounting for spatial uncertainty

One approach for addressing spatial uncertainty in the URM is to reconsider neighbor, or adjacency, definitions. The adjacency sets, Φ_i, are not necessarily accurate when spatial uncertainty is recognized. One way to deal with this is to define two sets, Ω_i and Ψ_i, representing the certain adjacency and possible adjacency conditions for unit i, respectively. If ε represents the spatial uncertainty of a polygon boundary, then the buffered area around units in Figure 3.1b is a realization of this uncertainty. The two sets can therefore be identified in a systematic fashion. An algorithm is detailed in Figure 3.2 to facilitate spatial relationship characterization and the relative probabilities that two units are adjacent.

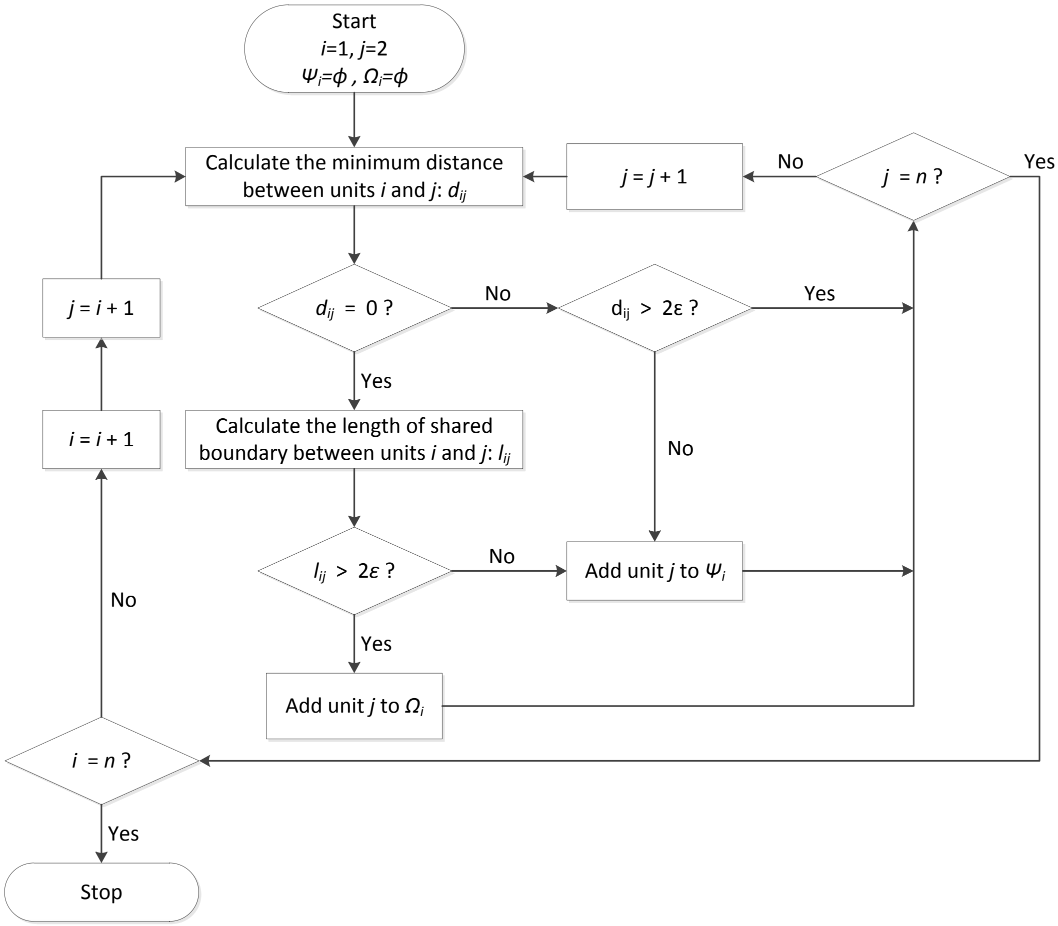


Figure 3.2 Ω_i and Ψ_i generation

If the minimum boundary distance between units i and j in the URM, d_{ij}, is larger than 0 and less than or equal to 2ε, two units may be neighbors. Further, if the length of shared boundary between units i and j, l_{ij}, is less than or equal to 2ε, then there is a chance that the units may not actually be neighbors. This suggests that some adjacency relationships, Ω_i, are known with certainty, but other potential adjacencies, Ψ_i, are not. The challenge is to explicitly incorporate known and potential adjacency into a URM based model in a manner where uncertainty can be explored and better understood. To accomplish this, one approach is to use decision variables y_{ij} to track whether a potential adjacency condition is relaxed, with a parameter p_{ij} reflecting a probabilistic significance of relaxation. The minimum distance and shared boundary length discussed previously could be employed to specify p_{ij}. For example, shorter minimum distance and longer shared boundary between two units suggest a higher probability of being adjacent. Considering Figure 3.1b, units 3 and 6 are more likely to be neighbors than units 2 and 6 because of the minimum boundary distance between units 3 and 6 is shorter than that of units 2 and 6. An approach based on certain and possible adjacency is structured as follows:

Error-Unit Restriction Model I (E-URM I)

$$Maximize \qquad \sum_i \beta_i x_i \tag{3.4}$$

$$Minimize \qquad \sum_i \sum_{j \in \Psi_i} p_{ij} y_{ij} \tag{3.5}$$

$$Subject\ to \qquad x_i + x_j \leq 1 \quad \forall i, j \in \Omega_i \tag{3.6}$$

$$x_i + x_j - y_{ij} \leq 1 \quad \forall i, j \in \Psi_i \tag{3.7}$$

$$x_i = \{0,1\} \quad \forall i \tag{3.8}$$

$$y_{ij} = \{0,1\} \quad \forall i, j \in \Psi_i$$

This model is structured with multiple objectives. The first objective, (3.4), is to maximize the benefits associated with harvesting units, as structured in the URM. The second objective, (3.5), minimizes the total risk of violating proximity constraints that are spatially uncertain. Constraints (3.6) ensure that no two harvest units conflict among the known restrictions. Constraints (3.7) track separation between those units that are potentially in conflict. These are the constraints that could be okay to relax or ignore, depending on uncertainty. When y_{ij} equals one, both x_i and x_j could be one in constraint (3.7), indicating both units could be harvested; otherwise, the proximity constraint is imposed, and at most one unit could be harvested. Constraints (3.8) impose binary restrictions on decision variables.

Of course, the E-URM I is not the only possible approach for addressing uncertainty. In the E-URM I each pair of potential adjacency restrictions, Ψ_i, may be relaxed. However, it might make sense to consider relaxing all uncertain restrictions associated with unit i, with no violation risk beyond the initial relaxation. That is, relaxations associated with unit i are allowed, or they are not. Consider the following additional notation:

$$\tilde{p}_i = probability\ of\ violating\ potential\ restrictions\ for\ unit\ i$$

$$z_i = \begin{cases} 1, & if\ unit\ i\ is\ relaxed, \\ 0, & otherwise. \end{cases}$$

The risk, $\tilde{p}_i$, is not related to each constraint relaxation but rather to each potential harvest unit. Since relaxing unit i indicates that all potential adjacency constraints associated with unit i could be relaxed. Specification of this potential unit risk, $\tilde{p}_i$, might be:

$$\tilde{p}_i = \sum_j p_{ij} \tag{3.9}$$

New decision variables, z_i, are utilized to determine whether the unit constraints are to be relaxed. An alternative model is therefore structured as follows:

Error-Unit Restriction Model II (E-URM II)

$$Maximize \qquad \sum_i \beta_i x_i \tag{3.10}$$

$$Minimize \qquad \sum_i \tilde{p}_i z_i \tag{3.11}$$

$$Subject\ to \qquad x_i + x_j \le 1 \quad \forall i, j \in \Omega_i \tag{3.12}$$

$$x_i + x_j - z_i - z_j \le 1 \quad \forall i, j \in \Psi_i \tag{3.13}$$

$$x_i = \{0,1\} \quad \forall i \tag{3.14}$$

$$z_i = \{0,1\} \quad \forall i$$

The first objective, (3.10), remains to maximize the total benefits of harvesting and the second objective, (3.11), is to minimize the total risk of violating potential harvest unit restrictions. Constraints (3.12) ensure the certain separation requirements are satisfied for each unit. Constraints (3.13) link the relaxation decision of unit i to all of the units in the set Ψ_i. When z_i equals one, all units in Ψ_i could be harvested concurrently. Constraints (3.13) impose binary integer restrictions on decision variables.

In the E-URM I each potential adjacency restriction pair is associated with a violation risk, independent of other uncertain adjacencies. Alternatively, E-URM II focuses on a unit and the entire set of potential adjacency conditions. However, the violation risk is associated with units and potential restrictions for one unit are bundled together in the E-URM II. The differences between the E-URM I and E-URM II are subtle. Consider unit 6 in Figure 3.1, with $\Psi_6 = \{2,3,9\}$. Constraints (3.7) for Ψ_6 in the E-URM I would be:

$$x_6 + x_2 - y_{62} \leq 1$$

$$x_6 + x_3 - y_{63} \leq 1$$

$$x_6 + x_9 - y_{69} \leq 1$$

Selecting both x variables in any one of these constraints would imply a relaxation, so the associated y variables would need to be one. Specifically, suppose $x_6 = x_2 = x_3 = 1$. This could imply that $y_{62} = y_{63} = 1$. The objective, (3.5), would therefore reflect that multiple pairs are relaxed. Consider this relationship in the E-URM II. Constraints (3.13) would be:

$$x_6 + x_2 - z_6 - z_2 \leq 1$$

$$x_6 + x_3 - z_6 - z_3 \leq 1$$

$$x_6 + x_9 - z_6 - z_9 \leq 1$$

Selecting both x variables in any one of these constraints would only require at least one associated z variables to be one. Specifically, suppose $x_6 = x_2 = x_3 = 1$. This could imply that only $z_6 = 1$ is sufficient to maintain feasibility. This means that the relaxation of potential adjacency is accounted for in objective (3.11) for a single unit, in contrast to multiple conditions above for the E-URM I. Research has shown the URM to be challenging to solve (Murray and Church 1996; Hochbaum and Pathria 1997; Weintraub et al. 2000; Constantino et al. 2008). Thus, extensions of the URM, such as the multi-objective E-URM I and II, are also expected to be computationally demanding. Given the two objectives, the weighting or constraint methods could be used to identify tradeoff solutions (see Cohen 1978), providing insight on the impacts of spatial uncertainty.

3.5 Application results

A forest region located in northern California is used to examine the impacts of spatial uncertainty on harvest scheduling (see Figure 3.3). A number of other studies have undertaken various types of spatio-temporal analysis for this region, including Murray and Weintraub (2002), Murray et al. (2004) and Goycoolea et al. (2005). This region has 351 harvest units averaging 25 acres in size. Investigation suggests that harvest units were likely delineated using an automated or semi-automated map conversion approach. Our analysis estimates ε to be 30 meters. That is, harvest unit boundaries are only accurate to within 30 meters of their indicated location.

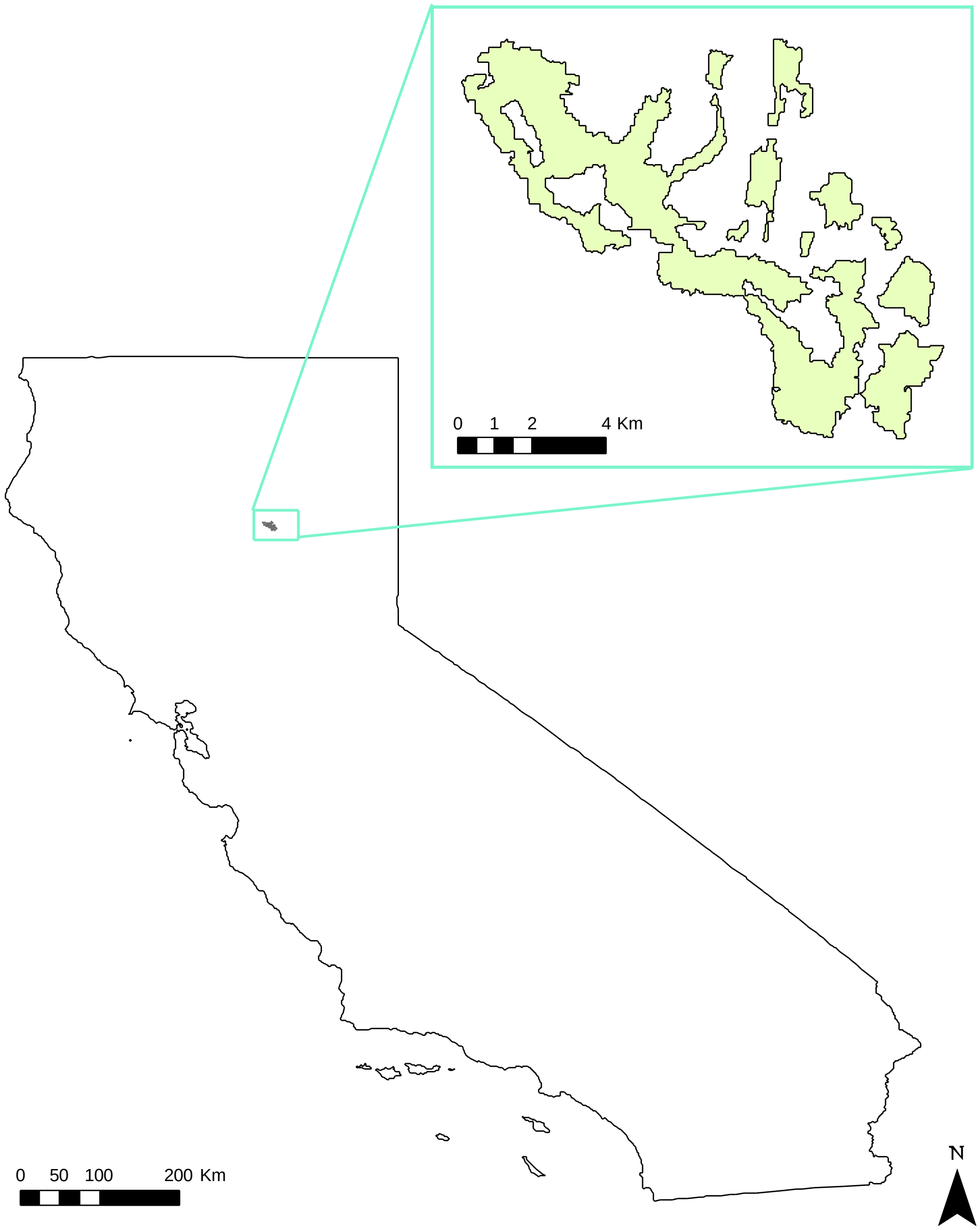
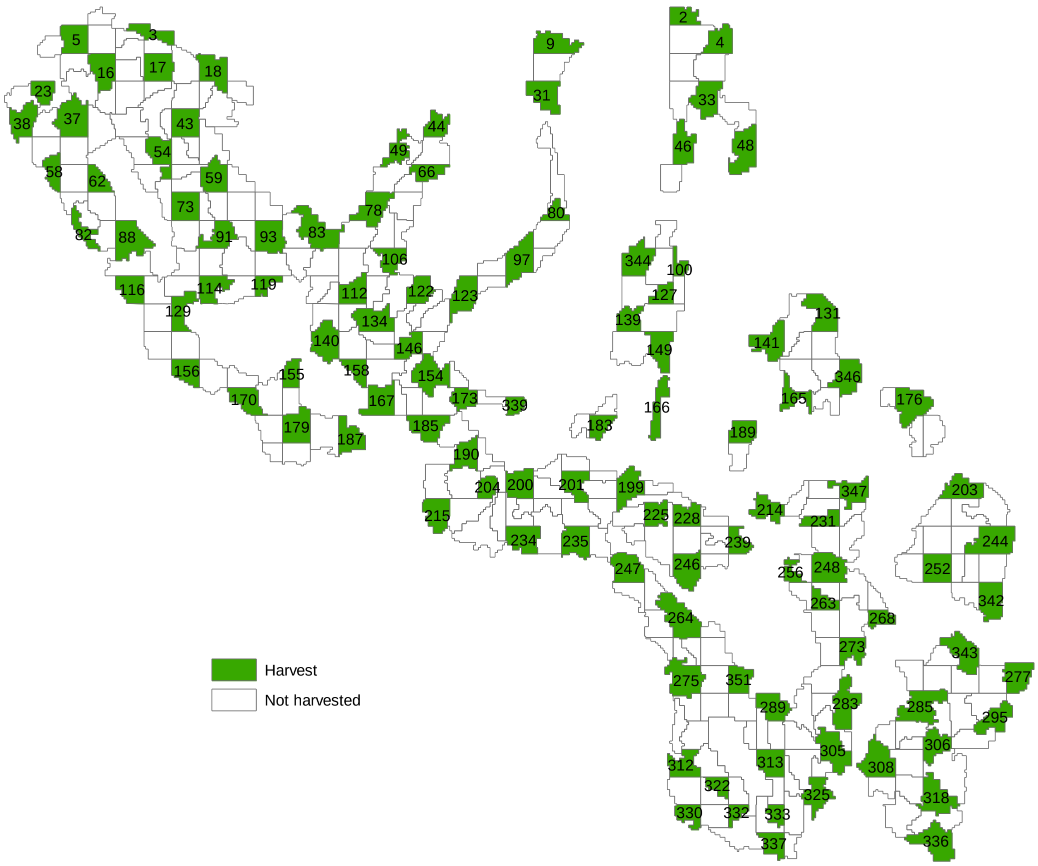


Figure 3.3 Butter Creek forest in Northern California

The spatial optimization models are structured using Python, and subsequently solved using a commercial optimization package, Gurobi. Computational processing was carried out on a MS Windows-based, Intel Xeon (2.53 GHz) computer with 6 GB of RAM.

Application of the URM indicates that the maximum return possible from harvesting activities is 5854.25 (114 units harvested). The spatial configuration of harvesting activity is shown in Figure 3.4a. This analysis result, as noted previously, assumes that the data is certain and spatially precise. However, spatial boundaries of units are only accurate to within 30 meters. There is a need to address this spatial uncertainty.

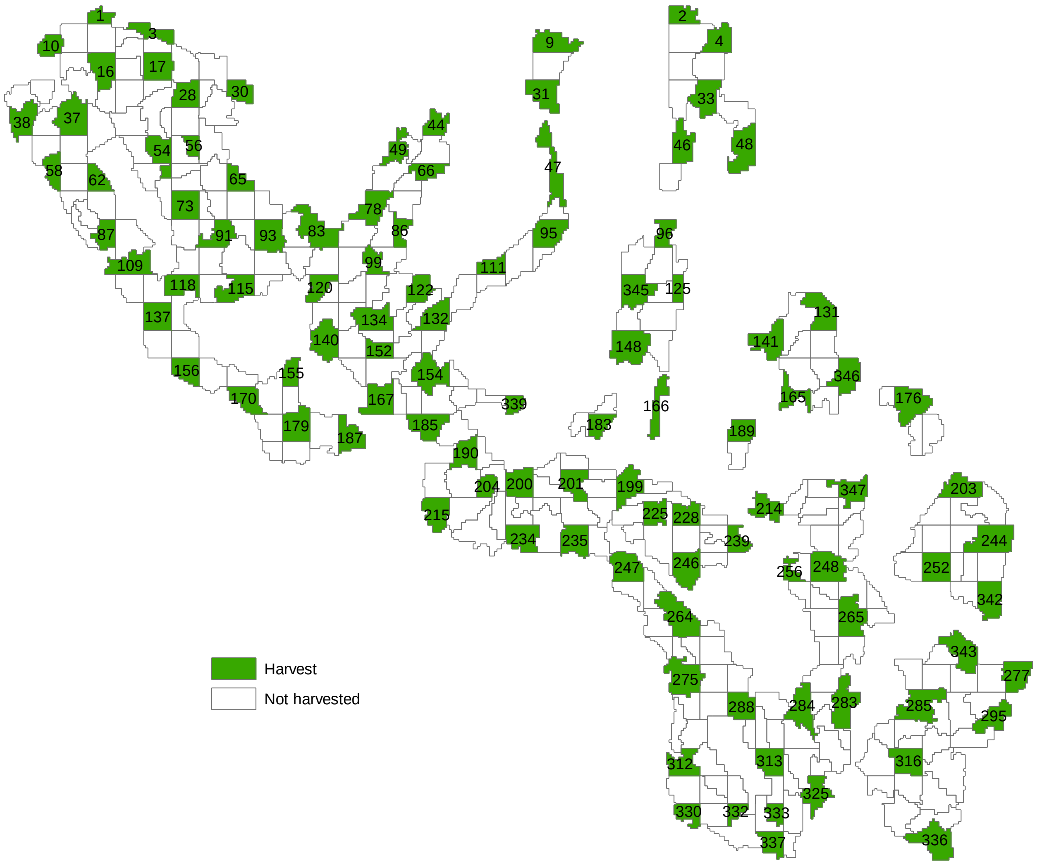
Both E-URM I and E-URM II are utilized to assess the impacts of data uncertainty. The risk of violating potential adjacency restrictions, p_{ij}, are set to range from 1 to 8 depending on the actual minimum boundary distance, d_{ij} and shared boundary length, l_{ij}. Specifically, when $\frac{3\varepsilon}{2} < d_{ij} \leq 2\varepsilon$, $p_{ij} = 1$; when $\varepsilon < d_{ij} \leq \frac{3\varepsilon}{2}$, $p_{ij} = 2$; when $\frac{\varepsilon}{2} < d_{ij} \leq \varepsilon$, $p_{ij} = 3$; when $0 < d_{ij} \leq \frac{\varepsilon}{2}$, $p_{ij} = 4$; when $0 \leq l_{ij} < \frac{\varepsilon}{2}$, $p_{ij} = 5$; when $\frac{\varepsilon}{2} \leq l_{ij} < \varepsilon$, $p_{ij} = 6$; when $\varepsilon \leq l_{ij} < \frac{3\varepsilon}{2}$, $p_{ij} = 7$; when $\frac{3\varepsilon}{2} \leq l_{ij} \leq 2\varepsilon$, $p_{ij} = 8$. The models all require less than 0.1 second to solve using Gurobi and the constraint method is used here to identify all tradeoff solutions associated with the two objectives.



(a)

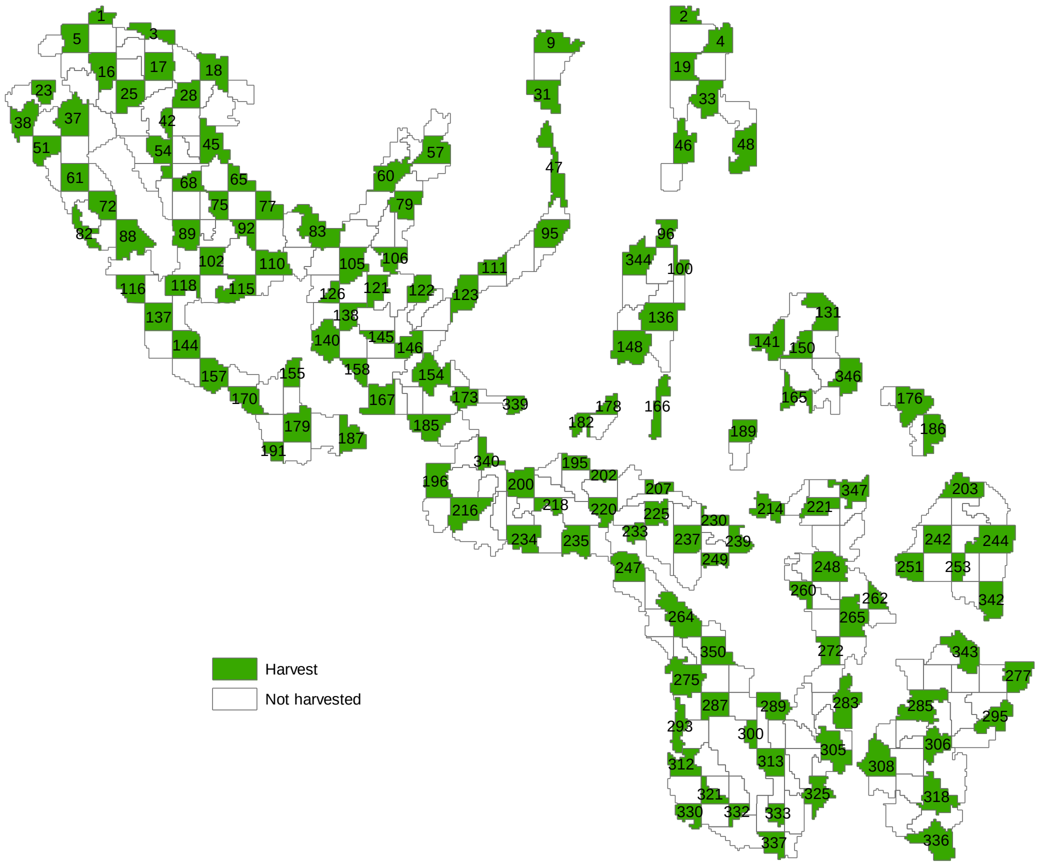
Figure 3.4 Spatial configuration comparison between the URM and E-URM I: (a) URM (5854,25 benefit), (b) E-URM I (5571.62 benefit and 0 violation risk), (c) E-URM I (6876.95 benefit and 387 violation risk)

Figure 3.4 continued



(b)

Figure 3.4 continued



(c)

The tradeoff solutions for the E-URM I are displayed in Figure 3.5. The x-axis ("Violation Risk") corresponds to objective (3.5) and (3.11), which equals the total risk of relaxed potential proximity constraints. The y-axis ("Benefit") corresponds to objective (3.4) and indicates total economic return. As Figure 3.5 shows, accounting for spatial uncertainty in the model means that the total economic returns 5571.62 (106 units harvested) but as high as 6876.95 (142 units harvested). This is in contrast to 5854.25 identified using the URM, assuming boundary certainty, which means a 4.8% reduction in total benefit in the lowest case and a 17.5% increase in total benefit in the most optimistic case. The Pareto optimal curve also shows that increasing benefits means that the violation risk would increase, or rather more potential proximity constraints would need to be relaxed. For the 5571.62 solution in Figure 3.5, all certain and uncertain proximity restrictions are actually imposed (zero violation risk). On the other extreme, the 6876.95 solution is achieved by relaxing all potential restrictions. The other solutions in Figure 3.5 therefore reflect a tradeoff ranging between these extremes of 5571.62 to 6876.95 for economic return from forest harvesting and 0 to 387 for violation risk. In addition to the implications for total benefits of harvesting schedules, each solution represents a different spatial pattern. Figure 3.4b shows the spatial configuration of the 5571.62 solution, while that of the other extreme (6876.95) is displayed in Figure 3.4c. There are clear spatial pattern differences in Figure 3.4a, 3.4b and 3.4c, resulting in different economic returns.

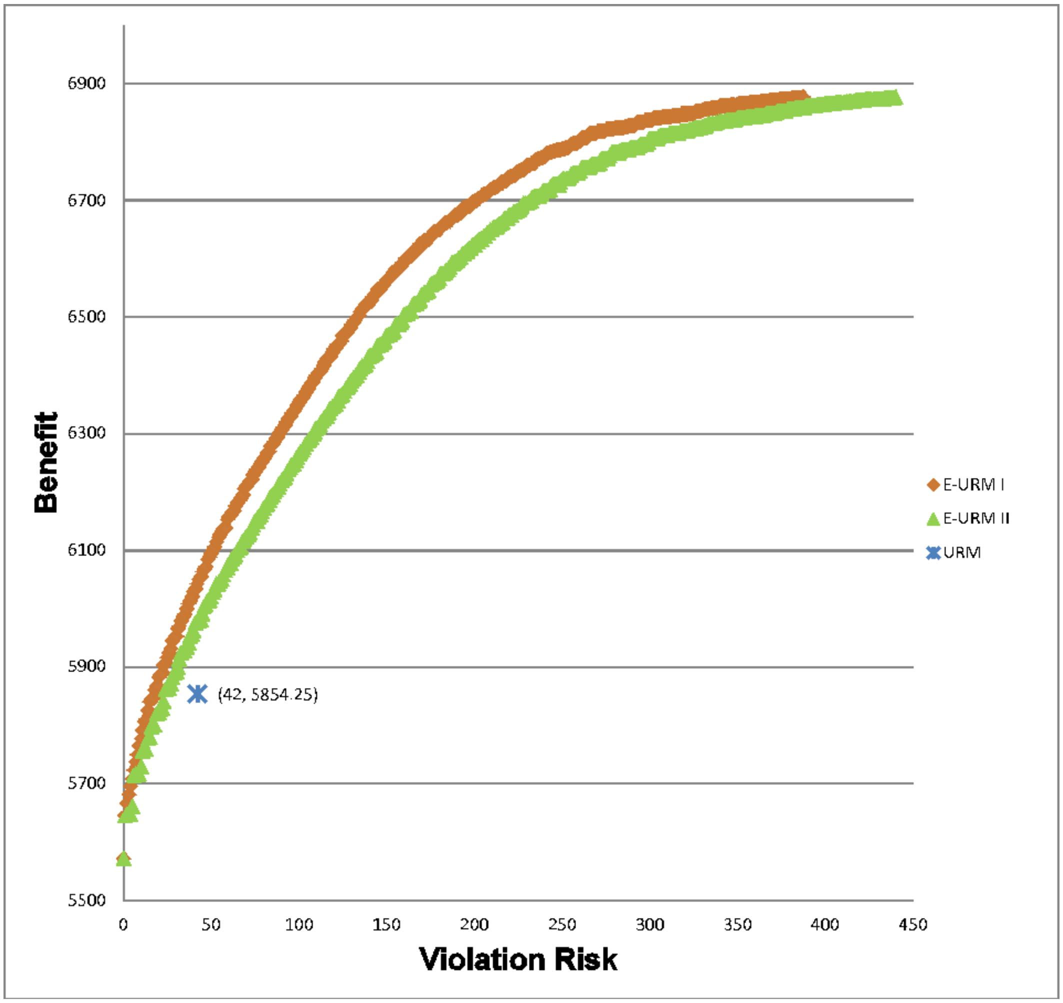
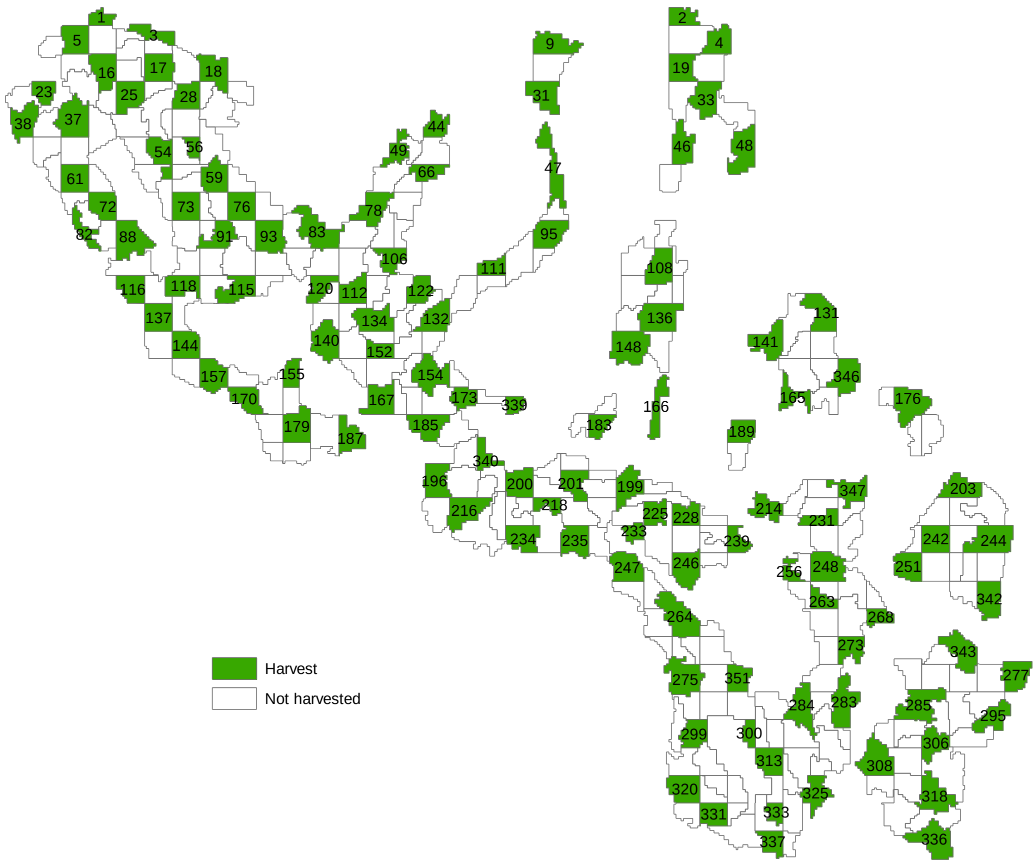


Figure 3.5 Tradeoff curves for the E-URM I and II

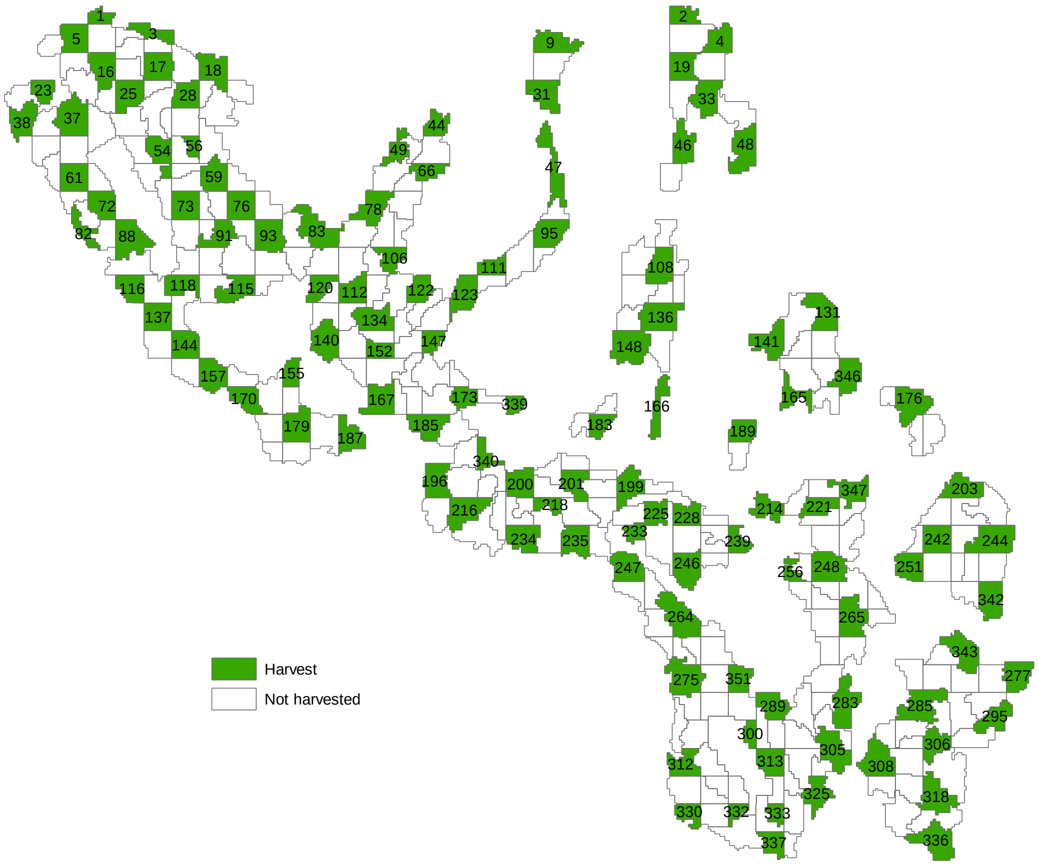
The tradeoff solutions for the E-URM II are also depicted in Figure 3.5, where each unit is associated with violation risk. The sum of violation risk associated with each unit ("Violation Risk") ranges from 0 to 440. Since the two extremes in Figure 3.5 characterize the scenarios of relaxing none and all potential restrictions, the range of total economic return is the same as that in E-URM I, and the spatial configurations of the two extremes are the same as for the E-URM I (Figure 3.4b and 3.4c). However, the other tradeoff solutions of the E-URM II do have higher violation risk than those of the E-URM I because all potential proximity restrictions associated with a unit is either relaxed or not relaxed but this may not be the case in the E-URM I. Additionally, they also exhibit varying spatial patterns. Figure 3.6b shows the spatial configuration of a solution giving 6504.22 benefit and 160 unit violation risk for the E-URM II in contrast to the E-URM I solution with 6502.03 benefit and 134 violation risk in Figure 3.6a. Unit 305 is relaxed in Figure 3.6b, indicating that the proximity restrictions between 305 and 283, and 305 and 308 are all relaxed. However, in Figure 3.6a only unit 283 and 308 are selected. This occurs because each pair of potential constraints is independent in the E-URM I while potential adjacency restrictions for a unit are combined in the E-URM II.



(a)

Figure 3.6 Spatial configuration comparison between the E-URM I and II: (a) E-URM I (6502.03 benefit and 134 violation risk), (b) E-URM II (6504.22 benefit and 160 violation risk)

Figure 3.6 continued



(b)

3.6 Discussion and conclusions

The results presented illustrate that spatial data uncertainty has real and significant impacts on the URM. If the potential uncertainty/error in spatial information is ignored, the modeling results could be erroneous, biased or misguided. The Error-Unit Restriction Model I and II (E-URM I and E-URM II) were developed to account for spatial uncertainty. In order to address spatial uncertainty, the E-URM I and II utilized multiple objectives. This approach enables a range of tradeoff solutions to be identified, which makes analysis more complicated, but they remain finite and can be evaluated to assess implications in harvest scheduling. Specifically, considering spatial uncertainty shows that the total harvest benefits could range widely from 5571.62 to 6876.95, in contrast to 5854.25 suggested using the URM.

An issue that might be raised is to consider derived potential risk and simply impose those adjacency restrictions above a certain threshold and ignore those below the threshold. Doing so would mean that the URM could be applied. While simple and straightforward, such an approach fails to incorporate important information about risk and economic return that is considered simultaneously. Comparison of the E-URM I and II results with that of the case where all potential adjacency restrictions with violation risk less than a prespecified threshold are relaxed confirms the inferiority of such an approach as only dominated solutions are identified. That is, the threshold approach would only enable solutions in the interior/dominated region of the tradeoff solution space, and in most cases these are far from the Pareto frontier (see Figure 3.7).

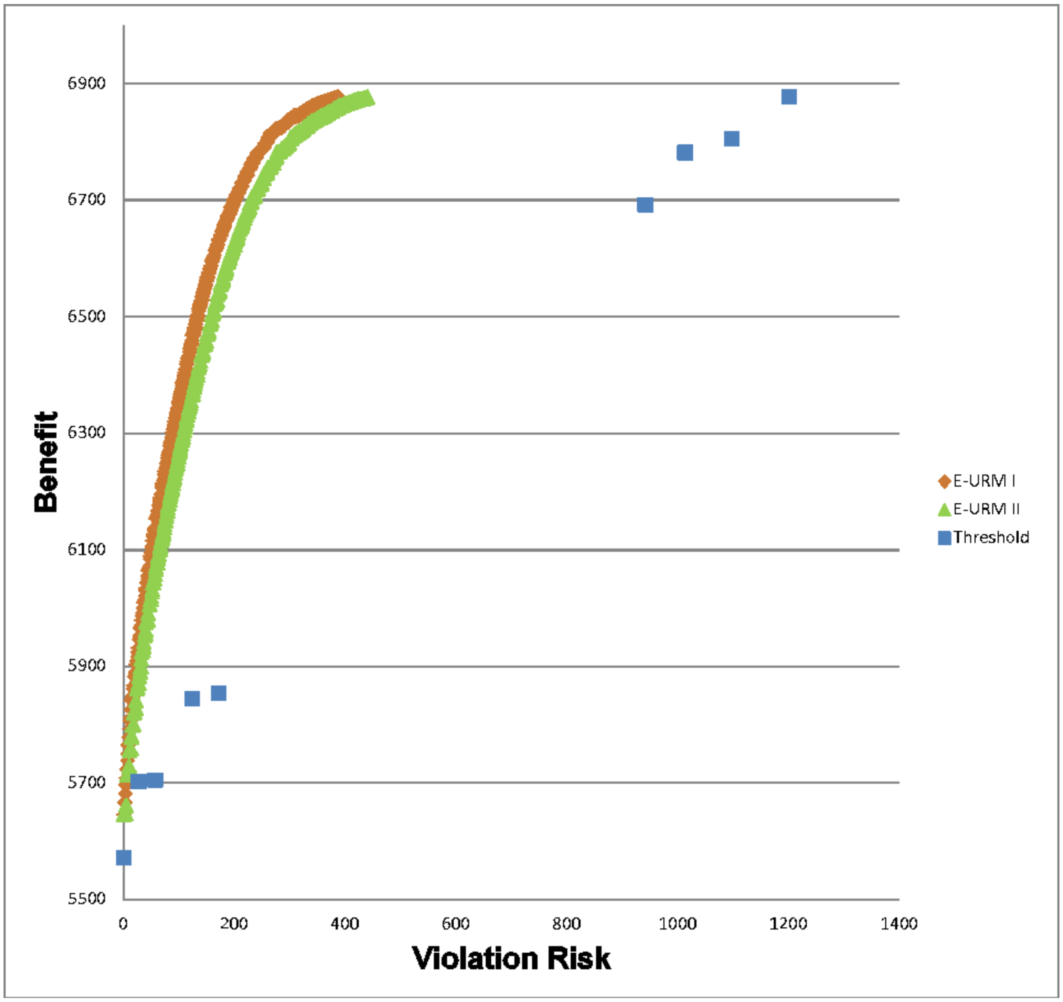


Figure 3.7 Comparison of threshold based solutions

Worth further discussion is the determination of violation risk. One possible way to impose violation risk by using minimum boundary distance and shared boundary length was detailed here, but there could be many other different ways to define such risk, depending on the practical application. Any approach for viewing potential adjacency in probabilistic terms may be of interest to consider. Both modeling approaches can readily accommodate this.

This chapter presented new ways to address spatial uncertainty in harvest scheduling by introducing the error unit restriction model I and II (E-URM I and E-URM II). We demonstrated that spatial data uncertainty could have significant impacts on forest planning. It remains to be seen how spatial uncertainty might be considered more generally in other spatial models used to support forest management planning. Clearly this is an important first step.

CHAPTER 4

A MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM FOR FACILITY DISPERSION UNDER CONDITIONS OF SPATIAL UNCERTAINTY*

As shown in Chapters 2 and 3, multi-objective models that explicitly account for spatial uncertainty are computationally challenging to solve. This chapter develops a multi-objective evolutionary algorithm to address the computational challenges posed by multi-objective approaches. The proposed algorithm incorporates problem-specific spatial knowledge to significantly enhance the capability of the evolutionary algorithm for solving these problems.

4.1 Introduction

Dispersion is essential in many environmental and urban planning contexts, including avoiding market saturation, determining forest harvest schedules, strategically placing military outposts, and locating undesirable facilities, among others. The intent is to identify outcomes that limit localized impacts, achieve sustainability and generally reflect an equitable distribution of services. An overview may be found in Church and Murray (2009).

One of the most widely applied dispersion approaches is the anti-covering location problem (ACLP), formally presented in Moon and Chaudhry (1984) to support location decision-making. In order to avoid concentration, this model

* This chapter represents a revised version of a paper submitted to *Journal of the Operational Research Society*, co-authored with Dr. Alan T. Murray.

aims to determine the placement of facilities that can generate maximum benefits while a prespecified spatial separation is imposed. The ACLP is effectively equivalent to node packing, vertex packing, r-separation, and maximal independent set problems. Technical discussion of these problems can be found in Nemhauser and Trotter (1975), Lawler et al. (1980), Nemhauser and Sigismondi (1992), and Erkut et al. (1996), among others. Though the ACLP and related problems are utilized extensively, they remain challenging to solve optimally for medium or large sized problem instances. This is not surprising as they belong to the class of NP-hard combinatorial optimization problems, indicating that there is no polynomial time algorithm to solve them (Garey and Johnson 1979). The challenges to solve the ACLP, combined with its broad application, have made it the focus of continued research efforts devoted to efficient solution. For exact solution approaches, improved mathematical structure has been sought. For example, Nemhauser and Sigismondi (1992) proposed a strong mathematical formulation by introducing clique and odd-hole inequalities. Murray and Church (1997) incorporated both neighborhood constraints and clique constraints, demonstrating significant computational benefits. Other examples can be found in Caprara et al. (2000), Strijk et al. (2000), Goycoolea et al. (2005) and Murray and Kim (2008). In addition to exact approaches, a variety of heuristic methods have also been developed. Examples include Greedy Search (Chaudhry et al. 1986; Feo et al. 1994; Cravo et al. 2008; Gamarnik and Goldberg 2010), Tabu Search (Gendreau et al. 1993; Strijk et al. 2000; Wu and Hao 2011), Simulated Annealing (Fleischer 1994; Strijk et al. 2000), Lagrangian Relaxation (Murray and Church

1997; Ribeiro et al. 2011), and Genetic algorithm (Hifi 1997; Chaudhry 2006). These techniques have considerably improved computational capabilities for solving this problem.

Beyond the computational difficulty in obtaining optimal or near optimal solutions, recent work highlights important data uncertainty issues. As the ACLP is a deterministic model, assuming the input of the model to be precise and accurate may be problematic in many ways. Chapter 2 has demonstrated that the ACLP and related models are sensitive to spatial data uncertainty, and proposed a multi-objective formulation that explicitly accounts for data uncertainty. The new model, referred to as the error-anti-covering location problem (E-ACLP), can identify trade-off solutions reflecting the range of potential impacts associated with data uncertainty on modeling results. While this model provides an effective approach to evaluate the implications of spatial uncertainty, it also requires significantly more computational effort to solve because it is a multi-objective extension of ACLP. This is particularly important given that the ACLP is itself a NP-hard problem and subject to considerable research on efficient solution techniques. There is clearly a need to develop heuristic approaches to solve the E-ACLP, as many practical problem instances simply cannot be solved otherwise. In this chapter we develop a multi-objective evolutionary algorithm to address the computational challenges of solving the E-ACLP. The next section introduces the spatial optimization model considered here and reviews solution approaches for this problem. Following this, genetic algorithms (GAs) and multi-objective genetic algorithms (MOGAs) are briefly discussed. Then, the design of a MOGA for the E-ACLP is detailed. Application results are then presented and discussed. The chapter ends with a summary and concluding remarks.

4.2 Modeling spatial uncertainty

As noted previously, the purpose of anti-covering location problem (ACLP) and related models is to maximize the benefits associated with selected units while maintaining a minimum spatial separation between them. Units in this context could, for example, refer to management units to be harvested or commercial parcels for placing waste recycling centers or other types of service facilities. Consider the following notation:

$\beta_i = benefit\ associated\ with\ selecting\ unit\ i$

$\Phi_i = set\ of\ units\ that\ are\ in\ conflict\ with\ unit\ i$

$$x_i = \begin{cases} 1, & if\ unit\ i\ is\ selected, \\ 0, & otherwise. \end{cases}$$

$i = index\ of\ units$

The ACLP can be formulated as follows (Church and Murray 2009):

$$Maximize \qquad \sum_i \beta_i x_i \tag{4.1}$$

$$Subject\ to \qquad x_i + x_j \leq 1 \quad \forall i, j \in \Phi_i \tag{4.2}$$

$$x_i = \{0,1\} \quad \forall i \tag{4.3}$$

The objective, (4.1), is to maximize the benefits associated with selecting units. Constraints (4.2) ensure that no two conflicting units can be selected simultaneously. Binary integer requirements are stipulated in constraints (4.3).

The defining feature of the ACLP is the dispersion orientation in siting/selecting units. The conflict sets, Φ_i, reflect this, and are assumed to be known precisely. However, it is widely acknowledged and accepted that spatial data is fraught with uncertainty and error (see Goodchild and Gopal 1989). Spatial data uncertainty, therefore, creates ambiguity in Φ_i. New sets can be introduced, Ω_i and Ψ_i, to represent certain and uncertain conflicts, as was done in Chapter 2 to propose an extension of the ACLP. The model is referred to as the E-ACLP and relies on the following additional notation:

$$\Omega_i = set\ of\ units\ that\ are\ certainly\ in\ conflict\ with\ unit\ i$$

$$\Psi_i = set\ of\ units\ that\ are\ possibly\ in\ conflict\ with\ unit\ i$$

$$p_{ij} = probablity\ that\ unit\ i\ is\ in\ conflict\ with\ unit\ j$$

$$y_{ij} = \begin{cases} 1, if\ \ the\ proximity\ restriction\ \ between\ units\ i\ and\ j\ is\ relaxed, \\ 0, \qquad otherwise. \end{cases}$$

Ω_i is introduced as a conservative conflict set that is deemed to be certain. Separation restrictions between unit i and members of Ω_i will always be imposed. Alternatively, the uncertain conflicts are represented by the set Ψ_i. It may or may not be necessary to impose restrictions between unit i and members of Ψ_i, depending on the probability of a conflict, p_{ij}. Binary decision variables y_{ij} track whether the uncertain proximity constraints are imposed or not.

Using the above notation, the E-ACLP is formulated as follows (Chapter 2):

$$Maximize \qquad \sum_i \beta_i x_i \tag{4.4}$$

$$Minimize \qquad \sum_i \sum_{j \in \Psi_i} p_{ij} y_{ij} \tag{4.5}$$

$$Subject\ to \qquad x_i + x_j \leq 1 \quad \forall i, j \in \Omega_i \tag{4.6}$$

$$x_i + x_j - y_{ij} \leq 1 \quad \forall i, j \in \Psi_i \tag{4.7}$$

$$x_i = \{0,1\} \quad \forall i \tag{4.8}$$

$$y_{ij} = \{0,1\} \quad \forall i, j \in \Psi_i$$

The first objective, (4.4), remains to maximize the benefits associated with selecting units, as structured in the ACLP. The second objective, (4.5), is to minimize the total probability of relaxing separation constraints that are spatially uncertain. The two objectives are in conflict in the sense that gains in benefits can be achieved only by relaxing uncertain proximity constraints. Constraints (4.6) ensure that no two selected sites conflict among the certain restrictions. Constraints (4.7) track separation of those that might be in conflict. These are the constraints that could be okay to relax or ignore, depending on uncertainty. When y_{ij} equals one, both x_i and x_j could be one in constraint (4.7), indicating both units could be selected; otherwise, at most one of them can be selected. Constraints (4.8) impose binary integer restrictions on decision variables.

When comparing the formulation of the ACLP with the E-ACLP, an equivalency can be observed when $\Omega_i = \Phi_i$ and $\Psi_i = \emptyset$. That is, both models are exactly the same in this case. In practice, however, it would be expected that $\Omega_i \subset \Phi_i$ and $(\Omega_i \cup \Psi_i) \supset \Phi_i$. The implication is that the total benefits would increase when uncertain proximity constraints are relaxed as fewer overall spatial restrictions would be imposed. Alternatively, total benefits would decrease when uncertain restrictions are imposed, because more restrictions would be imposed relative to the ACLP.

4.3 Solution approaches

The E-ACLP is a multi-objective model, requiring identification of tradeoff solutions using multi-objective solution techniques. One popular approach is the weighting method (Cohen 1978), where the two objectives are combined using a weight w. This is accomplished as follows:

$$Maximize \qquad w\sum_i \beta_i x_i - (1-w)\sum_i \sum_{j\in\Psi_i} p_{ij} y_{ij} \qquad (4.9)$$

Objectives (4.4) and (4.5) can be replaced by objective (4.9), and the model solved. By varying the weight from 0 to 1, different problem scenarios arise and tradeoff solutions can be found. Given that it is often impossible in practice to enumerate all possible values of the weight, w, techniques have been proposed to sample weight values to identify Pareto-optimal solutions (Eswaran et al. 1989; Solanki 1991; Ralphs et al. 2006). However, it may not be possible to find all nondominated solutions, significant time may be spend finding non-unique solutions, and the problem(s) may simply be too difficult to solve in a reasonable amount of time, if at all.

Another approach to find nondominated tradeoff solutions is the constraint method (Cohen 1978), where one objective is integrated into the model as a constraint and the second objective optimized. Assuming that K is a feasible solution for the second objective, then the E-ACLP can be represented as:

$$Maximize \qquad \sum_{i} \beta_i x_i \tag{4.10}$$

$$Subject\ to \qquad \sum_{i} \sum_{j \in \Psi_i} p_{ij} y_{ij} = K \tag{4.11}$$

$$x_i + x_j \leq 1 \quad \forall i, j \in \Omega_i \tag{4.12}$$

$$x_i + x_j - y_{ij} \leq 1 \quad \forall i, j \in \Psi_i \tag{4.13}$$

$$x_i = \{0,1\} \quad \forall i \tag{4.14}$$

$$y_{ij} = \{0,1\} \quad \forall i, j \in \Psi_i$$

Since the probabilities, p_{ij}, can always be scaled to integers and y_{ij} are binary decision variables, the possible values of $K$are finite. By iterating all potential values of K, different single-objective models result and can be solved using exact IP approaches. Given this, the constraint method can ensure the identification of all nondominated solutions but requires solving the transformed single-objective model many times. While nice in theory, it remains that the E-

ACLP is NP-hard and unlikely to be optimally solved in some instances or for large size problems. In fact, attempts to solve some of the reported problems here using an exact IP approach proved impossible.

4.4 Multi-objective genetic algorithms (MOGAs)

In multi-objective optimization problems the objectives are usually conflicting with each other. Thus no single solution represents a best case for all objectives. As a result, the ultimate goal of a solution technique for multi-objective optimization problems is to identify a complete set of nondominated or Pareto-optimal solutions that cannot be improved with respect to any objective without degrading at least one other objective (Cohen 1978; Konak et al. 2006). An efficient multi-objective optimization algorithm should therefore be capable of identifying or approximating the Pareto-optimal front in order to reflect the range and diversity of tradeoff solutions possibly.

Based on the principle of natural selection, genetic algorithms (GAs) are well suited to solve mutli-objective optimization problems (Deb 2001). GAs operate with a population of chromosomes (solutions) and can capture a diverse set of chromosomes (solutions) in a single generation (run) (Deb et al. 2002). As the evolution goes through crossover, mutation and reproduction, the population is able to identify or approximate the Pareto-optimal front. GAs theoretically provide the capacity to satisfy both convergence and diversity goals, important to multi-objective optimization problems, which explains why they have been widely applied. Jones et al. (2002) reported that 70% of multi-objective heuristic approaches are based on GAs. In the recent years, there has been increasing interest in designing multi-objective GAs (MOGAs) for spatial optimization problems. Examples can be found in Xiao et al (2002), Bennett et al. (2004), Kim et al (2009), Cao et al. (2011), Wu and Grubesic (2010), Roberts et al. (2011) and Wu et al. (2011).

Even though the basic idea of GAs is universal, it is the actual design and implementation that determine the success and performance of the algorithm, and whether the Pareto-optimal front can be found or sufficiently approximated (Alp et al. 2003; Konak et al. 2006). Beasley et al. (1993), Bennett et al. (2004), Xiao (2008) and Tong et al. (2009) demonstrated the necessity to incorporate problem-specific knowledge into GA design. To this end, we propose a new hybrid initialization procedure and greedy feasibility operator to enhance the performance of MOGA for solving the E-ACLP. The overall design of the algorithm follows Deb et al. (2002) but integrates the hybrid initialization procedure and greedy feasibility operator ideas proposed here. The overall design of the developed MOGA for the E-ACLP is summarized in Figure 4.1. While each GA component is important, more emphasis is placed on the initialization and feasibility operator given their novelty and capacity for improving the performance of MOGA. Details of GA components are now presented.

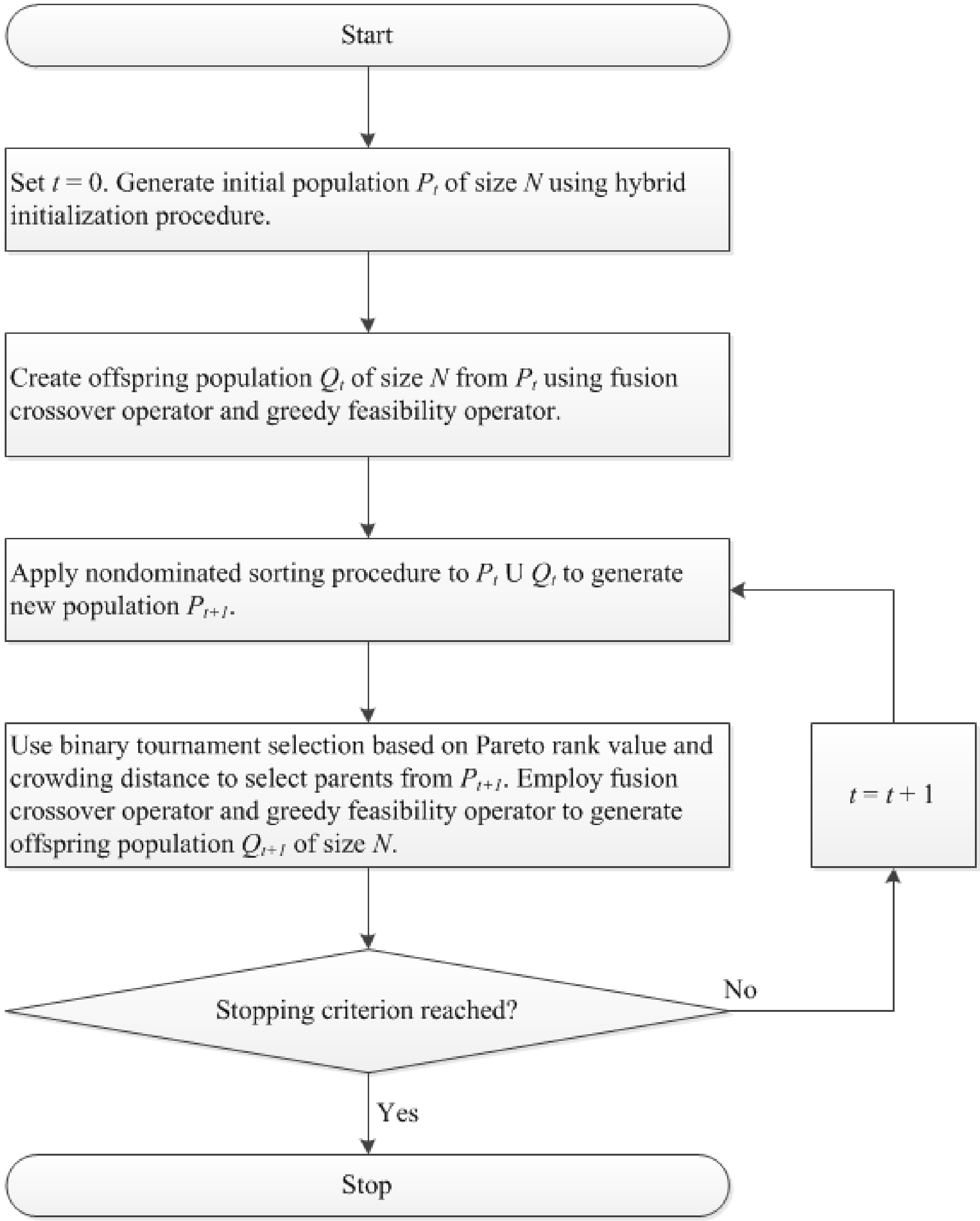


Figure 4.1 Flow diagram of the MOGA for solving the E-ACLP

4.4.1 Representation and Initialization

Solutions are represented using a simple n-bit binary string where n is the number of potential areas. A value of 1 at the ith bit indicates that unit i is selected, while 0 implies that it is not selected. An important issue regarding the binary encoding scheme is that the resulting offspring solutions after genetic operations may be infeasible. Infeasible solutions are modified, or repaired, using a greedy feasibility operator.

Being a population-based approach, GAs require generating a set of initial solutions to start the basic search process. A random initialization procedure is quite common in GA design. A straightforward approach for the E-ACLP could be:

- B-1. Solution X is a n-bit string of zeros and candidate set S contains all potential units: $S = \{1, \dots, n\}$.
- B-2. Randomly select a unit i from candidate set S and set $X[i]$ to equal 1.
- B-3. Remove the certain conflict units Ω_i from candidate set S: $S = S \backslash \Omega_i$.
- B-4. If candidate set S is empty, stop; otherwise, go to step B-2.

While this is easy to implement and can generate feasible solutions for E-ACLP, solution quality is not particularly good. Evidence suggests that good initial solutions will generally lead to better performance of GAs (Ahuja et al. 2000; Bennett et al. 2004), so a new hybrid initialization procedure is developed to generate more effective initial solutions. The new hybrid initialization procedure is as follows.

H-1. Impose all uncertain proximity constraints and solve the resulting model by exact or heuristic approaches.

H-2. Relax all uncertain proximity constraints and solve the resulting model by exact or heuristic approaches.

H-3. Add the solutions obtained in steps H-1 and H-2 to initial population.

H-4. Generate the remaining initial solutions randomly (B-1 through B-4).

The first two steps help ensure that the optimal or near-optimal solutions are found when relaxing all or none of the uncertain proximity restrictions for the initial population. Through the integration of good initial solutions, greater efficacy of the MOGA is expected compared to only random initialization.

4.4.2 Feasibility operator

As noted previously, offspring solutions generated using a crossover operator may be infeasible, indicating that some certain conflict conditions are not imposed. As a result, a feasibility operator is required to transform infeasible solutions into feasible ones. Xiao (2008) presents a general feasibility operator that keeps randomly selecting from the infeasible solution until none can be selected without violating constraints. Hifi (1997) developed a GA for ACLP, where infeasible solutions are repaired by sequentially removing selections that violate constraints (alter some genes from 1 to 0) and then adding all other potential areas that do not violate constraints (alter some genes from 0 to 1). Compared to the general feasibility operator in Xiao (2008), the repair procedure in Hifi (1997) might generate solutions with relatively higher benefits because it includes the areas initially not selected. However, sequentially removing conflicts could result in a poor solution, slowing down convergence. Of course, a good solution for the E-ACLP should not only maximize benefits but also minimize the probability of violating the potential proximity conflicts. To achieve quick convergence to a good solution and maintain solution diversity, the following greedy feasibility operator is proposed:

F-1. T is the total number of generations; t is the current generation; C is the binary string of input offspring solution; S represents the set of selected areas in C: $S = \{i|C[i] = 1, i = 1, \dots, n\}$.

F-2. Calculate the net benefit (NB) of each area in S by subtracting the total benefits of its neighbors in set S from its own benefits:

$$NB[i] = \beta[i] - \sum_{j \in S \cap (\Omega_i \cup \Psi_i)} \beta[j]$$

F-3. Pick the area i^* with the highest net benefit in S, remove its conflict set from set S and invert the corresponding indexes in C to 0:

$S = S \backslash \Omega_{i^*}$ and $C[j] = 0, \forall j \in \Omega_{i^*}$ with probability t/T, or

$S = S \backslash (\Omega_{i^*} \cup \Psi_{i^*})$ and $C[j] = 0, \forall j \in (\Omega_{i^*} \cup \Psi_{i^*})$ with probability $1 - t/T$;

F-4. If all areas in S has been picked, go to step F-5; otherwise, go to step F-3.

F-5. Consider the other areas that can be added into S, denoted as S':

$$S' = \{j | j \notin S \; and \; j \notin \bigcup_{i \in S} \Omega_i \, , j = 1, \ldots, n\}$$

F-6. Calculate the net benefit of each area in S' as step F-2.

F-7. Pick the area $i^{*\prime}$ with the highest net benefit in S', set $C[i^{*\prime}] = 1$, and remove it and its conflict set from set S':

$S = S \backslash (i^{*\prime} \cup \Omega_{i^{*\prime}})$ and $C[i^{*\prime}] = 1$ with probability t/T, or

$S = S \backslash (i^{*\prime} \cup \Omega_{i^{*\prime}} \cup \Psi_{i^{*\prime}})$ and $C[i^{*\prime}] = 1$ with probability $1 - t/T$;

F-8. If S' is empty, stop; otherwise, go to step F-7.

For the feasibility operator F-1 – F-8, NB ensures that areas with large benefits and less neighbor conflict will be preserved. This is in contrast to sequentially removing, along the lines suggested in Hifi (1997). The greedy feasibility operator can therefore generate higher quality solutions and speed up convergence. The proposed approach takes into account the minimization of conflict probabilities by incorporating the uncertain conflict set Ψ_i into both the calculation of net benefit (steps F-2 and F-6) and the exclusion of neighbor conflict (steps F-3 and F-7). At earlier generations (small t), when area i is selected, the areas in the uncertain conflict set Ψ_i will be highly likely to be excluded, leading to a small sum of conflict probabilities. Over time, only the certain conflict set will be excluded and large conflict probabilities could be expected. This greedy feasibility operator will enhance the diversity of solutions given its exploration of a range of conflict probabilities in different generations.

4.4.3 Other operator

In order to achieve efficient convergence, we select the best nondominated individuals from parent and offspring populations for the next generation. This is accomplished by implementing the nondominated sorting procedure detailed in Deb et al (2002). The crowding distance approach, preventing solutions in the same nondomination level from clustering in objective space, is also introduced to increase the diversity of solutions. Binary tournament selection is utilized to select parents for offspring production. Fusion crossover proposed by Beasley and Chu (1996) is employed to create new offspring solutions given its superiority in keeping good information of parents and enhancing the diversity of offspring solutions over other common operators. This crossover operator is performed with some probability (crossover rate) that determines how often two selected individuals will crossover. If not performed, the two selected individuals will be directly copied as offspring solutions.

4.5 Application results

The proposed MOGA for solving the E-ACLP was implemented in Visual C++ and executed on a Intel Xeon (2.53 GHz) computer running Windows with 6 GB of RAM. Two different planning problems are utilized to demonstrate the effectiveness of the proposed algorithm.

The first planning application involves harvest scheduling. The forest region (Butter Creek) is in Northern California, and contains 351 harvest units. Each unit has a delineated spatial boundary and an associated economic return. Harvest scheduling for this region has been reported in Murray and Weintraub (2002), Murray et al. (2004) and Goycoolea et al. (2005). The planning goal is to maximize the total economic return, but no two adjacent (spatially conflicting) units can be harvested simultaneously. When the ACLP is applied to this forest planning problem, it suggests that the maximum economic return that can be achieved without violating proximity conflicts is 5854.25. However, the boundaries of harvest units are inaccurate and could have implications for total economic return. In order to account for the observed 30 meter boundary uncertainty, the E-ACLP is applied to identify all nondominated solutions using the constraint method. Using Gurobi, a commercial optimization package, total solution time was 90 seconds to find the 275 nondominated solutions, with a maximum of 0.51 seconds required to solve any individual problem. These nondominated solutions indicate that the total economic return could be as low as 5571.62 with 0 conflict probability or as high as 6876.95 with 387 total conflict probabilities, which means a 4.8% decrease and 17.5% increase in total economic return, respectively, from the 5854.25 identified using ACLP. Having exact solutions in this case enables assessment of the proposed MOGA heuristic in terms of computation time as well as convergence to and diversity along the true Pareto-optimal front.

As the feasibility operator is applied to each infeasible child solution, the only required input parameters are population size and crossover rate. After empirical testing, the crossover rate was set as 1.0. The algorithm terminates when a prespecified number of generations is reached. Algorithm performance is evaluated using the following population size and generation characteristics: 100 population size and 100 generations, 100 population size and 500 generations, and 500 population size and 500 generations. Each parameter combination is run for 10 times.

Table 4.1 reports computational results for solving the forest planning problem using the proposed MOGA heuristic. The first two columns in Table 4.1 are associated with the heuristic parameters, Population and Generation. The next column corresponds to the number of nondominated solutions identified using MOGA. The “Average benefit gap” and Maximal benefit gap” columns in Table 4.1 evaluate how much the MOGA solutions deviate from the Pareto-optimal front. The final column is solution time in seconds. The information in Table 4.1 summarizes the average of ten runs of the heuristic beginning with different random initial solutions. Table 4.1 shows that for the 10 different applications of the heuristic 89 nondominated solutions are found on average in the case where population size is 100 and number of generations is 100. Further, the average benefit gap is 1.71%, the maximal benefit gap is 4.32%, and solution time is 4.89 seconds. When the number of generations is increased to 500, 115 nondominated solutions are found, the average benefit gap decreases to 1.04% and maximal benefit gap decreases to 2.55%, but solution time increases to 21.60 seconds. Finally, when both population and number of generations are increased to 500, the number of nondominated solutions increases to 166, but average and maximal benefit gap are almost the same as 100 population size. This also does require increased computation effort, with solution time of 145.68 seconds.

Table 4.1: Computational results of the proposed MOGA for harvest scheduling

Population	Generation	Nondominated solutions identified	Average benefit gap*	Maximal benefit gap**	Solution time (sec)
100	100	89	1.71%	4.32%	4.89
100	500	115	1.04%	2.55%	21.60
500	500	166	1.12%	2.65%	145.68

* The solution is compared to the Pareto-optimal front based on similarity of conflict probability value, objective (4.5). The benefit difference, objective (4.4), is then standardized and converted to a percentage. The average is based on the benefit gaps of all MOGA solutions.

** See above, but in this case the largest benefit gap of all MOGA solutions is reported.

Generally, the MOGA heuristic is able to identify high-quality nondominated solutions. It is also worth noting that more generations could significantly enhance the performance of MOGA, irrespective of the number of nondominated solutions identified or deviation from the Pareto-optimal front. Larger populations seem to not contribute much to improving the convergence to the Pareto-optimal front. This is important in practice because identifying a reasonable population size is critical to the efficiency of a GA.

The results in percentage terms are possible in Table 4.1 because the actual Pareto-optimal front is known. This can be visualized as well for the two objectives of the E-ACLP. Figure 4.2 shows the nondominated solutions of the best run of the MOGA heuristic (out of ten) as a function of benefit (objective 4.4) and conflict probability (objective 4.5) by different population size and generations in comparison to the true Pareto-optimal front. As Figure 4.2 illustrates, the solutions of the MOGA are evenly distributed and cover the entire objective space of the Pareto front, with benefits ranging from 5571.62 to 6876.95 and conflict probability ranging from 0 to 387. The divergence from the Pareto frontier is small, especially at the neighborhoods of minimum and maximum conflict probabilities. The ability of the MOGA heuristic to closely approximate the Pareto frontier is evident, supporting the observation that the heuristic is performing well.

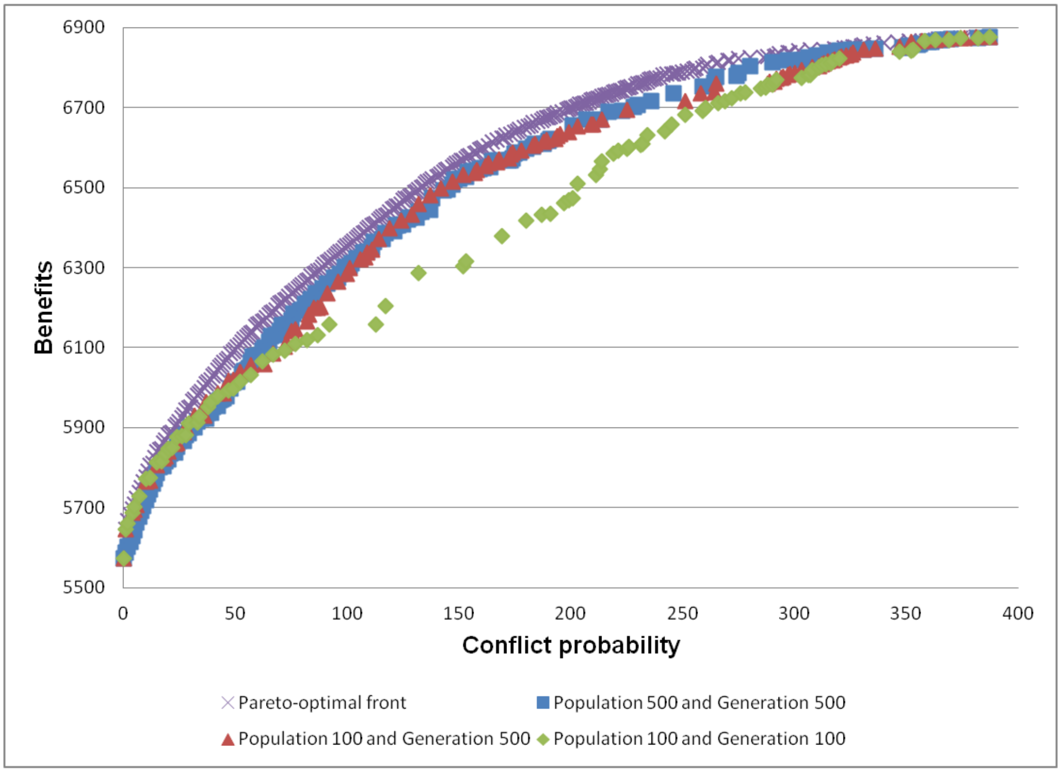


Figure 4.2 Nondominated solutions by different population size and generations for harvest scheduling

The second planning application involves estimating the number of alcohol outlets that could be located in Philadelphia after imposing a 200 feet proximity restriction between outlets. Grubesic et al. (2012) detailed how dispersion models can be used to assess public policy changes being considered in Pennsylvania to privatize alcohol sales, where distributors will seek to maximize customer access. Thus, the question arises regarding how many outlets can be expected and where will they be located, and ultimately what will the health and safety implications be. 11,226 potential outlet areas in Philadelphia are found through GIS-based suitability analysis. Based on parcel data accuracy and other sources of spatial uncertainty, the parcel boundary error is estimated to be $\pm$25 feet. To evaluate the impacts of spatial uncertainty on the number of expected alcohol outlets, the corresponding E-ACLP is structured. Unfortunately, it is not possible to optimally solve this problem using an exact IP method. For example, when K is 100 in equation (4.11), the resulting single objective model cannot be solved using Gurobi after running for seven days. As all nondominated solutions are needed, this is a significant problem. The MOGA heuristic is therefore essential.

Using the same parameters previously reported, the solutions using the MOGA heuristics are reported in Table 4.2. Given the large problem size (11,226 potential areas), the solution time ranges from 2631.56 seconds to 64,803 seconds when population size and generations increase from 100 to 500. In addition, 150, 198 and 197 nondominated solutions are found respectively, and are displayed in Figure 4.3. The differences between solution convergence for 500 population size and 100 population size are not large either, suggesting that it is possible to obtain good quality solutions for large size problem using a small population size. Though it is difficult to evaluate convergence of the MOGA solutions without knowing the true Pareto-optimal front, Figure 4.3 shows that a diverse set is found, where the maximum number of alcohol outlets could actually range from 2,896 (0 conflict probability) to 3,288 (1101 conflict probability). When compared to the 3,073 outlets identified using the ACLP, assuming parcel boundary certainty, these represents a 5.76% reduction and 7% increase in the total number of alcohol outlets.

Table 4.2: Computational results of the proposed MOGA for assessing alcohol outlets

Population	Generation	Nondominated solutions identified	Solution time (sec)
100	100	150	2631.56
100	500	198	13075.00
500	500	197	64803.00

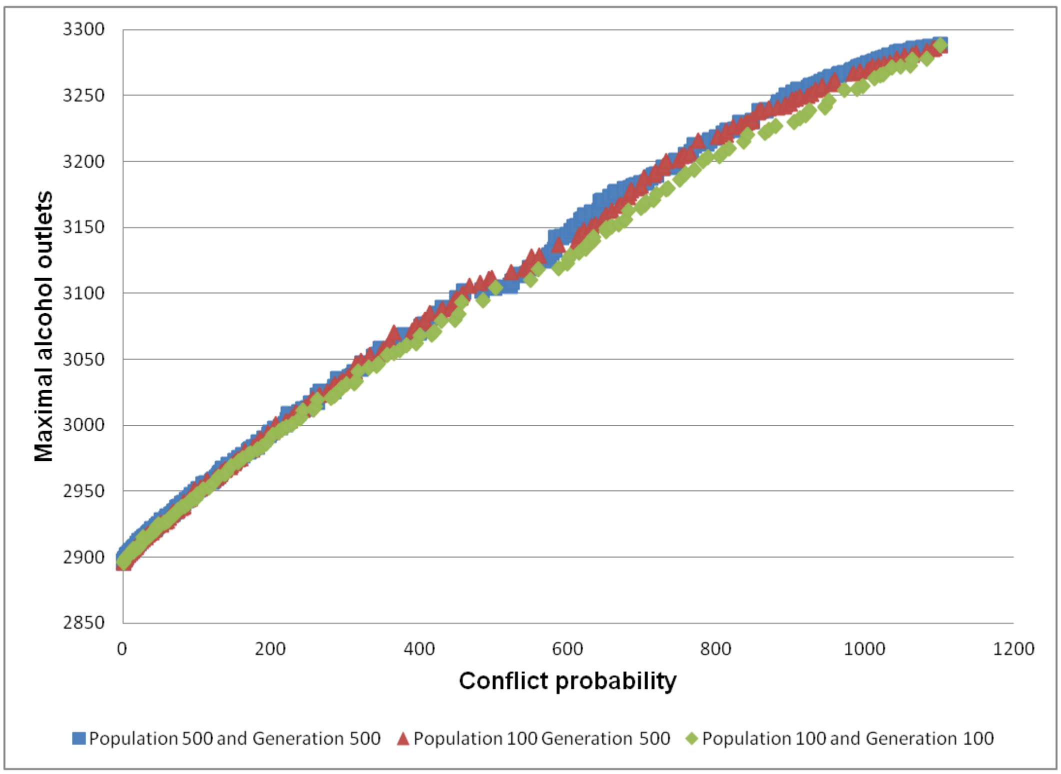


Figure 4.3 Nondominated solutions by different population size and generations for assessing alcohol outlets

4.6 Discussion

The application results presented in previous section demonstrate that the MOGA heuristic is effective in solving the E-ACLP in many ways. First, it is the first heuristic proposed to solve the E-ACLP and proved to be effective. Second, the average deviations of the MOGA solutions from the Pareto frontier are small, only 1.04% to 1.71% in the case of harvest scheduling, indicating that good quality solutions have been found. Third, the nondominated solutions identified using the MOGA are very diverse and evenly distributed (see Figures 4.2 and 4.3). Finally, the MOGA heuristic makes the large size applications computationally feasible to solve, using 2631.56 seconds to identify 150 nondominated solutions, as the E-ACLP could not be solved for assessing alcohol outlets using exact methods. This is important and meaningful because many practical planning problems often exceed the capabilities of exact approaches.

Worth further discussion is the distinction of the proposed MOGA from other GA approaches. As noted previously, a new hybrid initialization procedure and greedy feasibility operator are developed in the MOGA heuristic, both of which result from the exploration of problem-specific spatial knowledge. This is critical to good performance for the MOGA. To demonstrate this point, two other possible GA approaches are also assessed and compared. One is the general GA approach detailed in Xiao (2008) and the other is the GA approach used to solve ACLP in Hifi (1997). Since both approaches are designed for solving single objective optimization problems, their reproduction procedure is modified to use the nondominated sorting procedure utilized in the MOGA, but their genetic operators, like initialization, crossover and feasibility, remain intact. Nondominated solutions using 500 population size and 500 generations are compared against MOGA solutions using the same parameters and are displayed in Figure 4.4. The solutions identified using the other two GA approaches are obviously dominated by MOGA solutions. Further, these solutions tend to cluster in the medium conflict probability region. The superiority of the MOGA solutions can be attributed to the knowledge-based genetic operations. For example, the hybrid initialization increases the likelihood of GA exploration around low and high conflict probability values; the greedy feasibility operator allows more areas to be selected and larger benefits achieved.

The computational effort of the proposed MOGA also needs further discussion. Compared to other heuristics, GAs usually have larger computational loads (Xiao 2008). This can be observed when solving the large size problem. More work is therefore needed to improve its computational efficiency. One possible way is the parallelization of the algorithm, but is left for future research.

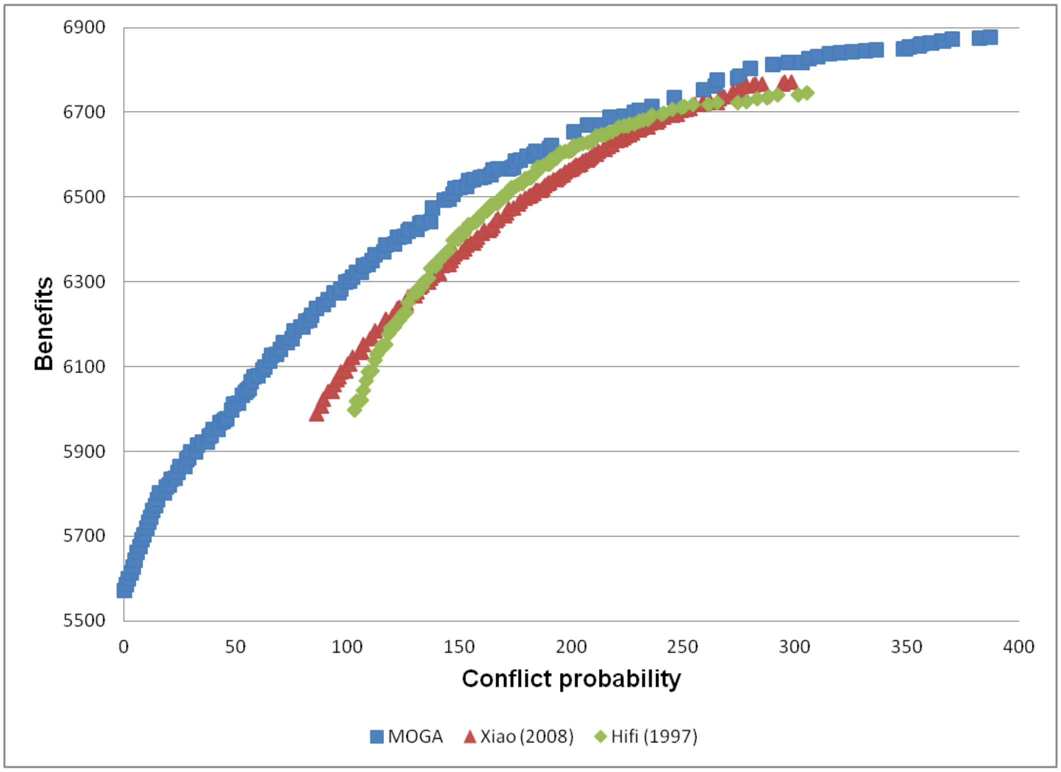


Figure 4.4 Nondominated solutions using different GA approaches for harvest scheduling

4.7 Conclusion

This chapter presents a new MOGA for solving the E-ACLP. This model addresses spatial data uncertainty issues in the selection of units to maximize the benefits associated with selected units while maintaining a minimum spatial separation between them. Even though the general idea of a GA is the same in most applications, there is no generic GA that works well for any problem. The proposed MOGA incorporates problem-specific spatial knowledge by introducing a hybrid initialization procedure and greedy feasibility operator to considerably enhance the capacity of GA to solve the E-ACLP. Compared with Pareto-optimal solutions derived by exact methods, the application results demonstrate that the new MOGA heuristic is able to generate a diverse set of solutions close to the Pareto-optimal front, achieving both diversity and convergence goals for solving multi-objective optimization problems. In addition, as the first heuristic approach for solving the E-ACLP, the MOGA is capable of identifying good quality solutions in a reasonable amount of time for large size planning problems that cannot be solved using exact approaches. The superior performance of the new MOGA is also affirmed by comparison to other possible GA approaches.

CHAPTER 5

CONCLUSIONS

5.1 Summary

It is well acknowledged that error or uncertainty always exists in geographic information (Longley et al. 2011). However, it is still a challenge to evaluate impacts on analysis and decision making. This research focused on addressing geographic uncertainty in spatial optimization. We first developed a multi-objective approach that explicitly accounts for spatial uncertainty in dispersion modeling, enabling the impacts of uncertainty to be evaluated with statistical confidence. Further, in the context of harvest scheduling, uncertainty in contiguity-based adjacency was assessed and an alternative modeling approach to integrate spatial uncertainty was proposed and compared. In addition, to address the computational challenges of the new multi-objective model, a new multi-objective genetic algorithm is developed and empirical results demonstrated its performance superiority in supporting facility and service planning.

Chapter 2 proposed a multi-objective extension of dispersion model to take into account spatial data uncertainty. It showed that geographic data uncertainty would have significant implications on proximity determination. Some proximity may not hold any longer and additional proximity may occur after considering data uncertainty. Such uncertainty in proximity is incorporated into the model and lead to the new multi-objective formulation. Solving this new model can identify trade-off solutions reflecting the range of potential impacts associated with data uncertainty on modeling results.

Chapter 3 discussed spatial uncertainty in harvest scheduling, where proximity is evaluated using contiguity-based measures. This is in contrast to the distance-based measures in Chapter 2. A new algorithm integrating both shared boundary length and minimum distance was developed to assess uncertainty in contiguity-based adjacency. In addition, we also proposed an alternative modeling approach to deal with uncertainty issues in dispersion modeling. The results of the two modeling approaches were also compared.

The new multi-objective models presented in Chapters 2 and 3 are NP-hard, requiring considerable computational efforts to solve optimally. Chapter 4 developed an efficient heuristic approach for solving these multi-objective models. This heuristic incorporated problem-specific spatial knowledge to significantly enhance the capability of the evolutionary algorithm for solving this problem. Application results also showed that high-quality nondominated solutions could be identified using the algorithm in a reasonable time.

5.2 Future work

5.2.1 Exploring alternative representations of geographic uncertainty

In this research, geographic data uncertainty is described by an error band, ε, around each study unit. However, a more accurate probabilistic description of error/uncertainty may be possible. For example, the error is likely to be normally distributed. In addition, spatial data uncertainties are assumed to be independent, but they could be highly correlated (Keefer et al. 1988). More work is needed to explore enhanced knowledge of geographic uncertainty.

5.2.2 Improving computational efficiency

While the proposed multi-objective genetic algorithm is capable of identifying good solutions in a reasonable amount of time, there is still a need to improve computational efficiency. As shown in the application results, the computational load of the algorithm for solving large sized problem is substantial. More work is therefore needed to seek out efficiencies. One possible way forward is parallelization of the algorithm. Of course, more research is needed to explore this possibility.

5.2.3 Incorporating uncertainty in other spatial optimization models

This research developed two multi-objective extensions for dispersion models to address spatial uncertainty issues. Such rationale could be applied to other spatial optimization models, such as the location set covering problem (LSCP) and the maximal covering location problem (MCLP). If spatial uncertainty is accounted for, facility coverage is no longer deterministic but probabilistic, which could be integrated into coverage modeling by constructing new multi-objective models. More research is therefore needed to investigate this.

BIBLIOGRAPHY

Abbaspour, R.A., Delavar, M.R. and Batouli, R., 2003, The issue of uncertainty propagation in spatial decision making. *Proceedings of the Scandinavian Research Conference on Geographical Information Science*, pp. 57-65.

Aerts, J.C.J.H., Goodchild, M.F. and Heuvelink, G., 2003, Accounting for spatial uncertainty in optimization with spatial decision support systems. *Transactions in GIS*, 7, pp. 211-230.

Aerts, J.C.J.H. and Heuvelink, G.B.M., 2002, Using simulated annealing for resource allocation. *International Journal of Geographical Information Science*, 16, pp. 571-587.

Ahuja, R.K., Orlin, J.B. and Tiwari, A., 2000, A greedy genetic algorithm for the quadratic assignment problem. *Computers & Operations Research*, 27, pp. 917-934.

Alçada-Almeida, L., Coutinho-Rodrigues, J. and Current, J., 2009, A multiobjective modeling approach to locating incinerators. *Socio-Economic Planning Sciences*, 43, pp. 111-120.

Alp, O., Erkut, E. and Drezner, Z., 2003, An efficient genetic algorithm for the p-median problem. *Annals of Operations Research*, 122, pp. 21-42.

AltInel, 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, pp. 790-799.

Andersson, G., Francis, R.L., Normark, T. and Rayco, M.B., 1998, Aggregation method experimentation for large-scale network location problems. *Location Science*, 6, pp. 25-39.

Arbia, G., Griffith, D. and Haining, R., 1998, Error propagation modelling in raster GIS: overlay operations. *International Journal of Geographical Information Science*, 12, pp. 145-167.

Ascough Ii, J.C., Maier, H.R., Ravalico, J.K. and Strudley, M.W., 2008, Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling*, 219, pp. 383-399.

Bach, L., 1981, The problem of aggregation and distance for analyses of accessibility and access opportunity in location - allocation models. *Environment and Planning A*, 13, pp. 955-978.

Barahona, F., Weintraub, A. and Epstein, R., 1992, Habitat dispersion in forest planning and the stable set problem. *Operations Research*, 40, pp. 14-21.

Beasley, D., Bull, D.R. and Martin, R.R., 1993a, An overview of genetic algorithms: Part 1. Fundamentals. *University computing*, 15, pp. 58-58.

Beasley, D., Bull, D.R. and Martin, R.R., 1993b, An overview of genetic algorithms: Part 2, research topics. *University computing*, 15, pp. 170-181.

Beasley, J. and Chu, P.C., 1996, A genetic algorithm for the set covering problem. *European Journal of Operational Research*, 94, pp. 392-404.

Beech, T., Dowd, M., Field, C., Hatcher, B. and Andréfou t, S., 2008, A stochastic approach to marine reserve design: Incorporating data uncertainty. *Ecological Informatics*, 3, pp. 321-333.

Bennett, D.A., Xiao, N. and Armstrong, M.P., 2004, Exploring the geographic consequences of public policies using evolutionary algorithms. *Annals of the Association of American Geographers*, 94, pp. 827-847.

Berman, O. and Odoni, A.R., 1982, Locating mobile servers on a network with Markovian properties. *Networks*, 12, pp. 73-86.

Bernhardsen, T., 1999, *Geographic information system: an introduction (second ed.).* (New York: Springer-Verlag.).

Bohle, C., Maturana, S. and Vera, J., 2010, A robust optimization approach to wine grape harvesting scheduling. *European Journal of Operational Research*, 200, pp. 245-252.

Bonneu, F. and Thomas-Agnan, C., 2009, Spatial point process models for location-allocation problems. *Computational Statistics & Data Analysis*, 53, pp. 3070-3081.

Bonneu, F., Thomas-Agnan, C. and Coelho, S., 2011, Educational planning: a simulation approach for the creation or closure of school classes. *Environment and Planning B: Planning and Design*, 38, pp. 595-615.

Boychuk, D. and Martell, D.L., 1996, A multistage stochastic programming model for sustainable forest-level timber supply under risk of fire. *Forest Science*, 42, pp. 10-26.

Boyland, M., Nelson, J. and Bunnell, F.L., 2005, A test for robustness in harvest scheduling models. *Forest Ecology and Management*, 207, pp. 121-132.

Boyland, M., Nelson, J., Bunnell, F.L. and D'Eon, R.G., 2006, An application of fuzzy set theory for seral-class constraints in forest planning models. *Forest Ecology and Management*, 223, pp. 395-402.

Brimberg, J. and Love, R.F., 1995, Estimating distance. In *Facility location: a survey of applications and methods.*, Z. Drezner (Ed.), pp. 9-32 (New York: Springer).

Brookes, C.J., 2001, A genetic algorithm for designing optimal patch configurations in GIS. *International Journal of Geographical Information Science*, 15, pp. 539-559.

Brown, J.D. and Heuvelink, G., 2007, The Data Uncertainty Engine (DUE): A software tool for assessing and simulating uncertain environmental variables. *Computers & Geosciences*, 33, pp. 172-190.

Bruin, D.S., Bregt, A. and van de Ven, M., 2001, Assessing fitness for use: the expected value of spatial data sets. *International Journal of Geographical Information Science*, 15, pp. 457-471.

Cánovas, L., García, S., Labbé, M. and Marín, A., 2007, A strengthened formulation for the simple plant location problem with order. *Operations Research Letters*, 35, pp. 141-150.

Cao, K., Batty, M., Huang, B., Liu, Y., Yu, L. and Chen, J., 2011, Spatial multi-objective land use optimization: extensions to the non-dominated sorting genetic algorithm-II. *International Journal of Geographical Information Science*, 25, pp. 1949-1969.

Caprara, A., Fischetti, M. and Letchford, A.N., 2000, On the separation of maximally violated mod-k cuts. *Mathematical Programming*, 87, pp. 37-56.

Carson, Y.M. and Batta, R., 1990, Locating an ambulance on the Amherst campus of the State University of New York at Buffalo. *Interfaces*, pp. 43-49.

Cayo, M.R. and Talbot, T.O., 2003, Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*, 2, p. 10.

Chaudhry, S.S., 2006, A genetic algorithm approach to solving the anti-covering location problem. *Expert Systems*, 23, pp. 251-257.

Chaudhry, S.S., McCormick, S.T. and Moon, I.D., 1986, Locating independent facilities with maximum weight: greedy heuristics. *Omega*, 14, pp. 383-389.

Church, R.L., 1999, Location modelling and GIS. In *Geographical information systems (second ed.)*, M.F.G. P. Longley, D. Maguire and D. Rhind (Ed.), pp. 293-303 (New York: John Wiley.).

Church, R.L., 2001, Spatial optimization models. In *International Encyclopedia of the Social & Behavioral Sciences*, N.J. Smelser and P.B. Baltes (Ed.), pp. 14811-14818 (Oxford: Elsevier).

Church, R.L., 2003, COBRA: a new formulation of the classic p-median location problem. *Annals of Operations Research*, 122, pp. 103-120.

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

Cohen, J.L., 1978, *Multiobjective programming and planning* (New York: Academic Press).

Constantino, M., Martins, I. and Borges, J.G., 2008, A new mixed-integer programming model for harvest scheduling subject to maximum area restrictions. *Operations Research*, 56, pp. 542-551.

Cooper, L., 1974, A random locational equilibrium problem. *Journal of Regional Science*, 14, pp. 47-54.

Cravo, G.L., Ribeiro, G.M. and Lorena, L.A.N., 2008, A greedy randomized adaptive search procedure for the point-feature cartographic label placement. *Computers & Geosciences*, 34, pp. 373-386.

Current, C. and Cohon, J., 1985, The maximum covering/shortest path problem: A multiobjective network design and routing formulation. *European Journal of Operational Research*, 21, pp. 189-199.

Current, J.R. and Schilling, D.A., 1990, Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis*, 22, pp. 116-126.

Das, A., Lele, S.R., Glass, G.E., Shields, T. and Patz, J., 2002, Modelling a discrete spatial response using generalized linear mixed models: application to Lyme disease vectors. *International Journal of Geographical Information Science*, 16, pp. 151-166.

Das, I. and Dennis, J.E., 1997, A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural and Multidisciplinary Optimization*, 14, pp. 63-69.

Daskin, M.S., Coullard, C.R. and Shen, Z.J.M., 2002, An inventory-location model: Formulation, solution algorithm and computational results. *Annals of Operations Research*, 110, pp. 83-106.

Daskin, M.S., Haghani, A.E., Khanal, M. and Malandraki, C., 1989, Aggregation effects in maximum covering models. *Annals of Operations Research*, 18, pp. 113-139.

De Bruin, S., Heuvelink, G. and Brown, J., 2008, Propagation of positional measurement errors to agricultural field boundaries and associated costs. *Computers and Electronics in Agriculture*, 63, pp. 245-256.

De Groeve, T. and Lowell, K., 2001, Boundary uncertainty assessment from a single forest-type map. *Photogrammetric Engineering and Remote Sensing*, 67, pp. 717-726.

De Jong, K.A., 1975, Analysis of the behavior of a class of genetic adaptive systems. University of Michigan.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T., 2002, A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, 6, pp. 182-197.

Devillers, R. and Jeansoulin, R., 2006, *Fundamentals of spatial data quality* (New York: John Wiley & Sons).

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, pp. 284-292.

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

Edwards, G. and Lowell, K., 1996, Modeling uncertainty in photointerpreted boundaries. *Photogrammetric Engineering and Remote Sensing*, 62, pp. 377-391.

Eid, T., 2000, Use of uncertain inventory data in forestry scenario models and consequential incorrect harvest decisions. *Silva Fennica*, 34, pp. 89-100.

Eiselt, H.A. and Laporte, G., 1995, Objectives in location problems. In *Facility location: a survey of applications and methods*, Z. Drezner (Ed.), pp. 151-180 (New York: Springer).

Emir-Farinas, H. and Francis, R., 2005, Demand Point Aggregation for Planar Covering Location Models. *Annals of Operations Research*, 136, pp. 175-192.

Eriksson, L.O., 2006, Planning under uncertainty at the forest level: A systems approach. *Scandinavian Journal of Forest Research*, 21, pp. 111-117.

Erkut, E., ReVelle, C. and Ülküsal, Y., 1996, Integer-friendly formulations for the r-separation problem. *European Journal of Operational Research*, 92, pp. 342-351.

Eswaran, P., Ravindran, A. and Moskowitz, H., 1989, Algorithms for nonlinear integer bicriterion problems. *Journal of Optimization Theory and Applications*, 63, pp. 261-279.

Feo, T.A., Resende, M.G.C. and Smith, S.H., 1994, A greedy randomized adaptive search procedure for maximum independent set. *Operations Research*, 42, pp. 860-878.

Fisher, P., 1999, Models of uncertainty in spatial data. In *Geographical Information Systems: Principles, Techniques, Management and Applications (Volume 1)*, P. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind (Eds.), pp. 191-205 (New York: John Wiley & Sons).

Fisher, P.F., 1991, Modelling soil map-unit inclusions by Monte Carlo simulation. *International Journal of Geographical Information Systems*, 5, pp. 193-208.

Fisher, P.F., 1992, First experiments in viewshed uncertainty: the accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing*, 57, pp. 1321-1327.

Fisher, P.F., 1998, Improved modeling of elevation error with geostatistics. *GeoInformatica*, 2, pp. 215-233.

Fleischer, M.A., 1994, Assessing the performance of the simulated annealing algorithm using information theory. Case Western Reserve University.

Fotheringham, A.S., Densham, P.J. and Curtis, A., 1995, The Zone Definition Problem in Location-Allocation Modeling. *Geographical Analysis*, 27, pp. 60-77.

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

Francis, R.L., Lowe, T.J., Rushton, G. and Rayco, M.B., 1999, A synthesis of aggregation methods for multi-facility location problems: strategies for containing error. *Geographical Analysis*, 31, pp. 67-87.

Galvao, R.D., Espejo, L.G.A. and Boffey, B., 2000, A comparison of Lagrangean and surrogate relaxations for the maximal covering location problem. *European Journal of Operational Research*, 124, pp. 377-389.

Gamarnik, D. and Goldberg, D.A., 2010, Randomized greedy algorithms for independent sets and matchings in regular graphs: exact results and finite girth corrections. *Combinatorics, Probability and Computing*, 19, pp. 61-85.

Garey, M.R. and Johnson, D.S., 1979, Computers and intractability: A guide to the theory of NP-completeness. (San Francisco: WH Freeman).

Gassmann, H.I., 1989, Optimal harvest of a forest in the presence of uncertainty. *Canadian Journal of Forest Research*, 19, pp. 1267-1274.

Gendreau, M., Soriano, P. and Salvail, L., 1993, Solving the maximum clique problem using a tabu search approach. *Annals of Operations Research*, 41, pp. 385-403.

Goldberg, D.E., 1989, *Genetic algorithms in search, optimization, and machine learning* (Boston: Addison-Wesley).

Goldberg, D.W., 2011, Advances in geocoding research and practice. *Transactions in GIS*, 15, pp. 727-733.

Goodchild, M.F., 1979, The Aggregation Problem in Location Allocation. *Geographical Analysis*, 11, pp. 240-255.

Goodchild, M.F., 1984, IL ACS: A Location-Allocation Model for Retail Site Selection. *Journal of Retailing*, 60, pp. 84-100.

Goodchild, M.F., 1989, Modeling error in objects and fields. In *Accuracy of spatial databases*, M.F. Goodchild and S. Gopal (Eds.), pp. 107-113 (London: Taylor & Francis).

Goodchild, M.F. and Noronha, V.T., 1987, Location-allocation and impulsive shopping: the case of gasoline retailing. In *Spatial analysis and location-allocation models*, A. Ghosh and G. Rushton (Eds.), pp. 121-136 (New York: Van Hostrand-Reinhold).

Goodchild, M.F. and Gopal, S., 1989, *The accuracy of spatial databases* (London: Taylor & Francis).

Goodchild, M.F., Guoqing, S. and Shiren, Y., 1992, Development and test of an error model for categorical data. *International Journal of Geographical Information Systems*, 6, pp. 87-104.

Goodchild, M.F., Shortridge, A.M. and Fohl, P., 1999, Encapsulating simulation models with geospatial data sets. In *Spatial accuracy assessment: land information uncertainty in natural resources*, K. Lowell and A. Jaton (Eds.), pp. 123–129 (Ann Arbor: Ann Arbor Press).

Goovaerts, P., 2006, Geostatistical analysis of disease data: visualization and propagation of spatial uncertainty in cancer mortality risk using Poisson kriging and p-field simulation. *International Journal of Health Geographics*, 5, p. 7.

Goycoolea, M., Murray, A.T., Barahona, F., Epstein, R. and Weintraub, A., 2005, Harvest scheduling subject to maximum area restrictions: exploring exact approaches. *Operations Research*, pp. 490-500.

Griffith, D., Haining, R. and Arbia, G., 1999, Uncertainty and error propagation in map analyses involving arithmentic and overlay operation: inverntory and prospects. In *Spatial accuracy assessment: land information uncertainty in natural resources*, K. Lowell and A. Jaton (Eds.) (Quebec City, Canada: Ann Arbor Press), pp. 11-25.

Griffith, D.A., Millones, M., Vincent, M., Johnson, D.L. and Hunt, A., 2007, Impacts of Positional Error on Spatial Regression Analysis: A Case Study of Address Locations in Syracuse, New York. *Transactions in GIS*, 11, pp. 655-679.

Grubesic, T.H., 2010, Sex offender clusters. *Applied Geography*, 30, pp. 2-18.

Grubesic, T.H. and Murray, A.T., 2008, Sex Offender Residency and Spatial Equity. *Applied Spatial Analysis and Policy*, 1, pp. 175-192.

Grubesic, T.H., Murray, A.T., Pridemore, W.A., Philip-Tabb, L., Liu, Y. and Wei, R., 2012, Alcohol beverage control, privatization and the geographic distribution of alcohol outlets. *Draft manuscript*.

Haimes, Y., Ladson, L. and Wismer, D., 1971, On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transaction on System, Man and Cybernetics*, 1, pp. 296 - 297.

Hansen, P., Kochetov, Y. and Mladenovic, N., 2004, *Lower bounds for the uncapacitated facility location problem with user preferences* (Montreal, Canada: GERAD-HEC).

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

Heuvelink, G.B.M. and Burrough, P.A., 2002, Developments in statistical approaches to spatial uncertainty and its propagation. *International Journal of Geographical Information Science*, 16, pp. 111-113.

Heuvelink, G.B.M. and Brus, D.J., 2009, Monte Carlo and spatial sampling effects in regional uncertainty propagation analyses. In *StatGIS 2009, Geoinformatics for Environmental Surveillance* (Milos Island, Greece).

Heuvelink, G.B.M., Burrough, P. and Stein, A., 1989, Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems*, 3, pp. 303-322.

Heuvelink, G.B.M., Brown, J. and Van Loon, E., 2007, A probabilistic framework for representing and simulating uncertain environmental variables. *International Journal of Geographical Information Science*, 21, pp. 497-513.

Heuvelink, G.B.M., Jiang, Z., De Bruin, S. and Twenhöfel, C.J.W., 2010, Optimization of mobile radioactivity monitoring networks. *International Journal of Geographical Information Science,* 24, pp. 365-382.

Hifi, M., 1997, A genetic algorithm-based heuristic for solving the weighted maximum independent set and some equivalent problems. *Journal of the Operational Research Society*, 48, pp. 612-622.

Hochbaum, D.S. and Pathria, A., 1997, Forest harvesting and minimum cuts: a new approach to handling spatial constraints. *Forest Science,* 43, pp. 544-554.

Hodder, J.E. and Dincer, M.C., 1986, A multifactor model for international plant location and financing under uncertainty. *Computers & Operations Research*, 13, pp. 601-609.

Hodgson, M.J., 1991, Stability of solutions to the p-median problem under induced data error. *INFOR*, 29, pp. 167-183.

Hof, J.G. and Pickens, J.B., 1991, Chance-constrained and chance-maximizing mathematical programs in renewable resource management. *Forest Science*, 37, pp. 308-325.

Hof, J.G., Bevers, M. and Pickens, J.B., 1996, Chance-constrained optimization with spatially autocorrelated forest yields. *Forest Science,* 42, pp. 118-123.

Hoganson, H.M. and Rose, D.W., 1984, A simulation approach for optimal timber management scheduling. *Forest Science*, 30, pp. 220-238.

Hoganson, H.M. and Rose, D.W., 1987, A model for recognizing forestwide risk in timber management scheduling. *Forest Science,* 33, pp. 268-282.

Holland, J.H., 1975, *Adaptation in natural and artificial systems* (Ann Arbor, MI: University of Michigan press).

Huang, B., Fery, P., Xue, L. and Wang, Y., 2008, Seeking the Pareto front for multiobjective spatial optimization problems. *International Journal of Geographical Information Science*, 22, pp. 507-526.

Hunter, G.J. and Goodchild, M.F., 1993, Managing uncertainty in spatial databases: Putting theory into practice. *Journal of the Urban and Regional Information Systems Association*, 5, pp. 55-62.

Johnson, D.S., Aragon, C.R., McGeoch, L.A. and Schevon, C., 1989, Optimization by Simulated Annealing: An Experimental Evaluation; Part I, Graph Partitioning. *Operations Research*, 37, pp. 865-892.

Jones, D., Mirrazavi, S. and Tamiz, M., 2002, Multi-objective meta-heuristics: An overview of the current state-of-the-art. *European Journal of Operational Research*, 137, pp. 1-9.

Karp, R.M., 1977, Probabilistic Analysis of Partitioning Algorithms for the Traveling-Salesman Problem in the Plane. *Mathematics of Operations Research*, 2, pp. 209-224.

Karuppiah, R. and Grossmann, I.E., 2008, Global optimization of multiscenario mixed integer nonlinear programming models arising in the synthesis of integrated water networks under uncertainty. *Computers & Chemical Engineering*, 32, pp. 145-160.

Keefer, B.J., Smith, J.L. and Gregoire, T.G., 1988, Simulating manual digitizing error with statistical models. In *GIS/LIS'88 Proceedings* (Falls Church: ASPRS/ACSM), pp. 475-483.

Kim, K., Murray, A.T. and Xiao, N., 2008, A multiobjective evolutionary algorithm for surveillance sensor placement. *Environment and Planning B: Planning and Design*, 35, pp. 935-948.

Klenner, W., Kurz, W. and Beukema, S., 2000, Habitat patterns in forested landscapes: management practices and the uncertainty associated with natural disturbances. *Computers and Electronics in Agriculture*, 27, pp. 243-262.

Klibi, W., Lasalle, F., Martel, A. and Ichoua, S., 2010, The Stochastic Multiperiod Location Transportation Problem. *Transportation Science*, 44, pp. 221-237.

Konak, A., Coit, D.W. and Smith, A.E., 2006, Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety*, 91, pp. 992-1007.

Kuby, M., Bielicki, J.M. and Middleton, R.S., 2011, Optimal Spatial Deployment of CO2 Capture and Storage Given a Price on Carbon. *International Regional Science Review*, 34, pp. 285-305.

Kyriakidis, P.C., Shortridge, A.M. and Goodchild, M.F., 1999, Geostatistics for conflation and accuracy assessment of digital elevation models. *International Journal of Geographical Information Science*, 13, pp. 677-708.

Larson, R.C. and Stevenson, K.A., 1972, On insensitivities in urban redistricting and facility location. *Operations Research*, 20, pp. 595-612.

Lawler, E.L., Lenstra, J.K., Kan, A.H.G.R. and Institute, E.U.E., 1980, Generating all maximal independent sets: NP-hardness and polynomial-time algorithms. *SIAM Journal on Computing*, 9, pp. 558-565.

Leung, S.C.H., Tsang, S.O.S., Ng, W.L. and Wu, Y., 2007, A robust optimization model for multi-site production planning problem in an uncertain environment. *European Journal of Operational Research*, 181, pp. 224-238.

Leung, Y., Ma, J.H. and Goodchild, M.F., 2004a, A general framework for error analysis in measurement-based GIS Part 1: The basic measurement-error model and related concepts. *Journal of Geographical Systems*, 6, pp. 325-354.

Leung, Y., Ma, J.H. and Goodchild, M.F., 2004b, A general framework for error analysis in measurement-based GIS Part 2: The algebra-based probability model for point-in-polygon analysis. *Journal of Geographical Systems*, 6, pp. 355-379.

Leung, Y., Ma, J.H. and Goodchild, M.F., 2004c, A general framework for error analysis in measurement-based GIS Part 3: Error analysis in intersections and overlays. *Journal of Geographical Systems*, 6, pp. 381-402.

Leung, Y., Ma, J.H. and Goodchild, M.F., 2004d, A general framework for error analysis in measurement-based GIS Part 4: Error analysis in length and area measurements. *Journal of Geographical Systems*, 6, pp. 403-428.

Leyk, S., Boesch, R. and Weibel, R., 2005, A conceptual framework for uncertainty investigation in map-based land cover change modelling. *Transactions in GIS*, 9, pp. 291-322.

Li, X. and Yeh, A.G.O., 2005, Integration of genetic algorithms and GIS for optimal location search. *International Journal of Geographical Information Science*, 19, pp. 581-601.

Liepins, G.E. and Vose, M.D., 1992, Characterizing crossover in genetic algorithms. *Annals of Mathematics and Artificial Intelligence*, 5, pp. 27-34.

Lilburne, L. and Tarantola, S., 2009, Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, 23, pp. 151-168.

Longley, P.A., Goodchild, M.F., Maguire, D. and Rhind, D., 2011, *Geographical information system and science (third version)* (Toronto: John Wiley & Sons).

Lowell, K. and Jaton, A., 1999, *Spatial accuracy assessment: land information uncertainty in natural resources* (Ann Arbor: Ann Arbor Press).

Ma, S., He, J., Liu, F. and Yu, Y., 2011, Land-use spatial optimization based on PSO algorithm. *Geo-Spatial Information Science*, 14, pp. 54-61.

Malizia, N., 2012, The effect of data inaccuracy on tests of space-time interaction. *Transactions in GIS*, In press.

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

Mirchandani, P.B. and Odoni, A.R., 1979, Locations of medians on stochastic networks. *Transportation Science*, 13, pp. 85-97.

Mirchandani, P.B. and Oudjit, A., 1980, Localizing 2 medians on probabilistic and deterministic tree networks. *Networks*, 10, pp. 329-350.

Moon, I.D. and Chaudhry, S.S., 1984, An analysis of network location problems with distance constraints. *Management Science*, pp. 290-307.

Murray, A.T., 1999, Spatial restrictions in harvest scheduling. *Forest Science*, 45, pp. 45-52.

Murray, A.T., 2003, Site placement uncertainty in location analysis. *Computers, Environment and Urban Systems*, 27, pp. 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, pp. 761-772.

Murray, A.T., 2007, Optimization. In *Encyclopedia of Geographic Information Science*, K.K. Kemp (Ed.), pp. 333-334 (Thousand Oaks: Sage).

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

Murray, A.T. and Church, R.L., 1995, Heuristic solution approaches to operational forest planning problems. *OR Spectrum*, 17, pp. 193-203.

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

Murray, A.T. and Church, R.L., 1997a, Facets for node packing. *European Journal of Operational Research*, 101, pp. 598-608.

Murray, A.T. and Church, R.L., 1997b, Solving the anti-covering location problem using Lagrangian relaxation. *Computers & Operations Research*, 24, pp. 127-140.

Murray, A.T. and Gottsegen, J.M., 1997, The Influence of Data Aggregation on the Stability of p Median Location Model Solutions. *Geographical Analysis*, 29, pp. 200-213.

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

Murray, A.T. and Weintraub, A., 2002, Scale and unit specification influences in harvest scheduling with maximum area restrictions. *Forest Science*, 48, pp. 779-789.

Murray, A.T. and Kim, H., 2008, Efficient identification of geographic restriction conditions in anti-covering location models using GIS. *Letters in Spatial and Resource Sciences*, 1, pp. 159-169.

Murray, A.T. and Grubesic, T.H., 2011, Spatial optimization and geographic uncertainty: Implications for sex offender management strategies. In *Community-Based Operations Research: Decision Modeling for Local Impact and Diverse Populations*, M. Johnson (Ed.), pp. 121-142 (Berlin: Springer).

Murray, A.T., Goycoolea, M. and Weintraub, A., 2004, Incorporating average and maximum area restrictions in harvest scheduling models. *Canadian Journal of Forest Research*, 34, pp. 456-464.

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

Mustaine, E., Tewksbury, R. and Stengel, K., 2006, Residential location and mobility of registered sex offenders. *American Journal of Criminal Justice*, 30, pp. 177-192.

Naesset, E., 1998, Positional accuracy of boundaries between clearcuts and mature forest stands delineated by means of aerial photointerpretation. *Canadian Journal of Forest Research*, 28, pp. 368-374.

Nemhauser, G.L. and Trotter, L.E., 1975, Vertex packings: structural properties and algorithms. *Mathematical Programming*, 8, pp. 232-248.

Nemhauser, G.L. and Sigismondi, G., 1992, A strong cutting plane/branch-and-bound algorithm for node packing. *Journal of the Operational Research Society*, 43, pp. 443-457.

Nozick, L.K. and Turnquist, M.A., 2001, Inventory, transportation, service quality and the location of distribution centers. *European Journal of Operational Research*, 129, pp. 362-371.

Orzanco, M., Lowell, K. and Fortin, M., 2004, Assessing the Spatial Uncertainty of Boundaries on Forest Maps Using an Ecological Similarity Index. In *Proceedings of the joint meeting of The 6th International Symposium On Spatial Accuracy Assessment In Natural Resources and Environmental Sciences and The 15th Annual Conference of The International Environmetrics Society*, R. McRoberts (Ed.) (Portland, Maine).

Padberg, M.W., 1973, On the facial structure of set packing polyhedra. *Mathematical Programming*, 5, pp. 199-215.

Palma, C.D. and Nelson, J.D., 2009, A robust optimization approach protected harvest scheduling decisions against uncertainty. *Canadian Journal of Forest Research*, 39, pp. 342-355.

Peter, B. and Nelson, J., 2005, Estimating harvest schedules and profitability under the risk of fire disturbance. *Canadian Journal of Forest Research*, 35, pp. 1378-1388.

Pickens, J.B. and Dress, P.E., 1988, Use of stochastic production coefficients in linear programming models: Objective function distribution, feasibility, and dual activities. *Forest Science*, 34, pp. 574-591.

Pickens, J.B., Hof, J. and Kent, B., 1991, Use of chance-constrained programming to account for stochastic variation in the *A*-matrix of large-scale linear programs: A forestry application. *Annals of Operations Research*, 31, pp. 511-526.

Pkukkala, T., 1998, Multiple risks in multi-objective forest planning: integration and importance. *Forest Ecology and Management*, 111, pp. 265-284.

Plastria, F., 2001, On the choice of aggregation points for continuous p-median problems: a case for the gravity center: A case for the gravity centre. *TOP*, 9, pp. 217-242.

Ralphs, T.K., Saltzman, M.J. and Wiecek, M.M., 2006, An improved algorithm for solving biobjective integer programs. *Annals of Operations Research*, 147, pp. 43-70.

Rardin, R.L. and Uzsoy, R., 2001, Experimental Evaluation of Heuristic Optimization Algorithms: A Tutorial. *Journal of Heuristics*, 7, pp. 261-304.

Ratick, S.J. and White, A., 1988, A risk-sharing model for locating noxious facilities. *Environment and Planning B: Planning and Design*, 15, pp. 165-179.

RCSO, 2005, Ramsey County Parcel Metadata Summary. Available online at: http://www.datafinder.org/metadata/rams_parcelpolygon.htm.

Reeves, L.H. and Haight, R.G., 2000, Timber harvest scheduling with price uncertainty using Markowitz portfolio optimization. *Annals of Operations Research*, 95, pp. 229-250.

Ribeiro, G.M. and Lorena, L.A.N., 2008, Lagrangean relaxation with clusters for point-feature cartographic label placement problems. *Computers & Operations research*, 35, pp. 2129-2140.

Ribeiro, G.M., Mauri, G. and Lorena, L., 2011, A lagrangean decomposition for the maximum independent set problem applied to map labeling. *Operational Research*, 11, pp. 229-243.

Roberts, S., Hall, G. and Calamai, P., 2011, Evolutionary Multi-objective Optimization for landscape system design. *Journal of Geographical Systems*, 13, pp. 299-326.

Rosing, K.E., 1992, An optimal method for solving the generalized multi-Weber problem. *European Journal of Operational Research*, 58, pp. 414-426.

Salema, M.I.G., Barbosa-Povoa, A.P. and Novais, A.Q., 2007, An optimization model for the design of a capacitated multi-product reverse logistics network with uncertainty. *European Journal of Operational Research*, 179, pp. 1063-1077.

SBSD, 2011, Santa Barbara County Parcel Metadata Summary. Available online at: http://www.countyofsb.org/pwd/surveyor/pm_pdfs/Parcel_Maps.shp.xml.htm.

Shi, W., Cheung, C.K. and Tong, X., 2004, Modelling error propagation in vector-based overlay analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59, pp. 47-59.

Snyder, L.V., 2006, Facility location under uncertainty: A review. *IIE Transactions*, 38, pp. 547-564.

Snyder, S. and ReVelle, C., 1997, Dynamic selection of harvests with adjacency restrictions: the share model. *Forest Science*, 43, pp. 213-222.

Solanki, R., 1991, Generating the noninferior set in mixed integer biobjective linear programs: an application to a location problem. *Computers & Operations research*, 18, pp. 1-15.

Strijk, T., Verweij, B. and Aardal, K., 2000, Algorithms for maximum independent set applied to map labelling. Utrecht University.

Syswerda, G., 1989, Uniform Crossover in Genetic Algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms*, J.D. Schaffer (Ed.) (San Mateo, CA: Morgan Kaufmann), pp. 2-8.

Thompson, E.F., Halterman, B.G., Lyon, T.J. and Miller, R.L., 1973, Integrating timber and wildlife management planning. *The Forestry Chronicle*, 49, pp. 247-250.

Tong, D., Murray, A. and Xiao, N., 2009, Heuristics in spatial analysis: A genetic algorithm for coverage maximization. *Annals of the Association of American Geographers*, 99, pp. 698-711.

Tong, D. and Murray, A.T., 2012, Spatial Optimization in Geography. *Annals of the Association of American Geographers*, 102, pp. 1290-1309.

Toregas, C. and ReVelle, C., 1972, Optimal location under time or distance constraints. *Papers in Regional Science*, 28, pp. 131-143.

Von Gadow, K., 2000, Evaluating risk in forest planning models. *Silva Fennica*, 34, pp. 181-191.

Wagner, M.R., Bhadury, J. and Peng, S., 2009, Risk management in uncapacitated facility location models with random demands. *Computers & Operations Research*, 36, pp. 1002-1011.

Weaver, J.R. and Church, R.L., 1983, Computational procedures for location problems on stochastic networks. *Transportation Science*, 17, p. 168.

Weintraub, A. and Abramovich, A., 1995, Analysis of uncertainty of future timber yields in forest management. *Forest Science*, 41, pp. 217-234.

Weintraub, A., Church, R.L., Murray, A.T. and Guignard, M., 2000, Forest management models and combinatorial algorithms: analysis of state of the art. *Annals of Operations Research*, 96, pp. 271-285.

Weintraub, A. and Vera, J., 1991, A cutting plane approach for chance constrained linear programs. *Operations Research*, pp. 776-785.

Wu, Q. and Hao, J.-K., 2011, An adaptive multistart tabu search approach to solve the maximum clique problem. *Journal of Combinatorial Optimization*.

Wu, X. and Grubesic, T.H., 2010, Identifying irregularly shaped crime hot-spots using a multiobjective evolutionary algorithm. *Journal of Geographical Systems*, 12, pp. 409-433.

Wu, X., Murray, A.T. and Xiao, N., 2011, A multiobjective evolutionary algorithm for optimizing spatial contiguity in reserve network design. *Landscape Ecology*, 26, pp. 425-437.

Xiao, N., 2008, A unified conceptual framework for geographical optimization using evolutionary algorithms. *Annals of the Association of American Geographers*, 98, pp. 795-817.

Xiao, N., Bennett, D.A. and Armstrong, M.P., 2002, Using evolutionary algorithms to generate alternatives for multiobjective site-search problems. *Environment and Planning A*, 34, pp. 639-656.

Zeller, R.E., Achabal, D.D. and Brown, L.A., 1980, Market penetration and locational confict in franchise systems. *Decision Sciences*, 11, pp. 58-80.

Zhao, Y. and Kockelman, K.M., 2002, The propagation of uncertainty through travel demand models: An exploratory analysis. *Annals of Regional Science*, 36, pp. 145-163.