USING LOCATION-ALLOCATION MODELS TO AID IN THE LOCATING OF PREVENTIVE HEALTH CARE FACILITIES FOR NEWFOUNDLAND & LABRADOR

By

© Weldon Brett Forsey

A Thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of:

Master of Science

Department of Geography
Memorial University of Newfoundland

October 2014

St. John's Newfoundland and Labrador

Abstract

The province of Newfoundland and Labrador is facing unprecedented demographic change. The population is aging at a faster rate than in any other province in Canada, and this is leading to dramatically increased costs to the province's health care system. One way to help alleviate the rising costs of health care is to promote preventive health care. Preventive health care can save lives and contribute to better quality of life by diagnosing serious medical conditions early. Unlike services for those who have urgent medical needs, preventive health services are intended primarily for healthy people who are less willing to travel long distances to access services. For this reason preventive health services, such as mammography units, require different locational decision methodology than other types of health care (Gu & McGregor, 2010).

This research provides a methodology to locate preventive health care facilities efficiently while ensuring spatial equity in distribution of services. Spatial equity refers to the locating of services for individuals equitably regardless of where they live. To achieve this, a variation is presented on the traditional maximal covering location problem that incorporates equity into location-allocation (LA) modeling. Using custom developed LA software, the variant algorithm is used to locate mammography facilities as a representative type of preventive health services for the island of Newfoundland. The solution set is compared to the locations of the current mammography program, which will show that the facilities of the province are well located. The results are compared to those of other models and shown to be the best in terms of equity in service delivery. This

study also helps demonstrate that LA models are an effective tool in public facility planning, especially when evidence-based decision making is important.

Acknowledgements

During my undergraduate and graduate studies I have been fortunate to have the support, encouragement and guidance of my supervisor, Alvin Simms. Alvin has provided me with the independence to develop my scientific capabilities and provided constructive criticism on my thesis. As well, thanks to my committee members, Roger White and Keith Storey, for their detailed review and constructive criticism of my thesis.

My friend and colleague, Conrad Power, has provided support and advice and has made a significant contribution to both my studies and my career. Robert Reid has provided support and flexibility in my professional life to allow me to pursue my education and for this, I express deep gratitude.

My wife Kelly has supported me and encouraged me in everything I do. Special thanks to Weldon and Madeline for going to bed on time, so daddy has time to work on his thesis.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1: Introduction

The demographics of the province of Newfoundland and Labrador have changed considerably over the past four decades. As the number of deaths continues to exceed the number of births, a natural decrease in population is occurring. This is coupled with a high rate of out-migration that has reduced the provincial population from a high of approximately 570,000 in 1991 to below 515,000 in 2011 (Statistics Canada, 2012). During this time, the province's median age has also been climbing significantly, from 20.9 in 1971 to 43.8 in 2011, and is projected to surpass 49 years by 2021 (Government of Newfoundland and Labrador, 2006). The population in the 65+ age group continues to grow, as presented in Figure 1.1. In fact, due to the excessive out-migration of younger people, many of whom are in their child-bearing years, the population of Newfoundland and Labrador is said to be aging faster than that of any other province or territory of Canada (Government of Newfoundland and Labrador, 2006).

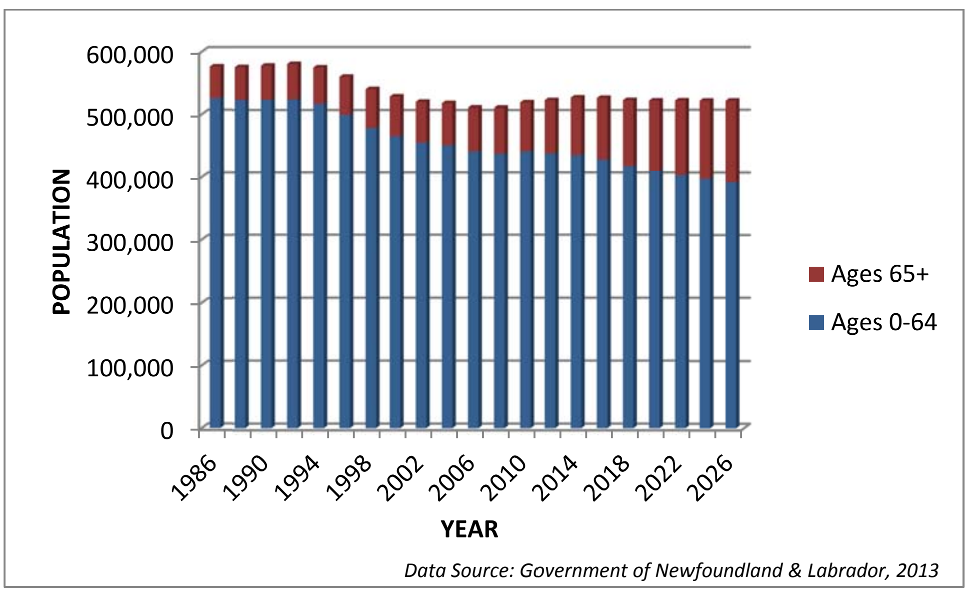


Figure 1.1: NL Population Projections to 2026f

The rapidly aging population of Newfoundland and Labrador has significant financial consequences on the health care system. As individuals advance towards their senior years, there is a substantial increase in the per capita cost of providing health care, as shown in Figure 1.2. In fact, spending for seniors accounts for 44% of all provincial government health care spending (Canadian Institute for Health Information, 2010). In 2010, health care expenditures for the Government of Newfoundland and Labrador were estimated at more than $4,766 per capita. This is more than incurred by other provincial government in Canada, and well above the national average of $3,691 per capita (Canadian Institute for Health Information, 2012). This is a strong indicator of the financial demands ahead as the population continues to grow older and live longer.

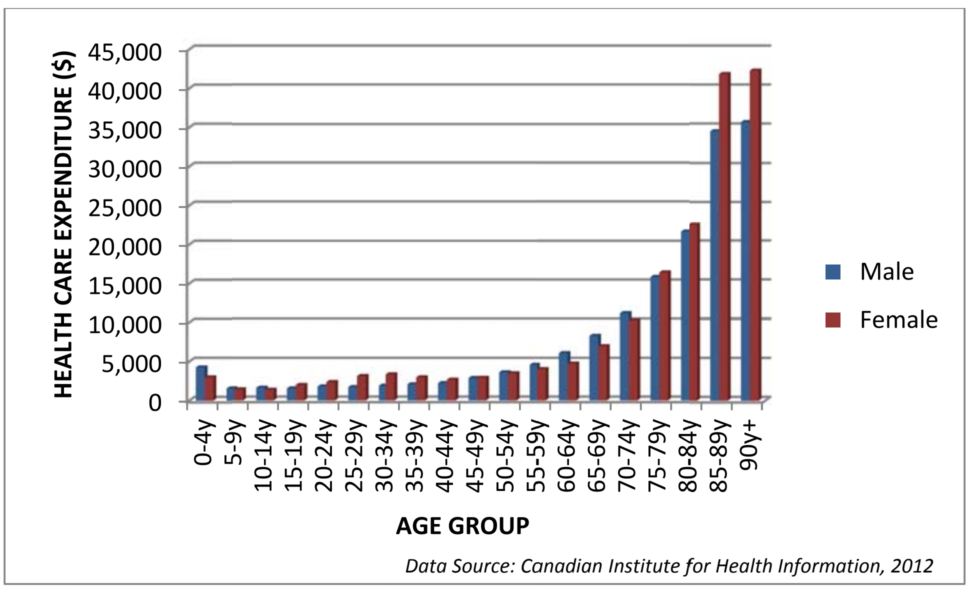


Figure 1.2: NL Health Care Expenditure per Capita by Age for 2010

The aging population combined with the associated high per capita cost of providing health care is leading to dramatic growth in health care expenditures.

Figure 1.3 indicates that trends for health care expenditures for seniors are forecast to grow from approximately $1.18 billion in 2010 to about $1.97 billion by 2026.

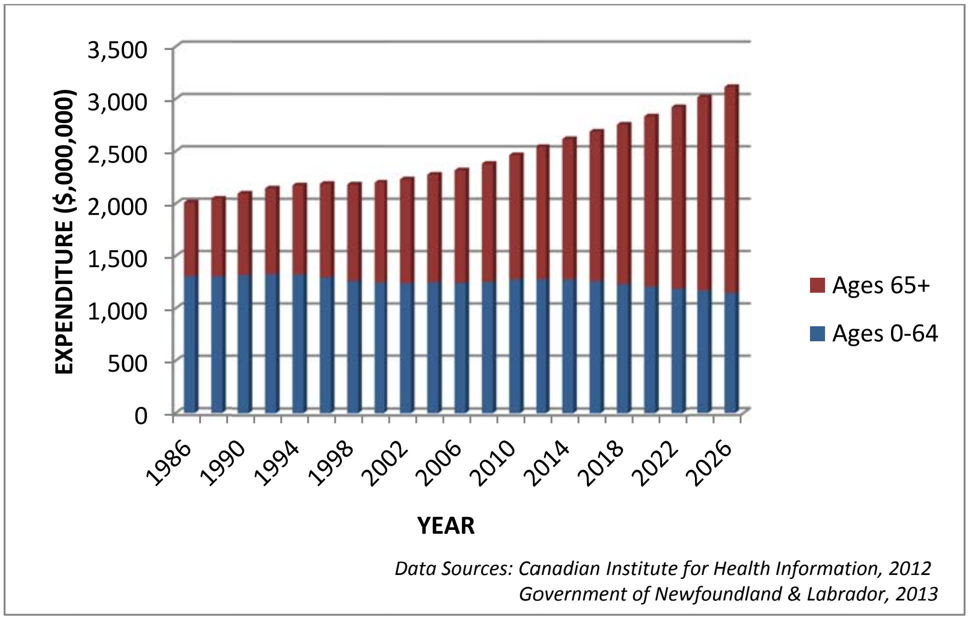


Figure 1.3: NL Health Care Expenditure by Age to 2026f, 2010 Dollars

One method of addressing the future financial burden of the province's aging population is to promote preventive health care. Preventive health care aims to reduce the likelihood of life-threatening illness and promote the early diagnosis of serious medical conditions. Commonly known preventive health services include immunizations, blood testing, and cancer screening exams. Health problems are much easier to prevent than to solve and recovery is more likely if the illness is diagnosed in an early stage (Zhang *et al.*, 2009). This also decreases overall health costs by placing greater emphasis on primary physicians for diagnosis, as opposed to higher costing specialists (Starfield *et al.*, 2005). Maximizing the number of people who receive preventive health care is one method to improve the effectiveness of the health care system (Goldsmith, 1989) and reduce its overall burden to government.

One of the key factors in maximizing the number of people who use preventive health care is accessibility. Proximity to services is a very important factor in an individual's decision to seek preventive health care (Zimmerman, 1997). Research has shown that healthy people are less willing to travel long distances for health care than those with urgent medical needs (Weiss *et al.*, 1971). Furthermore, there exists a maximum distance that a patient is willing to overcome to access a service (Farhan and Murray, 2006). Therefore, the physical location of a preventive health facility is a critical factor in patient participation (Verter & Lapierre, 2002).

One method that can help determine the optimal location of preventive health care facilities is location-allocation modeling. Location-allocation (LA) models are mathematical models used to determine optimal location of facilities based on a set of defined variables. It involves simultaneously selecting a set of locations for facilities and assigning spatially distributed demands to the facilities to maximize some measureable criterion (Rahman & Smith, 2000). The goal is to optimize the criteria specified by the objective function (Brandeau & Chiu, 1989). The objective function, or cost function, is often specified by the criterion under which the analysis will be conducted. Typical criteria for optimization include: minimizing average travel time, minimizing average response time, minimizing maximum travel time, or maximizing minimal travel time. In general, the objective function is set relative to travel distances or travel times in reference to facility-facility or facility-customer exchanges (Brandeau & Chiu, 1989).

Since the 1960s, various LA models have been proven effective and implemented in the placement of public facilities such as hospitals, schools, post offices, waste disposal sites, and public housing (Current *et al.*, 2002).

LA models can be valuable in identifying optimal locations of preventive health facilities to maximize utilization. They provide clear, defensible assessment methodologies that can be generalized using existing technologies (Messina *et al.*, 2006) and assist the decision-making process by allowing stakeholders to look at various scenarios and derive evidence-based results to help make better decisions. For example, a planner can consider how a new or existing facility will be utilized in the immediate future, as well as ten or twenty years in the future (Marianov & Serra, 2002), thus providing opportunities to consider alternative solutions and examine potential trade-offs of the various constraints. Consequently, this helps policy makers understand the implications on the demand and issues related to accessibility. In addition to being successful in locating new facilities, LA models have also been effective in measuring the efficiency of past facility decisions, and investigating alternatives to improve existing services (Rahman & Smith, 2000).

The spatial distribution of the population of Newfoundland and Labrador provides a good opportunity for the application of LA models. Outside of a high proportion of the population living within the St. John's Metropolitan Area, the province consists essentially of small rural communities scattered along an extensive coastline. The delivery of services in Newfoundland is therefore challenging due to its large geographic area and highly dispersed population. In traditional LA models the optimal location is defined in terms of efficiency, however the accepted standard of allocating public services is equity, which entails allocating services fairly, or in equal amounts to all citizens (Crompton & Lamb, 1983). Due to economies of scale, however, there are inevitably higher costs to maintain a standard level of access to services for individuals living in rural areas of the province. Efficiency and equity are not equivalent, and when faced with limited budgets service providers often seek to find a balance between the two. Therefore, to be useful for locating services such as preventive health care, traditional LA models should be modified to incorporate spatial equity. The incorporation of a spatial equity algorithm into traditional LA models for the island of Newfoundland will locate facilities efficiently while also promoting reasonable access to everyone based on some principle or standard of service, including those living in rural or remote areas of the province.

1.1 Statement of Objectives

This research presents a methodology to determine an optimal location configuration of a fixed number of facilities that maximizes accessibility to provincial preventive health facilities. A review of existing LA problem types and their associated characteristics will determine which model is best suited for the problem outlined. The selected model will be further examined and variations will be suggested for improved suitability for preventive health services. The original algorithm and proposed variants will be implemented in custom developed LA software to determine which models are appropriate for locating mammography facilities, as a representative type of preventive health service. The resulting solution sets will presented in thematic map and table formats. It is hypothesized that the variants proposed in this research will improve equity in the placement of preventive health care facilities in the province, with minimum degradation in efficiency. Furthermore, location models will be shown to be an effective tool in evidence-based decision-making.

1.2 Thesis Organization

This thesis is divided into six chapters.

1) *Introduction*

This chapter introduces the problem and provides direction on the methods that will be used to solve the problem. It will also outline the structure of the thesis.

2) *Related Research*

This chapter will review the common characteristics of LA models, these being: demand, facilities, space, and networks. This is followed by an overview of the traditional LA models and concludes with details on the various solution techniques.

3) *Applied Research*

This chapter will look at the suitability of the LA models proposed for locating preventive health services, including several variations to improve equity in facility placement. Also discussed are the methods available to integrate the new models with existing Geographic Information Systems (GIS) systems.

4) *Methodology*

This chapter outlines how each of the core components of the LA model will be represented, stored and structured for the analysis. It will also review the development of the custom LA software, including the graphical user interface and the underlying source code.

5) *Discussion and Results*

This chapter presents the results of the LA model. The results of several models will be presented to determine which model was most appropriate for the proposed problem. The results will be collected and tabulated to determine optimum placement of facilities in the study area. An extended analysis will manipulate the variables of the analysis to gain further insight into the problem.

6) *Conclusion*

This chapter presents the conclusions of the research regarding the problem of delivering preventive health services for Newfoundland and Labrador. Discussion will focus on the success of the models and role of spatial equity in facility placement. It will also outline the personal understanding and knowledge gained from the research.

Chapter 2: Related Research

This chapter reviews the various LA problem types and models to help determine which are most appropriate to examine accessibility to preventive health services. To better understand the differences in each type it is important to review the common characteristics of the traditional LA models, such as space, demand, facilities, and networks. This will also help to establish the data requirements for implementing the analysis. Once the characteristics are identified the model can be more precisely defined and formulated. The chapter concludes with a review of the solution algorithms available to solve these models.

2.1 Location-Allocation Problem Types

To solve the objective of this thesis the framework provided by Rahman and Smith (2000), shown in Figure 2.1, is useful in identifying the most appropriate problem type. LA models have been shown to be effective in both public and private sectors (Daskin, 1995), but there are philosophical and pragmatic differences that distinguish private sector from the more complex public sector location problems (ReVelle *et al.*, 1970). In the private sector, the objectives typically consider both costs and customer service (Daskin & Owen, 2002). However, the objectives in the public sector can be more difficult to identify and quantify. The goals of the public sector include social cost minimization, universality of service, efficiency, and equity (Marianov & Serra, 2002). Since these objectives are difficult to measure they are often surrogated by measures of accessibility. Typically, the goal would be to maximize accessibility by minimizing the average distance, or minimizing the greatest distance, between facility and its customers (ReVelle & Eiselt, 2005). Since this thesis is focused on delivery of preventive health care, it will explore models deemed most appropriate for public facility location problem types.

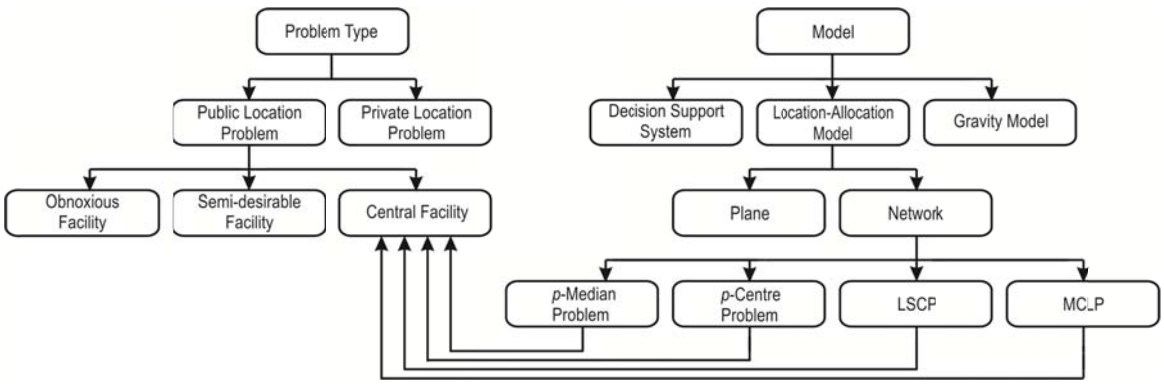


Figure 2.1: Problem Types in Public Facility LA Models
(Rahman and Smith, 2000)

Within public location problem types, Rahman and Smith (2000) offer several facility types, however preventive health facilities will be best considered a 'central facility'. This concept is derived from the notion that the spatially distributed demand will travel to a centralized facility to receive service (Rahman & Smith, 2000).

2.2 Location-Allocation Models

To solve various LA problem types, mathematical models have been developed to determine optimal locations based on identification and quantifying of realistic objectives (Church & ReVelle, 1974). This section will review several of the common traditional models that have been used for public facility applications, specifically the *p*-center, *p*-median, Location Set Covering Problem (LSCP), and the Maximal Covering Location

Problem (MCLP). Each of these models has different objectives and is suitable for different types of facilities. The purpose of this section will be to identify the most appropriate model for preventive health services.

2.2.1 *p*-Center and *p*-Median Models

Both the *p*-center and *p*-median algorithms are considered benchmarks in the development of location models. The objective of *p*-centre problems is to locate *p* servers on a general network in order to minimize the maximum distance between discrete demands and the service (Brandeau & Chiu, 1989). That is, facilities are located to minimize the maximum distance travel cost for the people who demand the good or service. A typical application of the *p*-centre problem is location of emergency service facilities, such as a fire station.

The objective of a *p*-median problem is to determine the locations of *p* facilities such that the aggregate weighted distance traveled between the demand points and the nearest facility are minimized (Chaudhry *et al.*, 1995). In other words, minimize the average distance to the nearest facility from the demand locations. This has been used to locate public facilities such as schools, as well as other types of public buildings (ReVelle *et al.,* 1970). The drawback to the *p*-median problem is that for the purposes of service delivery, some demand points may be outside a reasonable distance from the service (Rahman & Smith, 2000). This is not suitable for preventive health facilities where higher distances and time from the facility may be a deterrent to the people in the demanding locations.

2.2.2 Coverage Models

In many location problems, the level of service provided is only considered adequate if the distance between the individual and the facility is within an acceptable distance (Daskin, 1995). Coverage, or critical distance, refers to the maximum distance, or travel time, that a user is willing to overcome to access a service (Farhan & Murray, 2006). Coverage models are appropriate when the goal is to meet this critical distance or time (Daskin & Owen, 2002). A demand node is considered "covered", or "served adequately", by a facility located at another node, if the shortest path distance between the nodes is less than or equal to a specified critical distance, D_c (Daskin, 1995). Coverage models are binary in nature, that is, the demand at node x is considered satisfied at a distance less than or equal D_c, and unsatisfied at a distance greater than D_c.

2.2.2.1 Location Set Covering Problem

First developed by Toregas and ReVelle (1972), the Location Set Covering Problem (LSCP) is considered one of the simplest of facility location models. The objective of this covering problem is to find the minimum cost set of facilities among a finite set of candidates so that each demand node is covered by at least one facility (Daskin, 1995). Basically, the LCSP locates the minimum number of facilities that ensures all demand nodes are covered within a specified coverage distance. Each demand node must be serviced by at least one facility within a specified distance, D_c, from the closest facility (Marianov & Serra, 2002). The specified distance is often a proxy for a desired level of coverage, and is often stipulated in regulations or statutes. For example, finding the placement of the minimum number of fire stations within a community to ensure all residents are within a mandatory 15 minute response time. The formulation of the LSCP (Daskin & Owen, 2002) is as followed:

$$\text{Min} \sum_{j \in J} x_j \tag{2.1}$$

Subject to:

$$\sum_{j \in M} x_j \geq 1 \quad \forall i \in I \tag{2.2}$$

$$x_i \in \{0,1\} \quad \forall j \in J \tag{2.3}$$

where

I = the set of all demand nodes

J = the set of potential facility locations

M_i = the set of facility locations that cover the demand point *i*, within D_c.

D_c = the critical distance,

d_{ij} = the distance between facility *j* and demand point *i*,

$$x_j = \begin{cases} 1 & \text{if facility is sited at } j \\ 0 & \text{otherwise} \end{cases}$$

$$M_i = \{ j \mid d_{ij} \leq D_c \}$$

The objective function (2.1) minimizes the number of facilities to be located, constraint (2.2) ensures that each demand node is covered by at least one facility, and constraint (2.3) is a binary decision variable.

Although the LSCP has been effective in many legislation and planning scenarios, there are a few shortcomings that make it impractical for general public facility placements. First, it does not consider that providing coverage to all demand may be cost prohibitive. By definition the LSCP requires mandatory coverage of all demand by at least one facility; however if the demand nodes are sparsely distributed or very remote the number of facilities required may be unacceptably high. Secondly, there is no differentiation between the demand nodes that generate very little demand and those that generate a lot of demand (Daskin & Dean, 2004). For example, it is equally important to provide coverage to a community of 100 people as it is to cover a community of 10,000. In practice, when it is impossible to cover all demand nodes within the specified critical distance, it is often important to give priority to the nodes with the greater demand (ReVelle *et al.*, 2008).

2.2.2.2 Maximal Covering Location Problem (MCLP)

In reality there may not be sufficient resources (monetary and human capital) to achieve total coverage of all demand. Therefore, it may be necessary to relax the requirement of complete coverage and focus on maximizing what can be covered by a fixed number of facilities. The goal would be to locate facilities in such a manner that the fewest number of people are excluded from coverage (Verter & LaPierre, 2002).

The objective of the MCLP is to locate a predetermined number of facilities, P, in such a way to maximize the demand coverage within a specified coverage distance, D_c, from the closest facility. This approach has a condition that individuals will seek the closest facility if all the facilities provide the same quality of service (Verter & Lapierre, 2002). Assuming that there are not enough facilities to cover all demand nodes, this model aims to cover the most demand possible at a specified critical distance (Current *et al.*, 2002).

The formulation of the MCLP presented is based on the modified version by Karasakal and Karasakal (2004) as follows:

$$\text{Max} \sum_{i \in I} \sum_{j \in M_i} a_i c_{ij} x_{ij} \qquad (2.4)$$

Subject to:

$$\sum_{j \in J} y_j = P \qquad (2.5)$$

$$x_{ij} \leq y_j \quad \forall i \in I, j \in M_i \qquad (2.6)$$

$$\sum_{j \in M_i} x_{ij} \leq I \quad \forall i \in I \qquad (2.7)$$

$$y_j \in \{0,1\} \quad \forall j \in J \qquad (2.8)$$

$$x_{ij} \in \{0,1\} \quad \forall i \in I, j \in M_i \qquad (2.9)$$

where

I = the set of all demand points

J = the set of potential facilities locations

P = the number of sites to be located

a_i = the demand associated with point i

M_i = the set of facility locations that cover the demand node i

D_c = the critical distance

d_{ij} = the distance between facility j and demand node i,

$$c_{ij} = \begin{cases} 1 & \text{if } d_{ij} < D_c \\ 0 & \text{otherwise} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if facility is sited at } j \\ 0 & \text{otherwise} \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{if the demand at point } i \text{ is covered} \\ 0 & \text{otherwise} \end{cases}$$

The objective function (2.4) maximizes the coverage within the maximum critical distance, D_c. Constraint (2.5) will ensure that the total number of facilities located do not exceed the total number of facilities to be sited, *P*. Constraint (2.6) will limit x_{ij} to the facilities sited. Such that, if *j* is not sited, then all x_{ij} associated with *j* are equal to zero. Constraint (2.7) requires that all demand points may only be covered by one sited facility. If a demand point can be covered by more than one sited facility, then the facility with the maximum coverage is selected. The maximum coverage would be determined by the objective function. Constraints (2.8) and (2.9) are binary decision variables.

Unlike the LSCP, the MCLP relaxes the condition that all demand must be covered and maximizes the covered demand within a specified distance using a fixed number of facilities. It also has the ability to distinguish nodes by assessing the quantity of demand. Clearly, if it is too cost prohibitive to cover all demand, it would be beneficial to cover the nodes that generate the most demand (Daskin & Dean, 2004). The MCLP has been employed in many health care problems and is the approach utilized in this study. However, it will be later determined if this model can be improved for the purposes of preventive health care.

2.3 Characteristics of Location-Allocation Models

To understand the requirements of LA model implementation and to define data requirements it is important to consider the different components that are the basis of these models. There are four components that characterize location models (ReVelle & Eiselt, 2005) and they are:

(1) The geographic space that the demand and facilities will occupy;

(2) The demand, or the people, who seek services;

(3) The facilities that provide services; and,

(4) The conceptualized distance cost between the people and facilities.

Each component has a variety of specific aspects that distinguish between various classes of location problems.

2.3.1 Space

In LA models, space refers to the geographic space in which the demand and the facilities are to be located and is an important component in distinguishing between classes of location problems (ReVelle & Eiselt, 2005). Location models are primarily solved in two types of spaces: planar and network.

Continuous planar models types are the oldest type of location models, dating back to classical Weber (1929) problem. These models assume that demand is distributed continuously across a particular geographic space. As well, the facilities may be located anywhere throughout that space. This essentially creates an infinite number of possible facility locations and can result in unrealistic locations in the solution set, for example, locating a facility in the middle of a lake. Furthermore, due to the possibility of non-linear formulations these models can be computationally difficult to solve, especially when solving for more than one facility. In contrast, discrete planar models limit the demand to discrete points in the space. Distance is often measured between demand and facilities by a Euclidean or “straight line” distance (Daskin & Owen, 2002). Planar models are considered impractical, and should only be used to simply get a perspective of where facilities should be located and how many may be required (Daskin & Owen, 2002).

Network models assume that the location problem is embedded in an underling topology of links and nodes (*i.e.*, a transportation network). The demand is typically represented as nodes on the network which are connected by links. In a continuous network model, the facility locations can be situated anywhere on the network, including both the nodes and the links. This results in an infinite number of candidate locations for facilities. Discrete network models assume a discrete set of demand and a discrete set of candidate facility locations, both of which are limited to the nodes of the network (ReVelle *et al.*, 2008). Hakimi (1964) was the first to show that for one problem type, the *p*-median, limiting facilities only to the nodes does not degrade the quality of the solution. Regardless, limiting the solution set to network nodes is computationally advantageous, so it is often done even if the solution is possibly degraded (Daskin & Owen, 2002; Berman & Krass, 2002). Furthermore, eliminating locations that are deemed unsuitable can further reduce the candidate facility set. Discrete network models have been extensively used in the health care location problems (Daskin & Dean, 2004) and will be the model used in this research.

2.3.2 Demand

Demand in LA models refers to the individuals serviced by a facility. In the private sector demand might be potential customers, while in the public service demand may represent clients who seek a particular service. Demand may represent all individuals or a specific target group of individuals, such as particular age groups or other specific variables. For example, in the placement of a new school the demand may represent a target group of children ages 5 to 17. Discrete network models assume that demand is aggregated into a finite representative set of distinct points located on the nodes of the network. The demand assigned to each point is commonly referred to as a weight.

2.3.3 Facilities

Facilities in LA models provide services to those who would utilize them. The goal of the models is to seek the optimal location for a facility to satisfy the demand based on the problem type and criteria specified. The location of facilities is critically important for public sector facilities, such as schools, hospitals, libraries, and fire stations. A good location can provide high quality service to the community at lower costs (Daskin & Dean, 2004).

In discrete network models, the candidate set of facility locations is limited to demand nodes of the network. The candidate set will contain all the weight demand nodes or a reduced subset based on a suitability criteria. For example, communities of very low population may be deemed unsuitable for the candidate set due to anticipated lack of available workforce. Candidate sites may also only include locations that have been predetermined as options for service placement. For example, in finding optimal locations for the placement of mammography machines the candidate set could be limited to communities with a health centre. Therefore, in developing the LA system or model, it is important to provide constraints to the user to set minimum population thresholds and the ability to manually exclude locations in the candidate set. It is also beneficial to reduce the number of candidate sites to evaluate because it will reduce the computational complexity of the application.

There are scenarios where it is important to guarantee that certain candidate locations are in the solution set. These facilities are considered 'fixed' locations and are useful when attempting to add new facilities to an existing system. For example, decision-makers may seek to find the best location to add a new provincial library to augment the current library configuration that will improve overall accessibility. In these cases, it is important for location models to take into account existing facilities and solve the optimal location for the additional facilities. LA models require the ability to evaluate existing locations in the solution set.

When a facility has a restriction on the amount of demand it can service, it is considered capacitated. Alternatively, a non-capacitated facility will handle infinite demand. In a non-capacitated scenario all demand is serviced by the nearest facility, however in a capacitated scenario, the demand may not be necessarily serviced by the nearest facility if the demand exceeds service capacity. Therefore, the problem would not only be to determine the optimal number and location of facilities, but to also factor in the allocation of demand to those facilities (Daskin & Dean, 2004). Although capacitated models are relevant in many location-planning scenarios, they are much more difficult to solve (Zhou & Liu, 2003). Therefore, the models used in this research will be non-capacitated and all services locations presumed to offer equivalent services.

2.3.4 Distance

The distance between the demand and the facilities is a key component in LA modeling. Distance is the mathematical description of the concept of proximity, and distinguishes one place from another place in terms of its position from a fixed point (Plastria, 1996). In discrete planar models, distance is measured between demand and facilities using methods such as a Euclidean distance (Daskin & Owen, 2002). Using this type of model can result in impractical travel. Martin and Williams (1992) assert that straight-line measurements are not reasonable in estimating distances between patients and physicians. Realistically, people travel using existing transportation infrastructure, such as roads and highways. Therefore, network location models are more appropriate (Love & Lindquist, 1995).

In modern facility placement, with priority on vehicle travel and elaborate road networks, network models are suitable for LA modeling. Discrete network models are comprised of the roads and highways represented by the network links, while the communities are represented as nodes on the network. The connection between two nodes on a network can be solved in many ways, but generally we are interested in measuring the shortest network path distance between any pair of nodes. The Dijkstra (1959) $O(n^3)$ algorithm is commonly used to calculating the shortest network paths, which represent the minimal impedance of moving between two network nodes.

Network distance can be measured as physical or cost distances. Physical distance is often represented in units such as kilometers or miles, however calculation of distance as travel time (*i.e.*, minutes) may be a more realistic unit of measure (Kalogirou & Foley, 2006; Lovett *et al.*, 2002). Calculation of the shortest route in terms of time is computationally the same, where the physical distance is replaced by the travel time required to traverse each individual network link. In the analysis of road networks, travel time is often the quotient of distance divided by the posted speed limit. For example, the travel time required to travel a 60 km segment at the posted speed limit of 80 km/h is 0.750 hours, or 45 minutes. It is important to note the results using travel time versus road distance may not be the same. For example, traveling between the communities of St. John's and Carbonear can be completed using rural highway routes at distance of 106 km with a travel time of 122 minutes; however using the physically longer Trans-Canada Highway/Veteran's Memorial Highway route (110 km), the travel time can be reduced to 69 minutes. Furthermore, it will be assumed that all roads are in ideal travel condition and not influenced by weather, seasonal closure, or poor road conditions.

2.4 Solution Techniques

Small location problems can be solved with exact solution methods, but as problems are scaled towards realism, the number of variables and constraints can become very large. At some point, which will vary depending on the equipment, the computational resources required to solve the location problem will become unacceptable in terms of computer memory and time. Many LA models are described as NP-hard, or nondeterministic polynomial-time hard, therefore cannot be solved in polynomial time

(Garey & Johnson, 1979; Marianov & Serra, 2002). In such cases, methods have been developed to find the best "guesses" to optimal solutions. These methods are known as heuristics, or algorithms that can find very good solutions to decision problems; however they are not always guaranteed to find the optimal solution (Current *et al.*, 2002).

2.4.1 Greedy Algorithm

The simplest type of algorithms to solve location problems are the greedy heuristics. These are known as "greedy" algorithms since each step solves optimally without consideration on how the current decision affects subsequent step decisions (Fallah *et al.*, 2009). The greedy-add algorithm (Figure 2.2) starts with an empty solution set and uses a sequential approach to evaluate facilities to repeatedly select the one that yields the greatest impact on the objective function (Current *et al.*, 2002). The first facility would be chosen using total enumeration, and that location would then be considered fixed. The second facility is again chosen using total enumeration while being mindful of the location of the previous facility. This would continue until all facilities are located (Daskin & Owen, 2002).

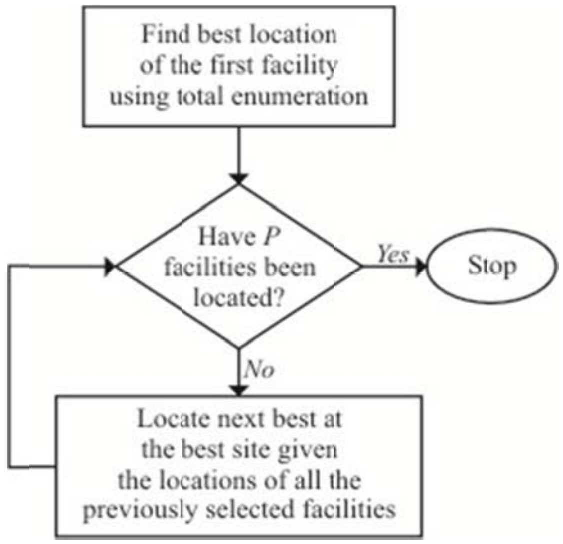


Figure 2.2: Greedy Adding Algorithm (Daskin & Owen, 2002)

Other solution techniques reviewed for this research include greedy adding with substitution, neighbourhood search, genetic algorithms, branch-and-bound, and Lagrangian relaxation. Lagrangian relaxation is particularly useful as it offers a means of evaluating the quality of the heuristics solution by providing upper and lower bounds on the objective function (Fisher, 1981). It replaces the original problem with an associated Lagrangian problem whose optimal solution provides the bounds on the objective function of the original problem (Current *et al.*, 2002). While acknowledging the potential in the aforementioned heuristics, this study concentrates on the utilization of the simpler greedy adding algorithm.

2.5 Conclusion

This chapter has presented the objectives of locating preventive health care facilities as a public centralized facility problem type. As shown in Figure 2.3, the LA model will be a discrete network model. Discrete network models assume that the demand is aggregated to a finite number of discrete network points. Likewise, the facilities to be located are limited to a finite subset of candidate network nodes. The network links that join the points will be weighted by time, as opposed to physical distance, as this is a more appropriate unit of measure in locating health care facilities (Kalogirou & Foley, 2006; Lovett *et al.*, 2002).

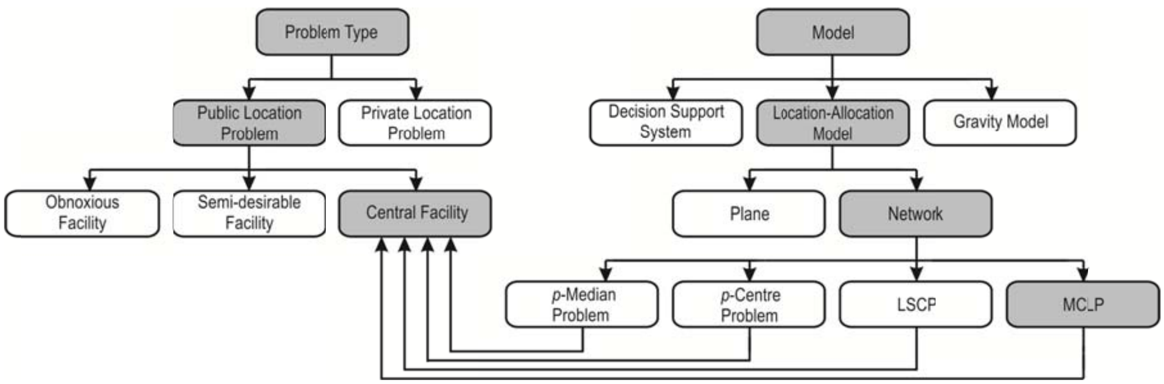


Figure 2.3: Problem Types in Public Facility LA Models
(Rahman and Smith, 2000)

In this chapter, a number of models were discussed for use with discrete network models. Both the *p*-centre and *p*-median models were shown to have limitations that made them unsuitable for locating health care services. These limitations were alleviated with the introduction of the notion of coverage, in which demand is assessed in terms of a specified critical distance to service locations. The LSCP and MCLP coverage models were shown to be appropriate in many planning situations, but the LSCP was shown to be impractical for general public facility placements. The MCLP acknowledges that in many situations, covering all demand regardless of distances may be unrealistic and would require excessive resources (Marianov & Serra, 2002). Therefore, it will be selected as the LA model for this thesis. To solve the MCLP, several solution methods were discussed that could potentially improve the solution set; however due to the forthcoming introduction of several variations of the MCLP, the simpler total enumeration method of the greedy adding algorithm will be preferred.

Chapter 3: Applied Research

In the previous chapter, the Maximal Covering Location Problem (MCLP) was identified as most appropriate for optimally locating preventive health care facilities. The MCLP is intended for applications in which the objective is to locate a predetermined number of facilities in such a way that the maximum demand is covered within a specified coverage distance. The assumptions of the MCLP have been readily accepted in many planning situations; however in some circumstances the definition of coverage has been questioned (Church and ReVelle, 1997). For example, the basic assumption of coverage models is that demand within the critical coverage distance is adequately covered and those that are outside that distance are not adequately covered. This assumes that all clients within the critical coverage distance are equally served. This chapter will demonstrate that utilization of preventive health care can be more realistically represented as a function of the distance (expressed in travel time) from the facility. Furthermore, the MCLP will be examined in terms of the balance between efficiency and equity in service delivery. This will show the need to incorporate spatial equity into the model. A supporting algorithm will also be presented. Finally, this chapter will outline the various options of incorporating LA models within standard Geographic Information System software packages and why it is necessary to develop independent LA software to optimally locate preventive health facilities on the island of Newfoundland. This software will provide the user with the means to interact with the model variables and generate solution sets from the algorithms proposed in this research.

3.1 Preventive Health Care

Many general demands for health care are based on response to acute problems, urgent patient needs, and immediate concerns. Health problems are much easier to prevent than to solve, and recovery is more likely when the illness is diagnosed in an early stage (Zhang *et al.*, 2009). Preventive health care aims to reduce the likelihood of life-threatening illness by the early diagnosis of serious medical conditions. It also has been proven to save lives and contribute to a better quality of life by reducing the need to use radical treatments, such as chemotherapy and surgery. Commonly known preventive health services include immunizations, blood testing, and cancer screening exams. In addition to the health benefits, the substantial cost savings to regional health care services through early detection and prevention of diseases, such as cancer, have been long recognized (Walker, 1977).

Preventive health services are inherently different from other health care services, because they are intended primarily for healthy people, who are often less willing to travel long distances for health services (Weiss *et al.*, 1971; Verter & Lapierre, 2002). For this reason, preventive health care facilities have a different location decision methodology that focuses on accessibility. Zimmerman (1997) showed that accessibility is a major factor in a patient's decision to have prostate cancer screening. Similarly, Maxwell (2000) concluded that a significant inverse relationship exists between travel distance and the likelihood of a patient attending a breast screening clinic. Unless services are offered at accessible locations, people are generally less likely to participate in preventive health practices.

Since preventive health care consists of a variety of services, each with slightly different service standards and intended target populations, mammography facilities are selected as a representative facility type for this study. Mammography is a well-known preventive health service for the screening of breast cancer. Studies indicate that women between the ages of 50 and 69 who receive regular mammograms are at a reduced risk of death by breast cancer (Canadian Task Force on Preventive Health Care, 2011). Mammograms are used in early detection of breast cancer and typically recommended every two years for the average woman aged 50 to 69 (Breast Cancer Society of Canada, 2013). However, the Government of Newfoundland and Labrador has recently broadened its screening program to include women aged 40 to 49 who are referred for screening by their primary health care provider (Government of Newfoundland and Labrador, 2012a). This amendment will increase the number of women eligible provincially in 2011 from 78,350 to 119,660. As depicted in Figure 3.1, this represents a significant proportion (22.3%) of the total 2011 Census population in Newfoundland and Labrador. In addition, the number of females in these age groups has increased by 17.6% from 2001, a sign of the aging population in this province. Considering that the province has the highest mastectomy rates in Canada, a focus on preventive health is timely for the province of Newfoundland and Labrador (Canadian Institute for Health Information, 2012).

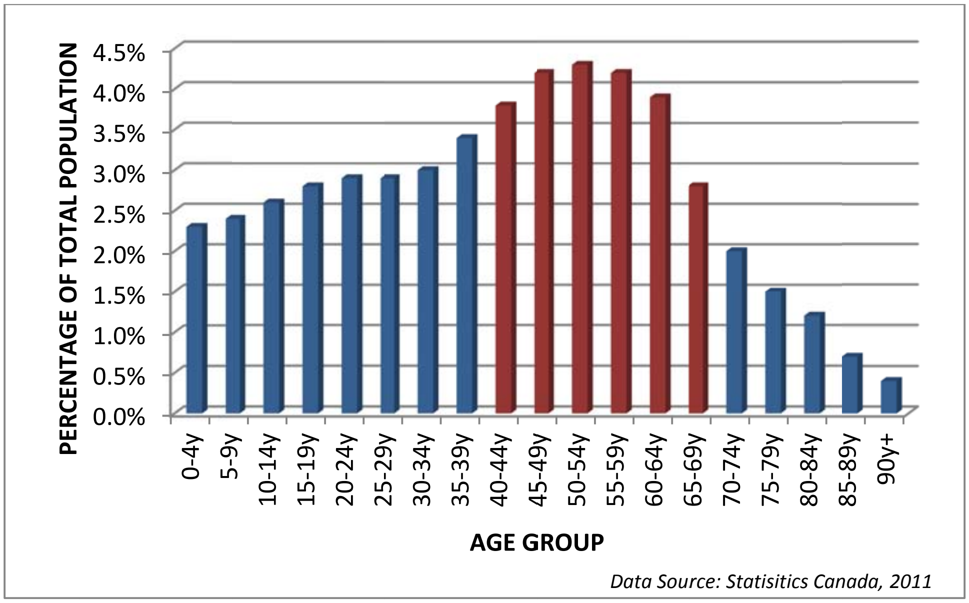


Figure 3.1: Female Population Percentage in Newfoundland and Labrador, 2011

Figure 3.2 depicts the spatial distribution of the number of the females in the target population group (ages 40-69) by Consolidated Census Subdivisions (CCS) for 2011. This map shows several regions highly populated by the target group, most notably the St. John’s area of the northeast Avalon Peninsula.

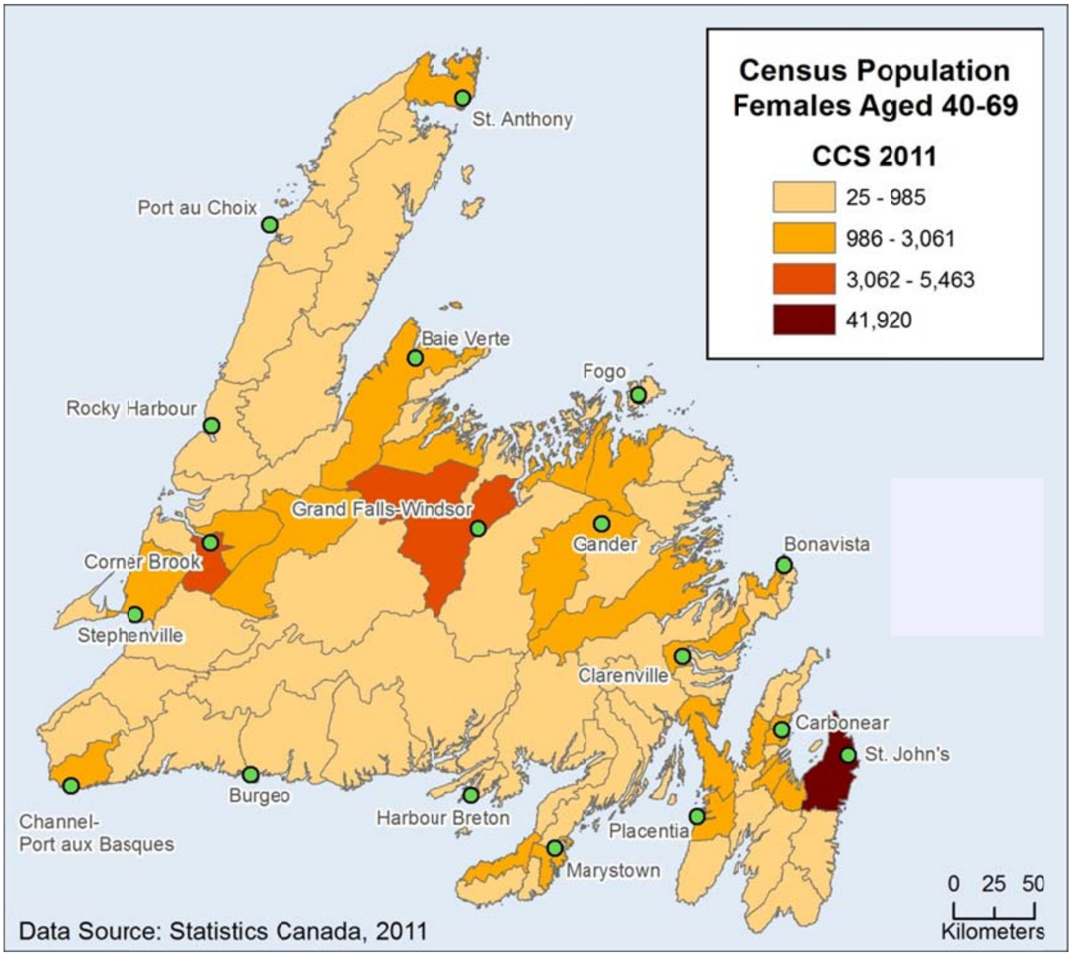


Figure 3.2: 2011 Census Population - Females Aged 40-69

The location quotient map (Fig. 3.3) shows the relative distribution of the target population group by CCS (Statistics Canada, 2011) for 2011. The dynamics of the aging trend on the Island of Newfoundland is accelerated in smaller communities because of out-migration and virtually no in-migration. This process has produced higher concentration of females aged 40 to 69.

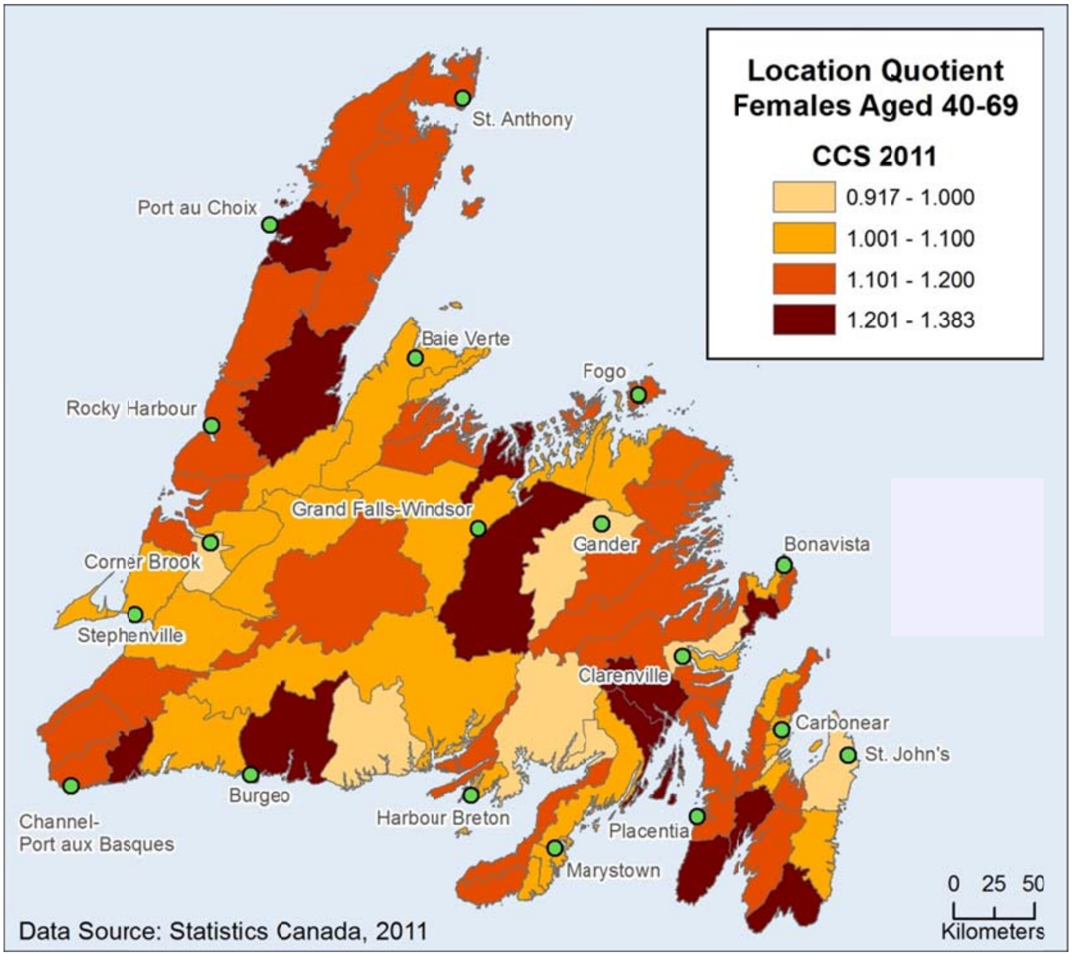


Figure 3.3: Location Quotient of Females Aged 40-69 by Consolidated Census Subdivision (CCS)

From a planning perspective, consideration must be given to future demands. In addition to the rapidly aging population described in Chapter 1, Newfoundland and Labrador is also experiencing significant regional population changes in which rural populations have declined, while urban areas have growth or remained relatively stable (Government of Newfoundland and Labrador, 2010). To effectively plan for the future, these changes should be considered to determine their impact on the results. In a more comprehensive study of preventive health services, the development of reliable projections of community-level population would be beneficial. With the current data available at the provincial level, the mammography target population of women aged 40 to 69 remains high and fairly constant in the years to come, as illustrated in Figure 3.4. This helps justify the use of the current community populations used in this study.

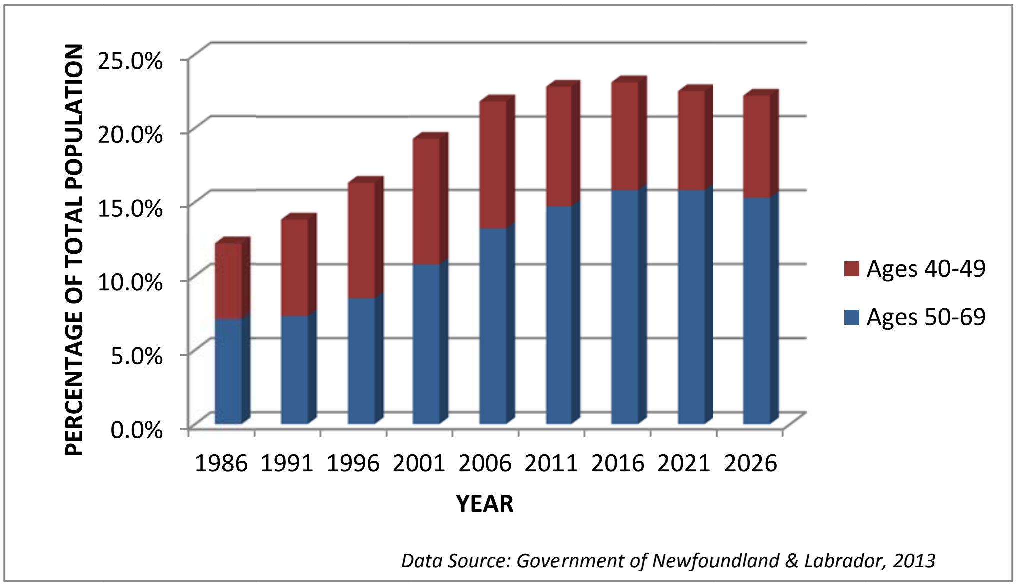


Figure 3.4: Percentage of Females Age 40-69 in Newfoundland and Labrador, 1986-2026f

3.1.1 Distance Decay

When locating preventive health services, such as mammography facilities, it has been contended that participation in these services will decline gradually according to some function of impedance, such as distance or cost. The rate at which a person's utilization of a facility declines, or likelihood of usage diminishes, is referred to as distance decay (Drezner & Eiselt, 2002). Ignoring distance decay in location modeling will result in selection of facilities that are not as close to demand as possible (Farhan & Murray, 2006). Johnston *et al.* (2000) describes distance decay as "the attenuation of a pattern or process with distance", and considers it a focal concept in various spatial models. Many classical models of spatial structure, such as the works of Christaller (1966) and Losch (1954), postulate a distance decay effect which is capable of a series of mathematical expressions. Linear, inverse power and negative exponential are some of the more commonly used functions (Longley *et al.*, 2010).

The precise nature of the function used to represent the effects of distance will vary between applications. This research will implement a linear distance decay function, so that maximizing participation and minimizing average travel distance are equivalent (ReVelle *et al.*, 1975; Holmes *et al.*, 1972). It is possible that participation is, in fact, a non-linear function of distance. With respect to preventive health care facilities, however, it has been argued by Verter and Lapierre (2002) that there are no empirical studies that establish a specific non-linear form. An area of future research would be a thorough sensitivity analysis of different distance decay functions on the solution set.

As distance increases from the facility, the patient's likelihood of utilizing the facility decreases. At a maximum critical distance the facility will be considered too far for the patient and therefore considered inaccessible. The use of a linear distance decay function will allow the decision maker to specify maximum impedance, or distance cutoff, for service delivery, which theoretically a patient would be not willing to travel to utilize a service facility (Hurst, 1972). Realistically, the distance a patient is willing to travel will vary from person to person, however, stipulating a maximum distance can be helpful in establishing a minimum level of service as a planning objective. The term coverage can be used as a proxy for diminishing service. Demand nodes are considered covered by a facility located at some other node if the distance between the two nodes is equal or less than the specified critical distance (D_c), while demand points beyond that distance are not considered covered (Daskin & Dean, 2004). A problematic condition of the traditional MCLP was the abrupt termination of coverage at the specified critical distance, as depicted in Figure 3.5. This is an unrealistic condition in the model and it would be more appropriate if coverage declined gradually as distance from the facility increased according to a distance decay function (Berman *et al.*, 2003).

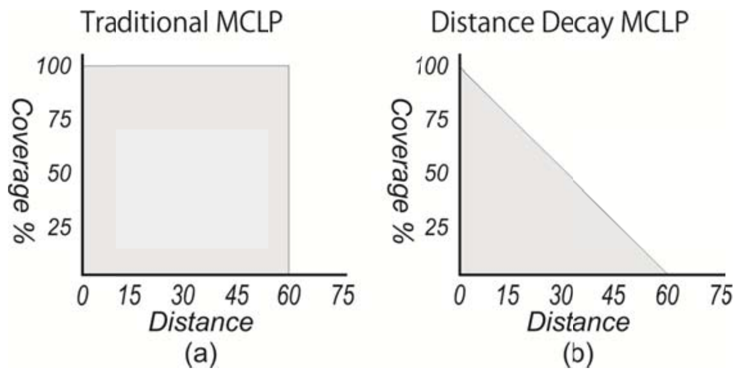


Figure 3.5: MCLP Distance Decay Functions

Similar to Karasakal and Karasakal (2004) the formalization of the MCLP with a linear distance function will be as follows:

$$\text{Max} \sum_{\text{i}\in\text{I}} \sum_{j\in J} a_i c_{ij} x_{ij} \qquad (3.1)$$

Subject to:

$$\sum_{j\in J} y_j = P \qquad (3.2)$$

$$x_{ij} \le y_j \quad \forall i \in I, j \in M \qquad (3.3)$$

$$\sum_{j\in M_i} x_{ij} \le I \quad \forall i \in I \qquad (3.4)$$

$$y_j \in \{0,1\} \quad \forall j \in J \qquad (3.5)$$

$$x_{ij} \in \{0,1\} \quad \forall i \in I, j \in M \qquad (3.6)$$

where

I = the set of all demand points,

J = the set of potential facilities locations,

P = the number of sites to be located,

a_i = the weight associated with point i,

M_i = the set of facility locations that cover the demand point i,

D_c = the critical distance,

d_{ij} = the distance between facility j and demand point i,

$$c_{ij} = \begin{cases} 1 & \text{if } d_{ij} = 0 \\ f(d_{ij}) & \text{if } d_{ij} \le D_c, (0 < f(d_{ij}) < 1) \\ 0 & \text{otherwise} \end{cases}$$

= the level of coverage provided by facility j on demand point i,

$$y_j = \begin{cases} 1 & \text{if facility is sited at } j \\ 0 & \text{otherwise} \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{if the demand at point } i \text{ is covered} \\ 0 & \text{otherwise} \end{cases}$$

The objective function (3.1) maximizes the coverage according to D_c. Constraint (3.2) will guarantee that the total number of facilities do not exceed the total number of facilities to be sited, P. Constraint (3.3) will limit x_{ij} to the facilities sited. Such that, if j is not sited, then all x_{ij} associated with j are equal to zero. Constraint (3.4) requires that all demand points may only be covered by one sited facility. If a demand point can be covered by more than one sited facility, then the facility with the maximum coverage is selected. The maximum coverage would be determined by the objective function. Constraints (3.5) and (3.6) are binary decision variables.

Since the distance function $f(d_{ij})$ is linear, the rate of coverage level is proportional to the ratio of the distance to the facility d_{ij} and the critical distance D_c. This can be expressed as:

$$c_{ij} = 1 - \left(d_{ij} / D_c\right) \tag{3.7}$$

The formulations have been developed under two assumptions (Verter & Lapierre, 2002). First, that all facilities offer equal services and each individual seeks the nearest facility for preventive services. Second, the probability of participation in a preventive health care program decreases with distance. However, there are a few exceptions to these rules that should be recognized. The first assumption may not hold if the individual is referred by a physician to a particular facility. Similarly, the individual may have a relationship to an alternative community due to work or to engage in other activities, such as shopping. The second assumption can be violated by personal issues than may influence participation, such as a family history of cancer. These exceptions to the assumptions are noted, but were not factored in the model formulation.

3.1.2 Spatial Equity

The generally accepted standard of allocating public services is equity (Crompton & Lamb, 1983). Spatial equity means to service individuals equally regardless of where they live (Bennett, 1983). In practice, the concept of equity, or fairness in respect to location, can typically be measured by an imposed minimum standard, such as a mandated critical distance (Morrill & Symons, 1977). For equity to exist, the distribution of the benefits of the service must be uniform (Bennett, 1980). Unfortunately, with limited resources this is often difficult to implement and service providers opt to emphasize efficiency to obtain the greatest yield from finite level of resources. In fact, equity and efficiency are two goals of locating public services that are often in conflict (Truelove, 1993). In the past it has been generally assumed that the efficient locations were equally distributed (Morrill & Symons, 1977). The difference which must be recognized is that an efficient location is concerned with the aggregate quantity of the service provided, whereas equity is concerned with who benefits from the service. In other words, efficiency deals with the distribution of service amongst the population and equity refers to the distribution of the effects of the service (Truelove, 1993).

For service providers, location decisions are often made in terms of minimizing operating costs, which can lead to fewer, larger facilities located in major population centres (Morrill & Symons, 1977). In terms of equity, this can be unfair, because it may result in some individuals having to travel unreasonable distances to access services. This group would have to incur more in terms of financial and psychic costs. In this case, psychic costs refer to a subset of social costs that represent added stress or losses to quality of life. In locating public services, accessibility and client costs (*e.g.,* travel expenses) must be considered and a more equitable approach would result in a more decentralized approach that promotes client accessibility (Truelove, 1993). Caution must be taken however, as too much decentralization could result in very high operating costs.

Decision makers often attempt to find a trade-off between high costs and providing services equitably.

In traditional LA models, the optimal location refers to the most efficient location. The MCLP determines the placement of an optimal facility based on maximum demand coverage within a specified distance. Generally speaking, each additional facility added to the solution set would be located to attain the highest population to maximize efficiency. As a consequence, lower density areas can be inadequately serviced and clients would incur higher costs to access facilities. To maximize equity, each additional facility placement should also consider the distance that the remaining uncovered demand must travel to access services. Therefore, the goal is to reformulate the MCLP to maximize efficiency given a set of demands that have been recurrently compensated for the distance required to access the nearest service.

This research presents a variation on the MCLP that will compensate clients for the travel needed to access services. On iteration of the solution algorithm, an equity subroutine (Figure 3.6) is executed to calculate and add additional weight to the uncovered demand as compensation for added costs of travel. Once the demand weights have been recalculated, the algorithm would proceed normally to solve for the next facility placement. This is repeated until all the facilities have been located.

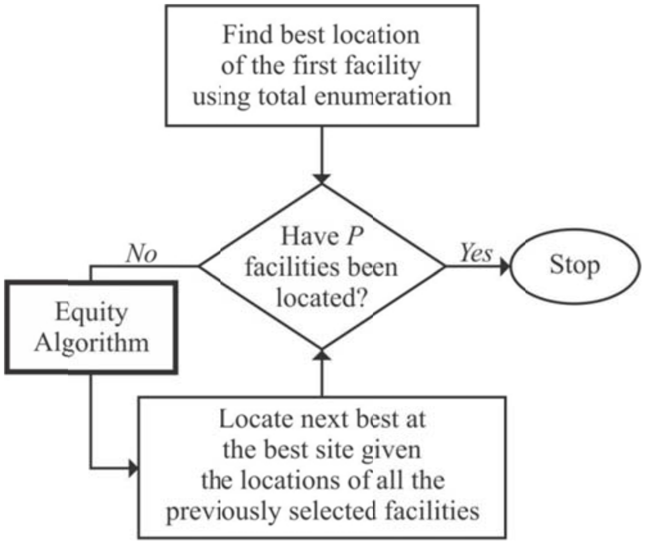


Figure 3.6: Greedy Algorithm with Spatial Equity

It is proposed that the equity algorithm will improve spatial equity in service delivery. The new weight of the demand node will be equal to the previous weight of the demand node, such as total population, multiplied by the distance (in minutes) between the demand node and the nearest facility, to the power of the equity variable. The recalculation of the demand weight variable is:

$$b_i = a_i d_{ij}^{\ e} \qquad (3.8)$$

where

b_i = the new weight associated with point i,

a_i = the previous weight associated with point i,

d_{ij} = the distance between facility j and demand point i,

e = equity variable.

By default, the equity variable is set to $e = 1$. If $e > 1$ then the compensation given for distance cost increases, contrariwise if $e < 1$ compensation will decrease. The equity variable is an important variable and requires careful assessment before the analysis.

The equity algorithm will recalculate the demand weight after each facility is located. As an example, if St. John's was the first located facility, the municipality of Corner Brook would be recalculated to the demand weight multiplied by the distance cost to St. John's. If the demand weight is the 2011 total population (19,886) and a driving time (447) in minutes, the new demand weight of Corner Brook would be $19{,}886 \times 447^1 = 8{,}889{,}042$. In the next iteration all remaining demand nodes would be recalculated using the minimum distance to any previously located facilities.

To determine if the spatial equity variant has improved equity in the delivery of services, the Schutz index (1951) will be used. The Schutz index is a method that measures the level of service in each region and compares the variation in service levels amongst regions (Truelove, 1993). The index is based on the Lorenz curve, and is described as the best simple index for measuring spatial equity by Gaile (1984). The Schutz index, S, of spatial equity is given by:

$$S = \sum_{i=1}^{n} \left| \frac{100 x_i}{\sum_{i=1}^{n} x_i} - \frac{100}{n} \right| \qquad (3.9)$$

where x_i is the measure of the benefits of the variable for the i^{th} region, and n is the number of regions.

The index is the sum of the deviation of the Lorenz curve from the diagonal, and the index will vary from 0 to 200. It will measure if the service is distributed equally amongst all regions. A value of $S = 0$ would indicate equity, and S = 200 is complete inequity. The spatial equity index will be used to compare several variations on the MCLP model and determine which variant provides the better equity. One important methodological issue with the Schutz index is the spatial scale used in measuring the equity. There is concern when using large areas that the index may appear more equitable. Truelove (1993) suggests using multiple geographic levels instead of one that might produce misleading results. For example, for the province of Newfoundland and Labrador one could compare the results on several administrative scaled geographies, such as: Economic Zones, Rural Secretariat regions, Regional Health Authorities, or Provincial Electoral Districts. The Schutz index will be used to measure equity in the discussion and results of this thesis.

3.2 Application of Location-Allocation Models

The importance of application research is debated within the field of LA modelling. In most published research, the literature has been directed towards development of new models and techniques, rather than specific applications (Current *et al.*, 2002). There are several reasons for this. First, many specific applications (*i.e.*, case studies) utilize current models and techniques, and therefore are often not viewed as scientific advances by the research community. Second, specific applications are frequently completed by planners and consultants who rarely publish in research journals. Third, advances made in the private sector can be considered proprietary and not shared with academia (Current *et al.*, 2002).

Rosing and Hodgson (1996) define two groups within the LA research field. The first group utilizes small randomly generated datasets or contrived examples to test and demonstrate new models and techniques. The second group uses real-world data to test and formulate models which have direct implications on real applications. Rosing and Hodgson (1996) suggest that increased interaction between these groups would be mutually beneficial and enhance the field of LA as a whole. To address the problem of locating preventive health care facilities for Newfoundland, this study focuses on both tasks; it will suggest a new solution technique and then utilize real-world data to study the problem.

3.3 Linking GIS to Location-Allocation

In the development of LA models, Geographic Information Systems (GIS) can play a significant role in the collection and organizing of spatial data. GIS software can maintain the attributes of the data layers which are necessary parts of these models and provide the processing tools to enable users to develop the data into appropriate formats for the modeling. For example, the Network Analyst extension of the ESRI ArcGIS® 10 provides tools to transform community and road data into an origin-destination cost matrix that is suitable for LA model purposes. In addition, GIS software is ideal for the mapping of the results from these models. However, standard GIS software packages sometimes do not offer the necessary tools to complete a specific analysis and occasionally data modelers have to develop customized applications.

3.3.1 Integration

Solving spatial problems can often require specific analytical tools that are not readily available within the standard GIS software packages. This may create the need to develop custom tools to meet the needs of the analysis and determine how they will be integrated with the GIS software. Goodchild *et al.* (1992) identified four integration strategies that can be used to integrate spatial data analysis tools with GIS: stand-alone, loose-coupling, close-coupling, and full integration.

The first method, stand-alone spatial software, ranges from a full comprehensive commercial spatial analysis package to software written to perform a single specialized piece of analysis. This method is not considered a good strategy because it doesn't take advantage of existing GIS technology. The stand-alone approach may need to recreate methods for data input, data editing, data management, and data display, which are already available in standard GIS packages (Goodchild *et al.*, 1992). This method may be preferred, for example, if the developer wanted to avoid expensive vendor costs.

A loose-coupled approach utilizes formats, such as ASCII text, exported from GIS software for use in external software (Goodchild *et al.*, 1992). This approach allows the GIS software to be used for tasks such as data development and management, while using the external software to process the exported files. If necessary, the GIS software can then be used again to present the model's results (Church, 2002). The ability to utilize the individual strengths of the separate software for the tasks in which they are most suited is one of the advantages of the loose-coupling approach (Goodchild *et al.*, 1992).

A close-coupled approach involves modifying the operations of the GIS software itself in some manner. Many commercial GIS software packages, such ArcGIS®, allow the use of macro languages or the ability to write routines in standard programming languages, such as Visual Basic or Python, to complete complex sequences of commands (Goodchild *et al.*, 1992). The routines can often be written and compiled separately while accessing the low-level data structures of the GIS by way of proprietary library functions (Anselin & Getis, 1992). This is potentially very powerful because they offer the developer access to the standard user interface, and often can appear to the user as simply extra commands. This coupling approach is a sensible option; however the possibility of limited programming languages and/or additional developer licensing costs can be a disadvantage.

The full integration approach involves completely embedding the analysis tool within the GIS software. An advantage of this approach is full support and documentation by the vendor. It also allows all users of the GIS software to have access to the newly developed analysis tool, not just those to whom access was given (Goodchild *et al.*, 1992). Some disadvantages of this method are that the vendor may require major changes or that the original developer may lose control of certain elements of the design. This method is not preferred as the developer is controlled and limited by the decisions of the vendor.

The method that best meets the needs of developing a LA model for this research is the loose-coupling approach. There are several reasons this method is preferred. Primarily, it was the aim of this research to produce an open-sourced application distributable to all users without the need for additional GIS software. This approach also allowed for development in the programming language of choice, which avoids any potential extended licensing costs associated with development within commercial GIS software.

Although the LA software will be developed independently, it will be important to ensure that the output produced can be imported back into the GIS software package. During the development stage it is necessary to ensure that the data structure of the output is consistent and usable by the external GIS software. If the output is not compatible, it may require additional programming or development of intermediate software to exchange data between the two programs.

3.3.2 Existing Location-Allocation Software

Creating independent LA software will provide the flexibility to modify, update, and adapt the application to the analysis of preventive health care. The existing options for LA analysis, such as tools found in ArcGIS® and reviewed for this research, were developed to solve for a broad range of services and lack customizability. This section will discuss the suitability of the problem types found in the ArcGIS® Network Analyst extension for the goal of locating preventive health services equitably.

The ArcGIS® 10 Network Analyst extension includes many features that have been identified in the previous chapters as important to LA modeling. Users can designate the demand weight, set the impedance variable, designate the number of facilities to be located, assign facilities as fixed or omitted, and modify many other advanced criteria. The problem types included are: *minimize impedance*, *minimize facilities*, *maximize* *coverage, maximize attendance, maximize market share,* and *target market share.* Each problem type will be reviewed individually, with the exception of the *maximize market share* and *target market share* problem types, which are intended for competitive private facility types, *e.g.*, as retail stores.

The *minimize impedance* problem type is similar to the *p*-median problem type. The goal is to locate facilities such that the sum of all weighted costs between demand points and solution facilities is minimized. It has been previously discussed that the drawback of the *p*-median problem is that for the purposes of service delivery, some demand points may be outside a reasonable distance from the service (Rahman & Smith, 2000). Therefore, this problem type is not suitable.

In the *minimize facility* problem type, facilities are located such that as many demand points as possible are allocated to the solution facilities within the impedance cutoff. Impedance cutoff is similar in nature to the term critical distance used in this research. The goal is to minimize the number of facilities to cover all demand points (ESRI, 2010). This problem type is basically the LCSP, which is more suitable for emergency services and not ideal for preventive health services.

The goal of the *maximize coverage* problem type is to locate facilities such that the greatest amount of demand is allocated to each facility within a specified impedance cutoff. All demand points within the impedance cutoff of a facility are allocated, or covered, while demand points outside the impedance cutoff are not allocated. If a demand point is within the impedance cutoff of multiple facilities, it is allocated to the nearest facility. This problem type is very similar to the MCLP and potentially useful for locating preventive health facilities.

The *maximize attendance* problem type locates facilities such that the maximum demand is allocated to each facility under the condition that the allocated demand weight will decrease in relation to the distance from the facility (ESRI, 2010). This problem type is appealing due to the use of a distance decay function, which is available in linear, power, and exponential transformations. This problem type may also be potentially suitable for locating preventive health facilities and will be compared to other models in this research.

In summary, the LA tools in the Network Analyst extension for ArcGIS® 10 offer problem types to solve a wide-range of planning scenarios for both public and private organizations. The presence of these analytical tools in large-scale commercial GIS software shows that there is a valid need for LA analysis in facility planning. With respect to locating preventive health services, such as mammography facilities, the maximize coverage and maximize attendance models have been identified as being potentially useful. Therefore, these two models will be implemented along with the MCLP variants proposed in this research and comparisons of the results will be made.

3.4 Conclusion

In this chapter, the benefits of preventive health care were discussed, as well as the need for a different LA methodology other than what is generally applied to public services. As a result, variations were proposed to the traditional MCLP to make it more applicable for preventive health care services, such as mammography facilities. Since proximity is a determining factor in an individual's decision to utilize these types of services, it was argued that accessibility to preventive health care is more realistically represented as a function of distance from the community to the nearest facility. Therefore, a diminishing linear distance decay function was implemented to more accurately represent accessibility in the traditional MCLP. It was also argued that incorporating spatial equity would improve fairness in service delivery. To accomplish this, the spatial equity variant was presented to compensate the demand nodes for the travel costs associated with accessing services.

It was argued that existing software does not offer the customizability to fully implement the LA models proposed for the study of preventive health care for Newfoundland. Therefore, there is a need to develop new, open-sourced LA software specifically for this research. This software will be developed in a loose-coupling approach to best utilize the strengths of the individual software and provide the user with the option to operate independently of existing GIS software, if required. The next chapter will show how ArcGIS® will be used to develop, manage, and visualize the data, while the custom LA software processes the data to determine optimal facility location.

Chapter 4: Methodology

The methodology for developing functional LA software for the optimal placement of mammography facilities, as a representative type of preventive health service, is described in sections 4.1 and 4.2. Section 4.1 outlines how each of the core components of LA models (demand, facilities, space, and networks) will be represented and how the respective data will be stored and structured in the software. There will also be a brief discussion on the effects of data aggregation when utilizing real-world data for LA modeling. Section 4.2 provides details on the development of the customized LA software. It outlines the development environment and explains how the software interacts with the user through the graphical user interface (GUI). This also includes the two forms of output generated by the software.

4.1 Model Characteristics

To successfully develop software for modeling preventive health care, each of the core components of LA modeling must be examined. As described in Chapter 2, these components are: space, demand, facilities, and distance (ReVelle & Eiselt, 2005). Examination of each of these components will identify the datasets required to complete the analysis. ArcGIS® 10 will be the software of choice to collect and manage the necessary datasets. The shapefile format is ideal in a loose-coupling integration approach due to the associated standard database (.dbf) formatted file for storing attribute data. More details on the precise database formatting and the loose-coupling approach are presented throughout this chapter.

4.1.1 Space

For this study, the primary research area will be the island of Newfoundland; the mainland region of Labrador will be excluded. The Labrador area offers special challenges in modeling due to the substantial distances between communities, low demand population, and issues related to health care delivery to remote regions. The delivery of services to the people of the island of Newfoundland presents its own unique geographic challenges due to coastal development patterns and areas of isolated populations. It is these challenges, however, that make the island an interesting area of study.

As previously described, space in this research will be represented in a discrete network model. This type of model will require a discrete set of demand data to be located on the nodes of a transportation network. This demand data will be also used to form the subset of candidate nodes for the facilities. More detailed information on demand, facilities, and the road network will be provided in the subsequent sections.

4.1.2 Demand

Demand in LA models refers to the individuals the facility is intended to serve. In this LA analysis, the demand is represented as communities. In this research, a community will be interpreted as a geographic area, and not a sociological or psychological concept. Communities are a suitable geographic level due to the availability of age specific population data. In addition, they can be easily associated to other data sources.

The primary source of community data is the 2011 Census of Population. The census is Canada's largest and most comprehensive data source, collecting demographic data on every individual in the country (Statistics Canada, 2011). The data collected is used by both the private and public sector to support decision-making in many areas, such as community services, forecasting consumer demand, and various other studies. Census data is disseminated in standard geographic units ranging from 13 provinces and territories down to 3,947,786 dissemination blocks. Municipalities are represented by statistical units known as census subdivisions (CSD). Many of the communities that do not meet the criteria established by Statistics Canada to be a CSD are defined as a designated place (DPL). These designated places are formed in co-operation with the provincial government and include many of the province's local service districts (LSD). The LSD is a unit of municipal government established to provide certain services to communities or areas that have similar needs within a geographic zone (Government of Newfoundland and Labrador, 2012b). Furthermore, there are smaller communities and settlements undefined by Statistics Canada. These populations are recognized by the provincial statistics agency and defined as localities (LOC). A locality is defined as a cluster of five or more dwellings (*i.e.*, a settlement), locally known by a specific name, but lacking legal limits or local government. All other populated areas not meeting any of the specified criteria may be considered as being part of an indistinct grouping known as "between communities".

The community listing and total population data for this research was provided by the Newfoundland and Labrador Statistics Agency (NLSA) (Government of

Newfoundland Labrador, 2012b). There are 544 identifiable communities; 513 of which are on the island of Newfoundland. The analysis also required the target population for mammography units. Therefore, the individual age cohorts were also acquired through the NLSA and were summated to get the total females between the ages of 40-69. The community data were imported into ArcGIS® 10 with the table design shown in Table 4.1.

Table 4.1: Population Database Table

Field	Type	Description
CommID	Integer	Community Id Number
Name	Text	Community Name
CCS	Text	Consolidated Census Subdivision
Type	Text	Community Type (ex. Town)
2011_ALL	Integer	Population 2011 – Total
2011_F4069	Integer	Population 2011 - Females Age 40-69

Using the ArcGIS® 10 software, a representative point for each community was digitized with special attention given to the positional accuracy of each location to reduce Source A error. Source A error is the result of locational information loss through the aggregation of the demand area into a single representation point (Hillsman and Rhoda, 1978). For each community in the dataset, Statistics Canada Census block population data was used in conjunction with satellite imagery from Google Maps (http://www.google.ca/maps) and topographic maps from the Atlas of Canada (http://atlas.nrcan.gc.ca/) to create approximate population-centered points. It should be acknowledged that the population center may have changed in these communities over the past 25 years due to infrastructure growth, thus the center has been chosen based on the latest data available.

4.1.3 Facilities

Facilities in LA models provide services to those who would potentially utilize them. In discrete network models, facilities would be located on the weighted nodes of the discrete network; this is in contrast with other models that allow facilities to be located anywhere on the network. In most cases, the candidate sites for facility location will be a subset of the communities of the province. One of the key features of the LA software will allow the user to automatically or manually determine which of the weighted demand nodes are to become candidate sites. The automatic method will allow specification of a minimum population threshold to exclude communities from the candidate set. This permits the exclusion of smaller communities that may not have the capacity to support a mammography facility. The manual method will allow the user to change a Boolean operator that sets a location as a candidate. This is useful when selecting locations that have been predetermined as suitable options for service placement. For example, in the analysis of the optimal locations mammography units this will allow the candidate set to be limited to current health care centres. An added benefit is that reducing the number of candidate sites to evaluate will also improve the computational speed of the application.

The LA software will also have the ability to ensure certain locations are included in the solution set. Setting fixed locations can be of particular value when locating new facilities to complement the set of existing sites. For example, in the analysis of mammography units it will be necessary to determine the optimal location of several additional units to the existing provincial breast screening program. Options will be provided within the software to select and include specific locations into the solution set. Furthermore, there will be conditional statements added to the program to ensure that the fixed location set does not surpass the number of facilities to be located.

4.1.4 Network

The LA analysis is applied to the constrained space of a transportation network for the island of Newfoundland. On this network, communities are represented by demand nodes with demand weight from the 2011 Census population. These demand nodes will also serve as candidate locations for potential facility placement in the LA model. Connecting the demand nodes will be network links representing the province's roads and ferry routes. These links will have a weight based on segment length expressed as physical distance or travel time.

The provincial road network was provided by the Newfoundland and Labrador Statistics Agency (NLSA). This road network contains all roads and ferry routes for the province. Each line segment has a number of attributes including distance and time. The time attribute represents the number of minutes it takes to transverse the line segment at the given speed limit. Each of the 513 communities used in this analysis are snapped to the nodes of this network in order to get the best results during the network analysis to create an origin-destination (OD) cost matrix.

The OD cost matrix is stored in a database which is accessible by the external LA software. The OD cost matrix is created using the ArcGIS® 10 Network Analyst extension. This extension finds and measures the least-cost paths along the network from multiple origins to multiple destinations (ESRI, 2010). The least-cost path is calculated by travel time to avoid the inclusion of routes that are geographically shorter, but may take long to traverse due to slower speed limits. The final matrix will consist of 263,169 (513^2) OD pairs than include the shortest travel time (minutes), and the respective distance (km), between each of the 513 communities. To summarize, the resulting OD cost matrix is stored in a database containing the shortest route between all possible combinations of the demand nodes.

Each community is given a unique three digit identifier ranging from 100-999. This will allow the OD identifier to be stored with a unique six digit value, where the first three digits represent the origin community and the last three digits represent the destination community. This approach simplified programming within the LA software and improved execution time. It differs slightly from the default ArcGIS® Network Analyst extension, which prefers to use a string-based identifier where the origin and destinations are separated by a hyphen. The format of the final exported table is shown in Table 4.2.

Table 4.2: Distance Matrix Database Table

Field	Type	Description
OrigDest	Integer	Origin & Destination Id
T_Time	Double	Total Time
T_Dist	Double	Total Distance

4.1.5 Data Aggregation

When utilizing real-world data for an applied LA model, it is important to consider the uncertainty introduced through data aggregation. Goodchild (1979) recognizes that LA solutions based on aggregated data are subject to error, because of the loss of locational information during the aggregation. Nevertheless, data aggregation is necessary in LA models for several reasons. First, micro-level data is inherently sensitive, so efforts are required to protect individual information. Data that can identify individuals, such as income levels and ages, are generally not available to researchers. Secondly, treating households as distinct demand points would be overwhelming with current solution techniques (Hodgson & Hewko, 2003).

In discrete network models, continuously distributed demand is aggregated to a finite number of nodes. Data aggregation into demand nodes facilitates data collection and data analysis efforts and expedites model computational times (Daskin *et al.*, 1989). However, there has been a great deal of research on the effects of spatial aggregation on the accuracy of LA modelling. When geographic areas, such as communities, are represented spatially as single points there is loss of locational information (Cromley & Mrozinski, 2002). Recognizing that data aggregation can lead to erroneous solutions is an important consideration when interpreting the results in LA modeling.

Hodgson and Hewko (2003) state that, in general, observed error increases with increased aggregation. There are three sources of error associated with demand point aggregation, classified as Source A, B, and C error (Hillsman & Rhoda, 1978). Source A errors occur when the distance between the representative point and the service facility is miscalculated due to the position of the aggregated point. For example, if the aggregated demand node is located at the geometric centre of the demand area, then the average distance between the facility and the disaggregated demand will be misestimated. Source

B error occurs when the facility is placed at an aggregated point and the distance from the facility to the demand is measured as zero. However, because the disaggregated demand is distributed throughout the polygon, the distance must in reality be greater than zero. Finally, Source C errors are a result of Source A and B error. It occurs when the aggregation causes the misallocation of demand to the incorrect facility. Potentially, the allocation of demand may lead to erroneous facility placement (Cromley & Mrozinski, 2002).

Recognizing sources of errors is necessary in LA modelling because the distance between demand points is essential to the formulation of the objection function. Goodchild (1979) states that solutions calculated with aggregated data are open to extensive manipulation, and “cast some degree of doubt on the usefulness of some LA models”. Aggregation errors are of particular concern when dealing with the traditional MCLP due to the coverage nature of the model. Potentially, positional errors may erroneously include or exclude demand nodes in the service area of a candidate facility, which could affect the selection of facilities. However, there are several ways to potentially reduce error. First, the use of the lowest level of aggregated data will reduce the impacts of aggregation error. For example, in a large urban area, data may be available for smaller suburban areas or other geographical units such as postal codes. A second method is to reduce positional error. A municipality, for example, with legislated boundaries often comprises a geographic area much larger than the actual populated area. Using the geometric centroid of the municipality could position the demand node in an area of very little population. Locating the demand node at the populated centre of the municipality would reduce positional error (Berke & Shi, 2009). The populated centre can be calculated using disaggregated data, such as postal codes or census blocks, or if sub-municipal data is unavailable it can be approximated with the use of secondary data, such as satellite imagery or topographical maps. For example, the geometric centroid for the municipality of Clarenville is approximately 7 kilometers from the core of the population, as illustrated in Figure 4.1.

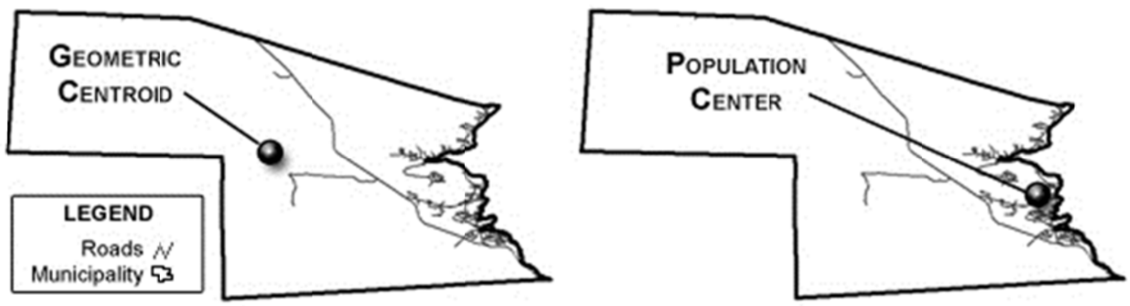


Figure 4.1: Clarenville: Geometric Centroid Vs. Population-Weighted Center

4.2 Location-Allocation Software Development

Development of the LA software was necessary to implement the MCLP variants presented in this research. The decision to create independent open-source software allowed full control over the source code. It also provided the opportunity for in-depth understanding of LA algorithms and greater appreciation for the theory behind the mathematical formulas that are the foundation of the various problem types.

4.2.1 Development Environment

There are many suitable development platforms available that can be freely used, copied, modified, and redistributed; however the platform chosen was Visual Basic within the Visual Studio Express 2012® environment. This free software has all the functionality required to complete this project. Furthermore, once the LA software is completed it permits the programmer to create an installation package to distribute the software without the need for the user to have Visual Basic or any GIS software.

The software was developed with a rapid application development (RAD) methodology that utilizes rapid prototyping over extensive planning and pseudo-coding (Maurer & Martel, 2002). This allowed the software to be developed quickly and evolve as the requirements changed during the progress of the research. The Visual Studio® environment was ideal for this approach as it offers many tools and data wizards to connect to datasets quickly and efficiently.

4.2.2 User Interface

The graphical user interface (GUI) allows users to manipulate the variables of the LA model. The goal of GUI design is to make the user's interaction as simple and efficient as possible. The design must take into account the needs, experience, and capabilities of the system user (Sommerville, 1995). As displayed in Figure 4.2, the GUI for the LA software is divided into four grouped panels: *Parameters*, *Database*, *Algorithms*, and *Output*. The *Parameters* panel contains the values that are to be set by the user. The *Database* panel allows pre-manipulation of the candidate set and access to edit the raw database. The *Algorithm* panel enables the user to select the solution algorithm, while the *Output* panel displays the results of the analysis. These grouped panels are accompanied by a “Run Model” button to execute the analysis, and a progress bar to track progress as the software executes. Each of the variables and column headers will be explained in detail in the subsequent sections.

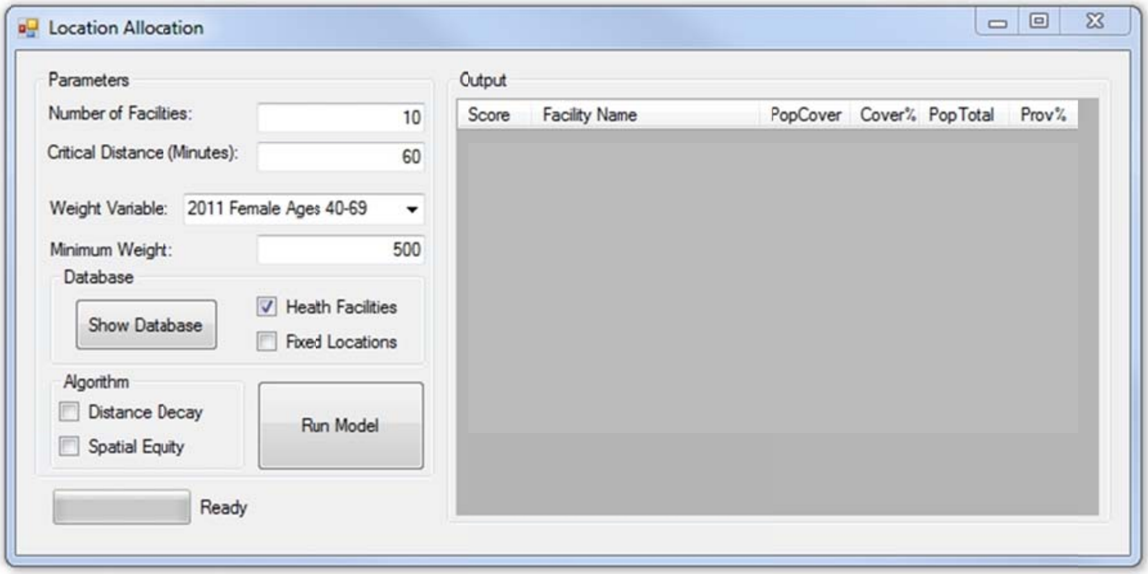


Figure 4.2: Graphic User Interface (GUI) at Initialization

4.2.2.1 Parameters

The *Parameters* panel allows the user to manipulate many of the key variables of the analysis. The first parameter is simply the number of facilities to be located. The next parameter is the critical distance; expressed in minutes of driving time. The weight variable is the weight given to the demand nodes during the analysis. By default the weight will be the 2011 Census population of the target group of females aged 40-69, however total population (both sexes) is available through the drop-down box. This is followed by the minimum weight variable which sets a population threshold for communities to be considered for the candidate set. By default this value is set to a total 2011 Census population (both sexes) of 500, but it is recommended that this value be reassessed prior to any analysis. This variable is based on total population of the community and not the weight variable selected. In a more thorough study of mammography facilities the value of this variable should be selected in consultation with stakeholders.

4.2.2.2 Database

The *Database* panel allows the user to directly access the underlying database to have more control over selection of the candidate and fixed facilities. Clicking the "Show Database" button will enlarge the interface to show the database (Figure 4.3). It contains information on the community names, the Consolidated Census Subdivision (CCS), community type, total 2011 Census population, fixed facility Boolean variable, candidate facility Boolean variable, and a coverage variable. Setting the 'Fixed' Boolean variable in the database row will ensure it is included in the solution set. Similarly, setting the 'Candidate' Boolean variable will ensure it is included in the candidate set. The software has been programmed with conditional statements to ensure the number of fixed facilities doesn't surpass the total number of facilities to be located. The final 'Covered' data column will be determined by the software, and will track the demand coverage during runtime.

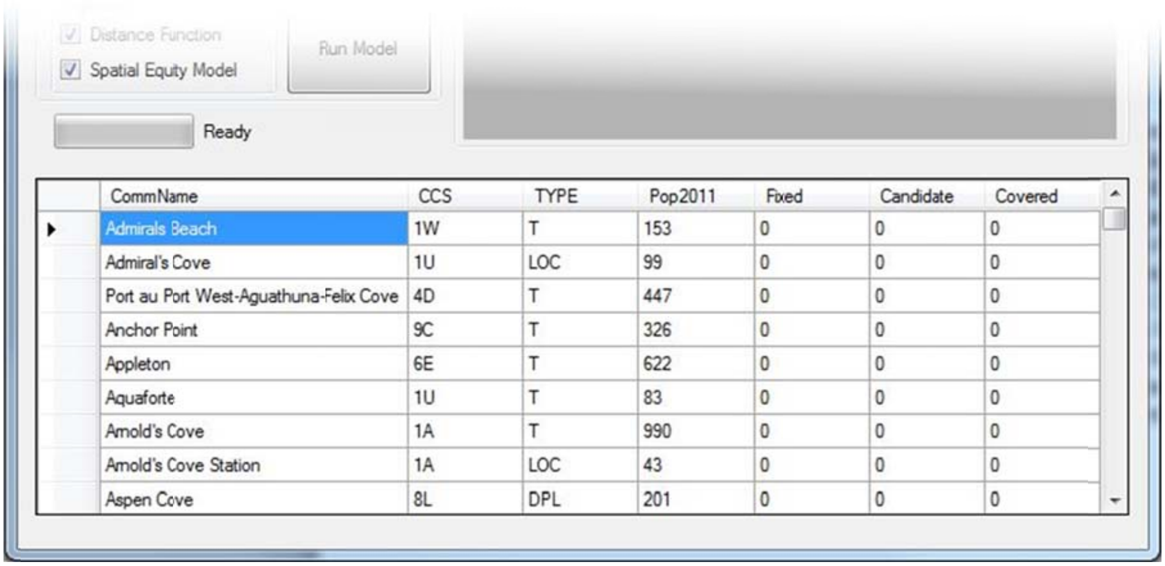


Figure 4.3: GUI – Database Editing

4.2.2.3 Algorithms

The *Algorithms* panel provides the user with the ability to select the solution algorithm. By default the software will solve the MCLP, or the user can activate the distance decay and spatial equity algorithms via checkboxes. The distance decay function will implement a linear distance decay function based on the specified critical distance. As previously outlined, this function will decrease the weight of the community in proportion to its distance from the proposed facility location. The 'Spatial Equity Model' checkbox implements the spatial equity algorithm. When selecting the spatial equity variant the distance decay function is also implemented.

4.2.2.4 Output

The output process will involve two parts. First, the results of the analysis are visible as a summary table in the *Output* panel within the LA software. The summary table presented in Figure 4.4 provides an overview of a solution set from the analysis. The table columns are:

1. **Score –** the score value formulated by the algorithm to select an optimal facility;
2. **Facility Name –** community in which a facility has been located;
3. **PopCover –** target population within the specified critical distance of the facility;
4. **Cover%** - percentage of **PopTotal** within the critical distance of the facility;
5. **PopTotal** - target population to which the facility is closest (regardless of critical distance); and,
6. **Prov%** - the percentage of total target population to which the facility is the closest.

This output table allows the user to see the results of the LA analysis in tabular format before proceeding into the visualization process.

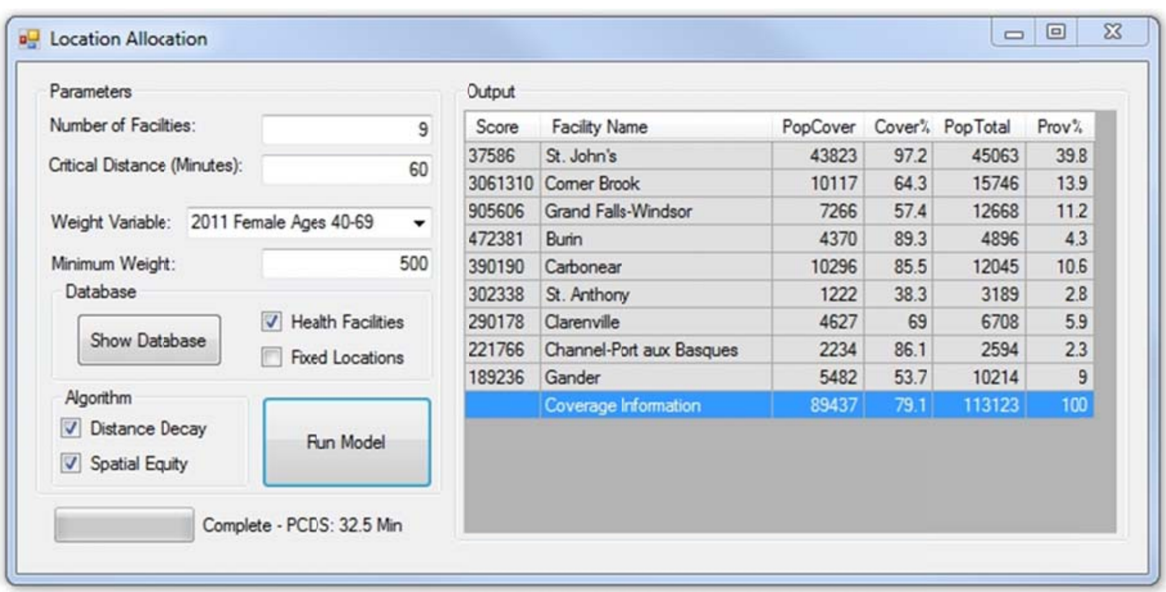


Figure 4.4: GUI – Output

Another variable calculated in the LA software is the Per Capita Distance to Services (PCDS), which is situated next to the progress bar once the analysis is completed. The PCDS value is the average per capita distance to a facility for all demand individuals. It is expressed in minutes and provides a useful comparison measure to assess improvements in global coverage. For example, a PCDS of 32.5 minutes is the average distance a client must travel to access a mammography facility on the island of Newfoundland.

The second output method is the Access® database itself. This database doesn't contain summary information from the analysis; instead it provides the specific community data. This data contains the original community information and an additional variable to indicate facility assignment for each community. This variable will be expressed in a one decimal format (#.#), where the whole digit is the facility identifier. The decimal digit will be 0, 1 or 2; where 0 indicates that a facility is located in the community; 1 indicates that the community is within the critical distance of a facility; and 2 represents a community that is outside the critical distance. For example, a classification of 2.0 signifies that the 2nd facility is located in that community; a classification as 3.2 signifies that this community has been assigned to the 3rd facility, but is outside of the specified critical distance. The purpose of this classification system is to aid the visualization process when importing the Access database into the ArcGIS® software.

4.2.3 Source Code

The source code for this research is provided in the Appendix. As discussed, the software was developed with a RAD methodology, and therefore extensive pseudo-code is not available. There is, however, documentation found within the source code to aid in the understanding of the various methods and procedures. Some of the components of the LA software were completed using visual data wizards within the Visual Studio Express® development platform and therefore not available in the provided source code.

4.3 Conclusion

The methodology for developing functional LA software to locate preventive health services, specifically mammography facilities, on the island of Newfoundland was presented in this chapter. The first section outlined each of the core components of LA models and it was determined how each would be acquired, stored, and structured for use in the LA software. This included details on utilizing the ArcGIS® Network Analyst extension to create an OD cost matrix from the provincial road network. Some of the concerns regarding the use of aggregated data in LA models and techniques used in this research to minimize error were also discussed. An objective for future research is to acquire sub-community age and sex population data to assess if Source A error can be reduced and determine if it improves the model results.

The details of the customized LA software were outlined in the second section, with specifics on the software development environment and developmental methodology. The graphical user interface (GUI) was presented in great detail with an in-depth review of all the input and output features of the software. The results of the analysis were presented in the summary table which also provides information on the facility locations and target population coverage totals. In addition to a summary table, the user can access the internal database that was manipulated during the analysis. This database provides detailed community information related to facility allocation, which can be imported to GIS software for visualization and further analysis.

Chapter 5: Discussion and Results

This chapter will discuss the results of the analysis for the optimal locating of mammography facilities on the island of Newfoundland. The analysis is conducted in three parts. The primary analysis will seek the optimal locations for the redistribution of the nine current mammography facilities. For this, the candidate location set will be limited to the 31 hospitals and health care centres on the island. These will be considered the typical facilities that would ideally support diagnostic equipment, rather than doctor's offices (Wang *et al.*, 2008). An extended analysis will further examine the mammography program by increasing the number of facilities and easing the restrictions on candidate set. Increasing the number of mammography facility locations will be used to investigate potential future expansion of the current program. Expanding the candidate set will further determine if limiting the candidate set to health care centres has hindered the objective function. The solution set to each analysis will be compared to the current real-world locations of facilities to assess the efficiency and equity implications of the current mammography program. Finally, there will be an assessment of the models discussed to determine which model produced the most equitable results.

5.1 Primary Analysis

The LA software developed for this study will be used to implement the original MCLP, distance decay variant, and spatial equity variant. The results are presented in tabular and thematic map formats. The tables will be snapshots of the actual results from within the LA software and provide detailed coverage information, while the thematic maps will help visualize the spatial distribution of the facilities. In addition to the results generated by the custom LA software, the results of the maximize coverage and maximize attendance problem types from the ArcGIS® Network Analyst extension are shown. Each of the resulting solution sets is discussed and comparisons are made to the current real-world locations.

Throughout the primary analysis, the input parameters will remain constant. This analysis considers the redistribution of the current nine mammography facilities located on the island of Newfoundland. It should be noted that several of these facilities actually have more than one mammography unit; however, this will be considered a patient capacity issue and be omitted as additional units to locate. As shown in Figure 5.1, the facilities are located in the communities of St. John's, Carbonear, Burin, Clarenville, Gander, Grand Falls-Windsor, Corner Brook, Stephenville, and St. Anthony. These nine facility locations service the 513 communities that were identified in the 2011 Census for the study area. The weight of the community demand nodes will be the target population group of females ages 40-69, as compiled from the 2011 Census. The minimum demand weight variable will be arbitrarily set at a total census (both sexes) population of 100, meaning communities of less than 100 people are not eligible to become candidates.

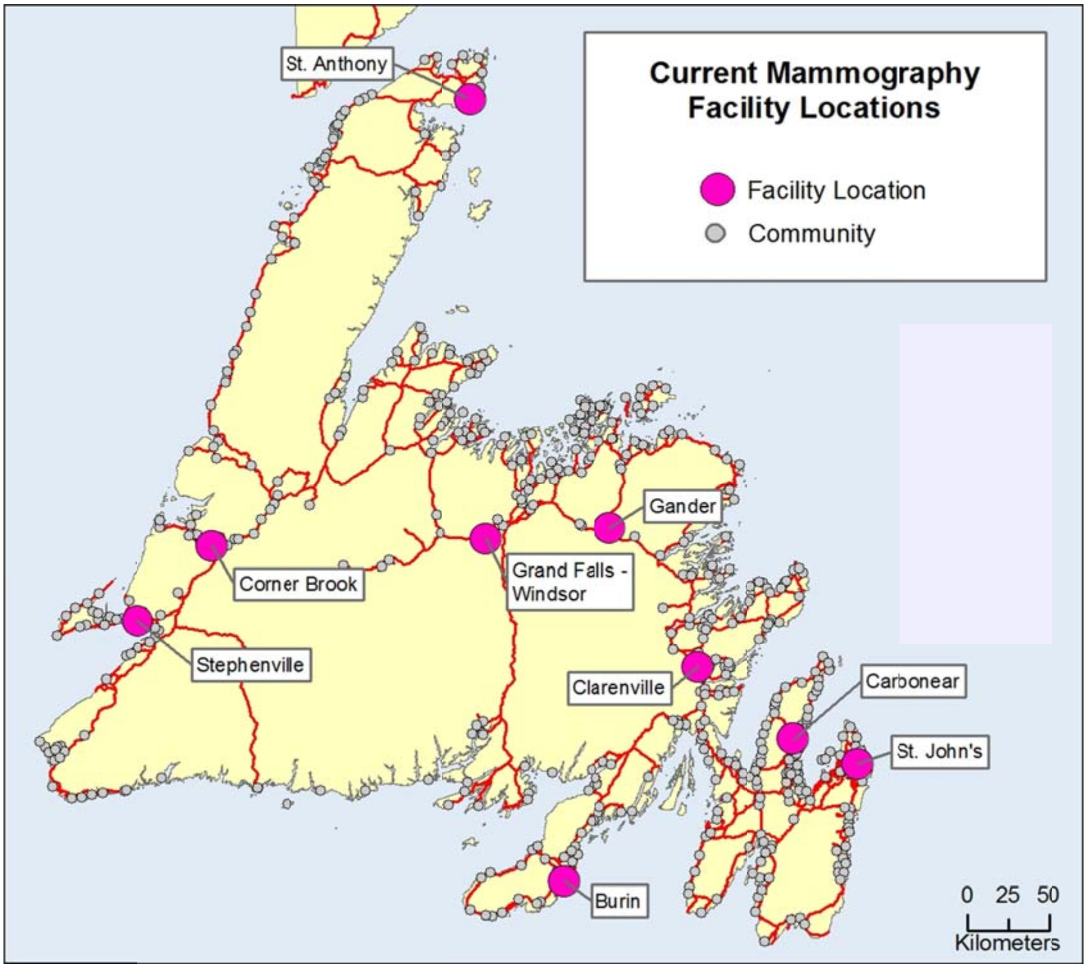


Figure 5.1: Current Mammography Facilities

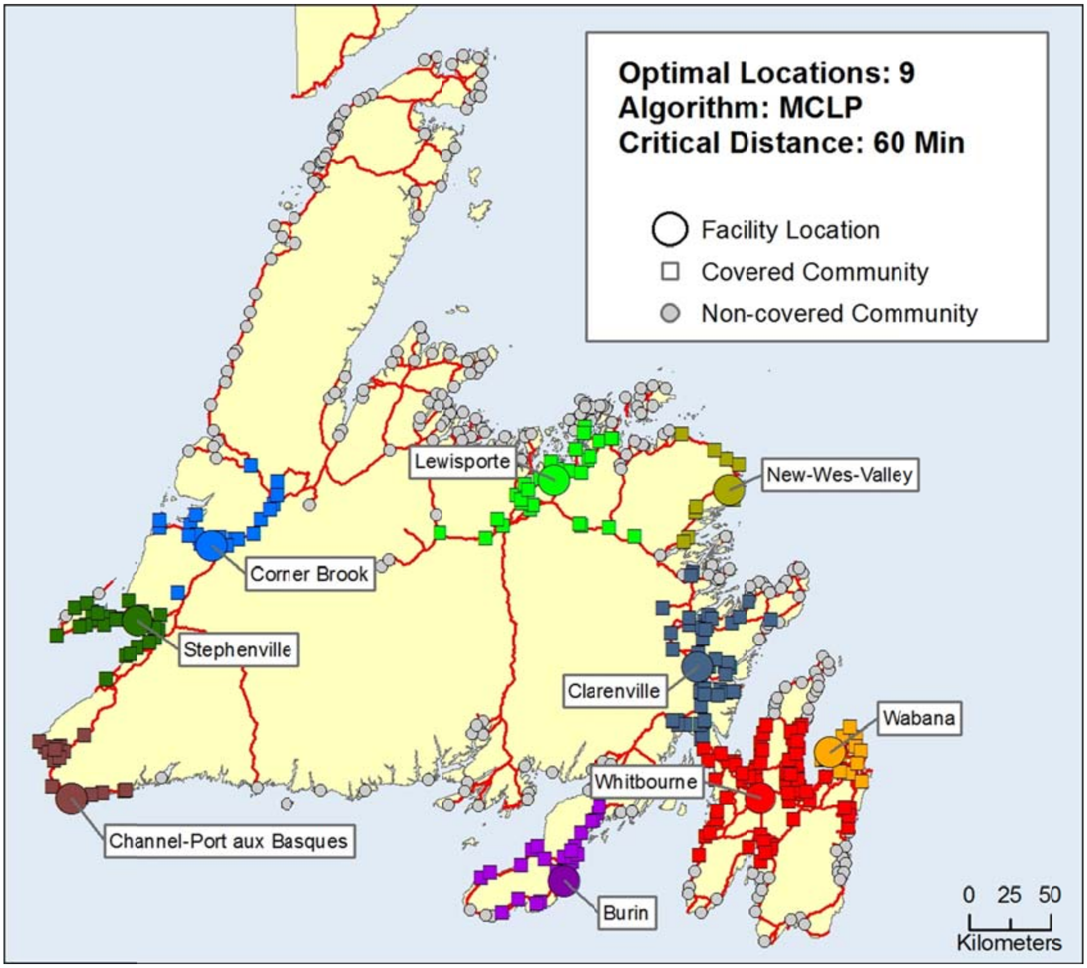
The most important variable to the solution set is the critical distance. Review of the literature failed to reveal an industry standard or consistent recommendation of a critical distance for preventive health care facilities or mammography units. A study of preventive health services by Gu *et al.* (2010) and Wang *et al.* (2008) define a critical distance of 30 minutes, a standard set by the U.S. Department of Health and Human Services for defining service areas, however, this threshold was originally developed for primary health and is possibly too stringent for non-urgent preventive health services. Using the ArcGIS Network extension it was found that the mean distance between communities and mammography services for the island of Newfoundland is 59.3 minutes. Therefore, the primary analysis of mammography facilities will proceed with a critical distance set to 60 minutes.

5.1.1 Current Mammography Locations

As a basis for comparison, the LA software was used to generate summary statistics for the current mammography program. The analysis was completed by setting the current locations of the nine real-world mammography facilities as fixed locations in the software. The results indicated that the total target female (ages 40-69) population with access to services, or covered, within 60 minutes driving time is 80.3% (90,827) of the total 113,123, while the per capita distance to service (PCDS) is 32.2 minutes. For comparative purposes between models, population coverage will be considered to represent the efficiency of the solution set, while PCDS will represent equity in distribution of services.

5.1.2 Maximum Coverage Location Problem (MCLP)

By default, the LA software runs the MCLP with the greedy algorithm. The results of this analysis are presented in Figure 5.2. The provincial population coverage was 81.4% (92,073) within the critical distance, with a PCDS of 50.8 minutes. A detailed explanation of the table column headers in Figures 5.2, 5.3, 5.4, 5.8 and 5.9 can be found in Section 4.2.2.



Score	Facility Name	PopCover	Cover%	PopTotal	Prov%
50032	Whitbourne	17993	86.1	20888	18.5
10796	Lewisporte	10796	61.6	17520	15.5
10117	Corner Brook	9225	59.6	15467	13.7
5209	Wabana	36609	100	36609	32.4
4370	Burin	4370	89.3	4896	4.3
3705	Clarenville	4344	59.1	7349	6.5
3624	Stephenville	4516	85.9	5258	4.6
2234	Channel-Port aux Basques	2234	98.9	2259	2
1986	New-Wes-Valley	1986	69	2877	2.5
	Coverage Information	92073	81.4	113123	100

Figure 5.2: LA Software – MCLP

At first glance, many of the facilities are not intuitively located where one would expect. For example, the highest scoring facility is Whitbourne which is approximately

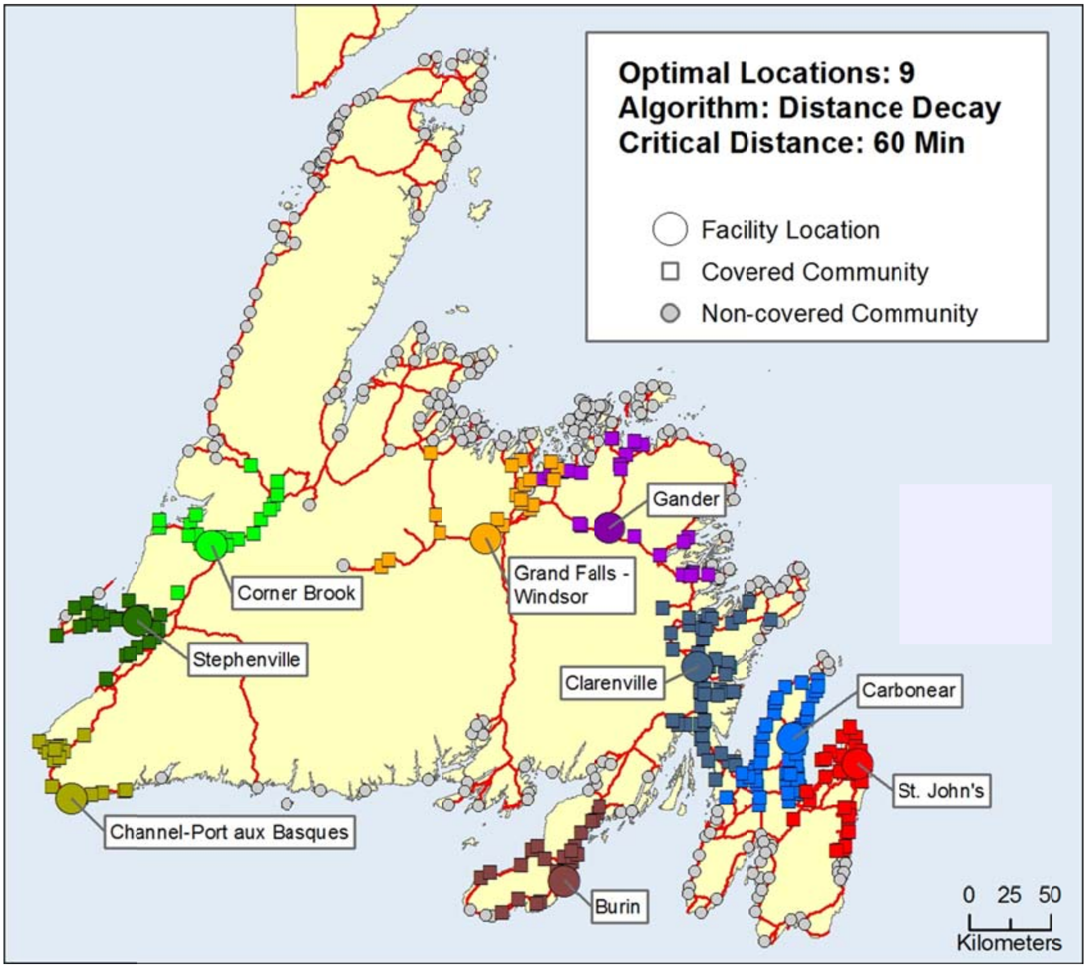
59 minutes from the large urban centre of St. John's. The objective of the MCLP, however, is to locate nine facilities in such a way that the maximum demand is covered within the specified coverage distance of 60 minutes from the closest facility. The selection of Whitbourne as an optimal location provides coverage to the St. John's area as well as the population centers of Carbonear and Placentia. Similarly, the facility located in Lewisporte can provide coverage to the large communities of Gander and Grand Falls-Windsor. Considering the objective of the MCLP algorithm, these locations are justifiable for maximizing pure benefit to the greatest amount of demand.

Another notable outcome was the placement of a facility in Wabana. This community is located on Bell Island and was the fourth facility to be located. Its selection is the result of the communities being the only candidate location remaining in the northeast Avalon Peninsula able to provide coverage to a large population that still remained uncovered from placement of the Whitbourne facility. This location would not be practical because of the ferry link, which is susceptible to barriers, such as wait times and weather delays. It is also important to mention the lack of facility placement on the Northern Peninsula. With four candidate locations in this area, neither provided enough total population coverage to warrant a facility placement. This will result in residents of this area travelling up to five hours to access mammography services, which may be a strong deterrent for usage.

5.1.3 Distance Decay

The distance decay algorithm is applied with a checkbox option in the LA software. The results of the analysis with the distance decay variant are shown in Figure 5.3. The provincial target population coverage within the critical distance of 60 minutes is 81.2% (91,839), which is a slight drop from the previous MCLP model. On the other hand the PCDS improved dramatically to 35.9 minutes, over the previous 50.8 minutes. This is a substantial increase and would be argued as adequate tradeoff for the 0.2% (234) drop in the coverage area population.

It can be argued that the total population covered variable calculated in the MCLP model is not directly comparable to the results of the distance decay model due to the attenuation of the target population that is considered covered through the use of the distance decay function. The total target population covered is still helpful in comparison of overall accessibility to preventive health services, regardless of the individual's decision to seek services. To aid in the comparison of models that incorporate the distance decay function, such as the spatial equity model discussed in the next section, the attenuated target population coverage is also calculated in a supplementary analysis. For the distance decay model the attenuated target population is calculated to be 60.9% (68,957).



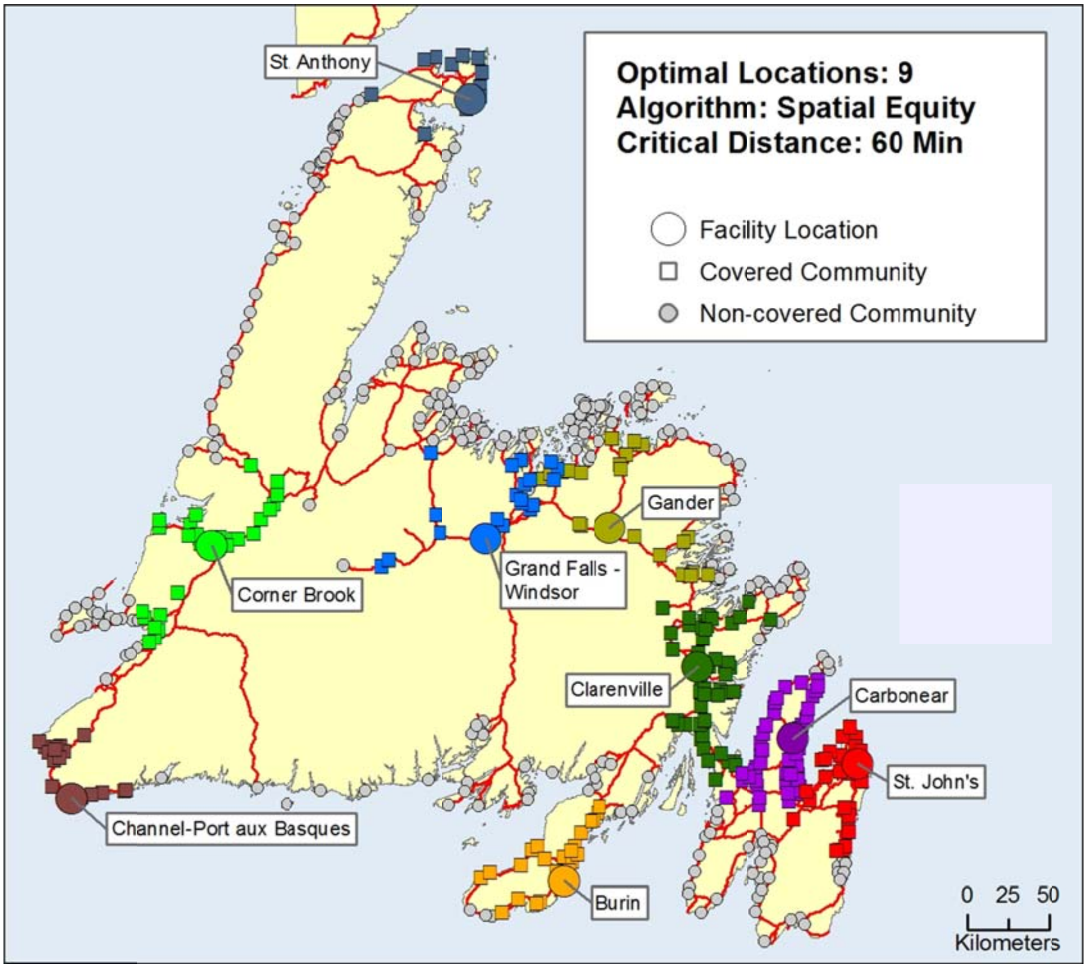
Score	Facility Name	PopCover	Cover%	PopTotal	Prov%
37586	St. John's	43823	97.2	45063	39.8
7166	Corner Brook	9225	65.8	14012	12.4
5497	Carbonear	10296	85.5	12045	10.6
5354	Grand Falls-Windsor	7266	57.4	12668	11.2
2940	Gander	5482	53.7	10214	9
2709	Clarenville	4627	69	6708	5.9
2572	Stephenville	4516	85.9	5258	4.6
2208	Burin	4370	89.3	4896	4.3
1584	Channel-Port aux Basques	2234	98.9	2259	2
	Coverage Information	91839	81.2	113123	100

Figure 5.3: LA Software – Distance Decay Variant

The distance decay analysis produced a more equitable solution set than that of the MCLP, with locations that are similar to the real-world locations of mammography facilities. The only difference was the placement of a facility in Channel-Port aux Basques versus the real-world St. Anthony location. A secondary analysis of this model would show that St. Anthony actually places behind Bonavista, Twillingate, Placentia, Springdale, Placentia, and New-Wes-Valley as an optimal facility location. With the lack of a facility situated on the Northern Peninsula it is evident that there is a need to introduce the spatial equity algorithm into the analysis.

5.1.4 Spatial Equity Variant

The results of the analysis of the spatial equity variant are displayed in Figure 5.4. The provincial target population coverage within the critical distance has decreased slightly to 79.1% (89,437); similarly the attenuated target population coverage has dropped to 59.0% (66,763). The PCDS has improved to 32.5 minutes.



Score	Facility Name	PopCover	Cover%	PopTotal	Prov%
37586	St. John's	43823	97.2	45063	39.8
3061310	Corner Brook	10117	64.3	15746	13.9
905606	Grand Falls-Windsor	7266	57.4	12668	11.2
472381	Burin	4370	89.3	4896	4.3
390190	Carbonear	10296	85.5	12045	10.6
302338	St. Anthony	1222	38.3	3189	2.8
290178	Clarenville	4627	69	6708	5.9
221766	Channel-Port aux Basques	2234	86.1	2594	2.3
189236	Gander	5482	53.7	10214	9
	Coverage Information	89437	79.1	113123	100

Figure 5.4: LA Software – Spatial Equity Variant

The spatial equity algorithm produced a solution set similar to the distance decay variant. The critical difference is the selection of the St. Anthony facility versus the previous locating of a facility in Stephenville. Also notable was the increase to the 4th ranking for the Burin facility in the solution set, which had previous ranked 8th. The inclusion of St. Anthony and the increased ranking of Burin indicate that the algorithm has successful compensated the demand population for the considerable distance required to access services.

The solution set of the spatial equity variant is also very similar to the current location of mammography units, the only difference being the selection of Channel-Port aux Basques versus the real-world Stephenville location. This will be examined further in the discussion of the primary analysis; however a supplementary analysis would indicate that Stephenville would be the 10th optimal facility, just one placement outside the nine.

There was a slight decline in total target population coverage from 81.2% (91,839), in the distance decay variant, to 79.1% (89,437); as well in the attenuated target population coverage from 60.9% (68,957) to 59.0% (66,763). This indicates that the facilities located are less efficient within the critical distance. In terms of equity however, the PCDS has improved to 32.5 minutes from 35.9 minutes. This is significant considering the PCDS is heavily influenced by the larger population centres. Follow-up analysis in ArcGIS indicates that the PCDS of the population outside of the critical distance has improved considerably; 127.4 minutes using the distance decay variant to 97.7 minutes using the spatial equity variant.

These results demonstrate an improvement in access to services in terms of equity. Specifically, individuals within the target population group will on average travel shorter distances to access mammography units. To the service provider, however, there has been a compromise in efficiency or decrease in the aggregate quantity of the service (Truelove, 1993).

5.1.5 ArcGIS® Network Analyst

The ArcGIS® Network Analyst extension offers several models for LA analysis. For the goal of locating mammography facilities for Newfoundland, two problem types were identified as potentially suitable. First, the maximize coverage problem type which attempts to locate facilities such that the greatest amount of demand is covered within the specified impedance, or distance, cutoff. Secondly, the maximize attendance problem type which incorporates a distance decay function to decrease the amount of demand allocated to a chosen facility as distance increases.

These problem types are implemented within ArcGIS® using the same criteria as the previous analyses, including a linear distance decay function. The results of these problem types are depicted in Figure 5.5 and Figure 5.6.

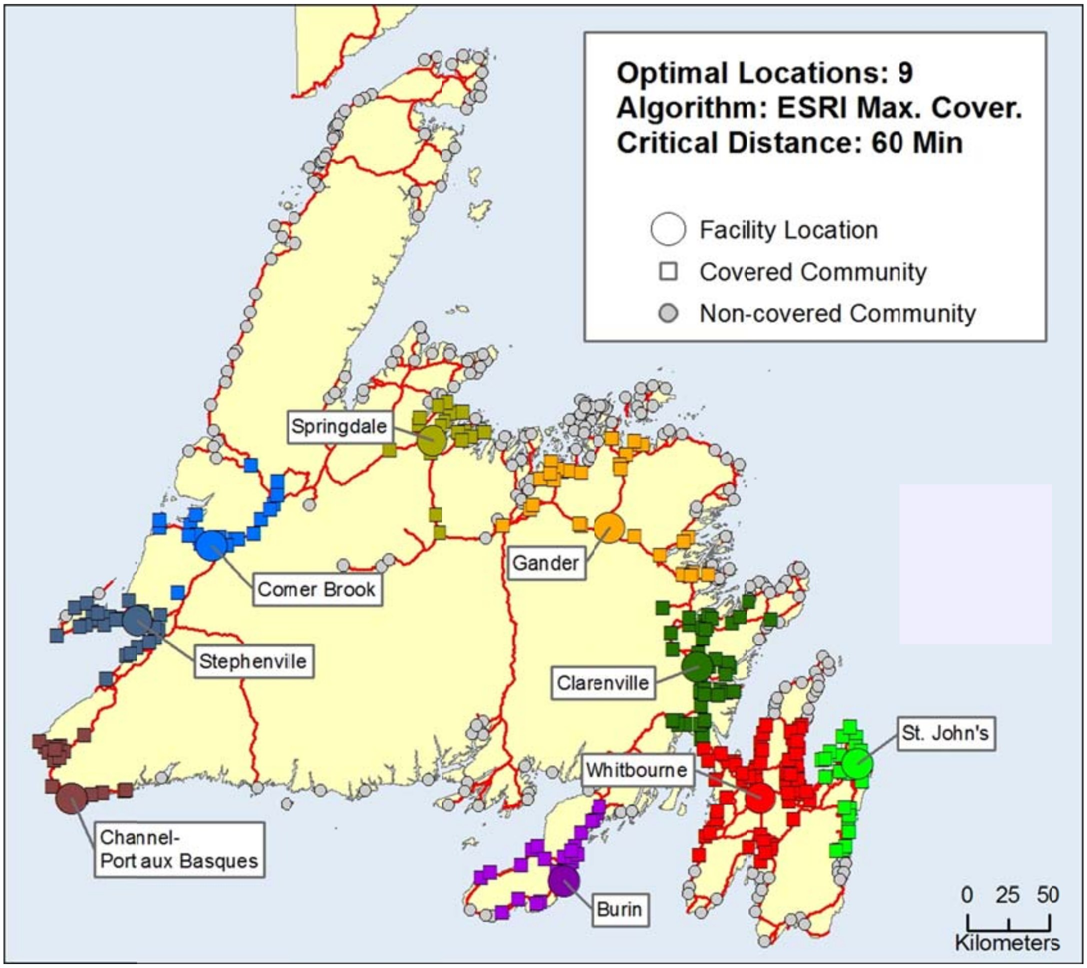


Figure 5.5: ArcGIS® 10 Network Analyst – Maximize Coverage

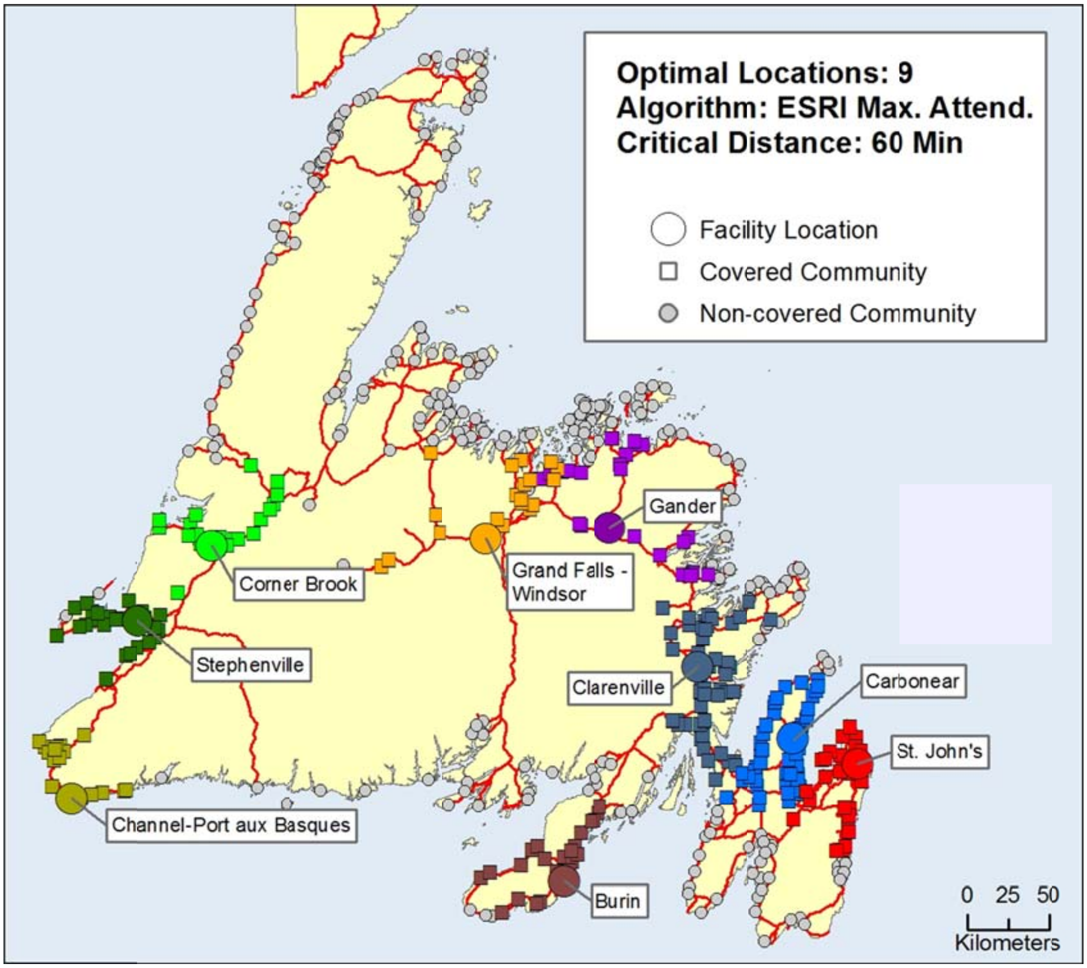


Figure 5.6: ArcGIS® 10 Network Analyst – Maximize Attendance

Both problem types produced similar results; however there were two notable differences. The maximize coverage model determined optimal facilities in Springdale and Whitbourne, which differed from the Grand Falls-Windsor and Carbonear locations of the maximum attendance model. The differences can be attributed to the slight differences in the solution algorithm. According to documentation the maximize coverage problem type assigns all demand to the nearest facility, while the maximize attendance model assigns only a partial demand calculated from the decay function (ESRI, 2010).

An interesting result is that the maximize attendance problem type produced exactly the same solution set as the distance decay variant with the LA software. It would be expected that using common parameters the results of each analysis would be similar; however an exact match was unexpected given the different solution techniques. Further analysis would be needed to determine if this is a product of a limited candidate set, or if the results are possibly due to the use of similar distance decay functions. Furthermore, the lack of facilities on the Northern peninsula indicates that both models do not appear to consider equity in facility placement.

5.1.6 Discussion of Primary Analysis

The success of each model depends on the goal of the analysis. Some models yielded higher total target population coverage, while others produced better per capita travel times to mammography facilities. Since the goal of this research is to incorporate equity into the delivery of services to residents of the province, the spatial equity variant performed better by producing a lower PCDS than the other models discussed. The spatial equity algorithm successfully compensated the demand nodes for the distance costs to access the nearest service location. The best example of this was the locating of the St. Anthony facility. The spatial equity variant was the only model to situate a facility at this location. Table 5.2 summarizes the results generated by the custom LA software.

Table 5.2: Results - Custom LA software

Rank	Current	MCLP	Distance Decay	Spatial Equity
1	St. John's	Whitbourne	St. John's	St. John's
2	Corner Brook	Lewisporte	Corner Brook	Corner Brook
3	Carbonear	Corner Brook	Carbonear	Grand Falls-Windsor
4	Grand Falls-Windsor	Wabana	Grand Falls-Windsor	Burin
5	Gander	Burin	Gander	Carbonear
6	Clarenville	Clarenville	Clarenville	St. Anthony
7	Stephenville	Stephenville	Stephenville	Clarenville
8	Burin	C.-Port aux Basques	Burin	C.-Port aux Basques
9	St. Anthony	New-Wes-Valley	C.-Port aux Basques	Gander
Pop. Coverage	90,827	92,073	91,839	89,437
PCDS	32.2	50.8	35.9	32.5

One of the most significant results of the analysis is that the original real-world locations performed better than the spatial equity model in terms of target population coverage with an insignificant difference in the PCDS. The real-world locations have 80.3% (90,837) total target population coverage versus 79.1% (89,437) for the spatial equity variant. Furthermore, the attenuated target population coverage for the real-world locations, calculated with the distance decay function, is 60.3% (68,228) versus 59.0% (66,763) for the spatial equity model. The difference between the two solution sets was the locating of a Channel-Port aux Basques facility as opposed to the real-world Stephenville facility. Stephenville ranked higher than Channel-Port aux Basques during the distance decay analysis, therefore it is a reasonable assumption that the compensation given through the spatial equity algorithm allowed this change to occur. Another consideration is that Stephenville is in close proximity (62 minutes) to the Corner Brook facility, which reduces the demand available in locating a facility in Stephenville. This suggests that there may be further opportunity for refinement in the spatial equity algorithm. A potential area of future research may be to seek further improvement in the objective function with other solution algorithms or possibly adding a facility substitution to the greedy adding algorithm.

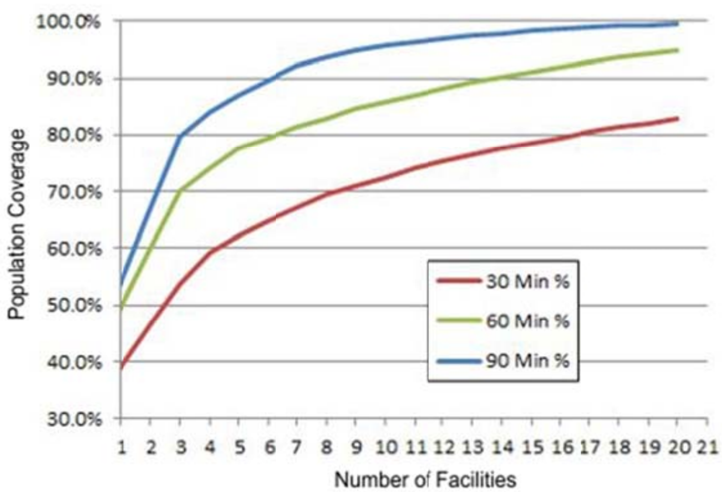
5.2 Extended Analysis

In the previous analysis, the solution set was limited to the optimal locating of nine mammography facilities. Each model produced different results; however, the spatial equity variant was identified as the best model for optimally locating of mammography facilities. The extended analysis will examine the solution sets of the spatial equity variant when increasing the number of facilities to be located and then removing the restrictions on the candidate set. Adding additional facilities to the solution set will help determine which locations are optimal for future expansion of the mammography program, while removing the restrictions on the candidate set will help determine if provincial health care centres are indeed the optimal location for placement of mammography units.

5.2.1 Additional Facilities

Consider a scenario where funding is made available and Government is seeking to open five additional mammography facilities in optimal locations, with consideration of the current facilities. These new facilities will be located taking into account the locations of the existing health care centers. Within the LA software, this means that setting the number of facilities to locate to 14, with the current nine mammography locations set as fixed facilities. This scenario will implement the spatial equity variant with a critical distance of 60 minutes. The target population group will remain females aged 40-69.

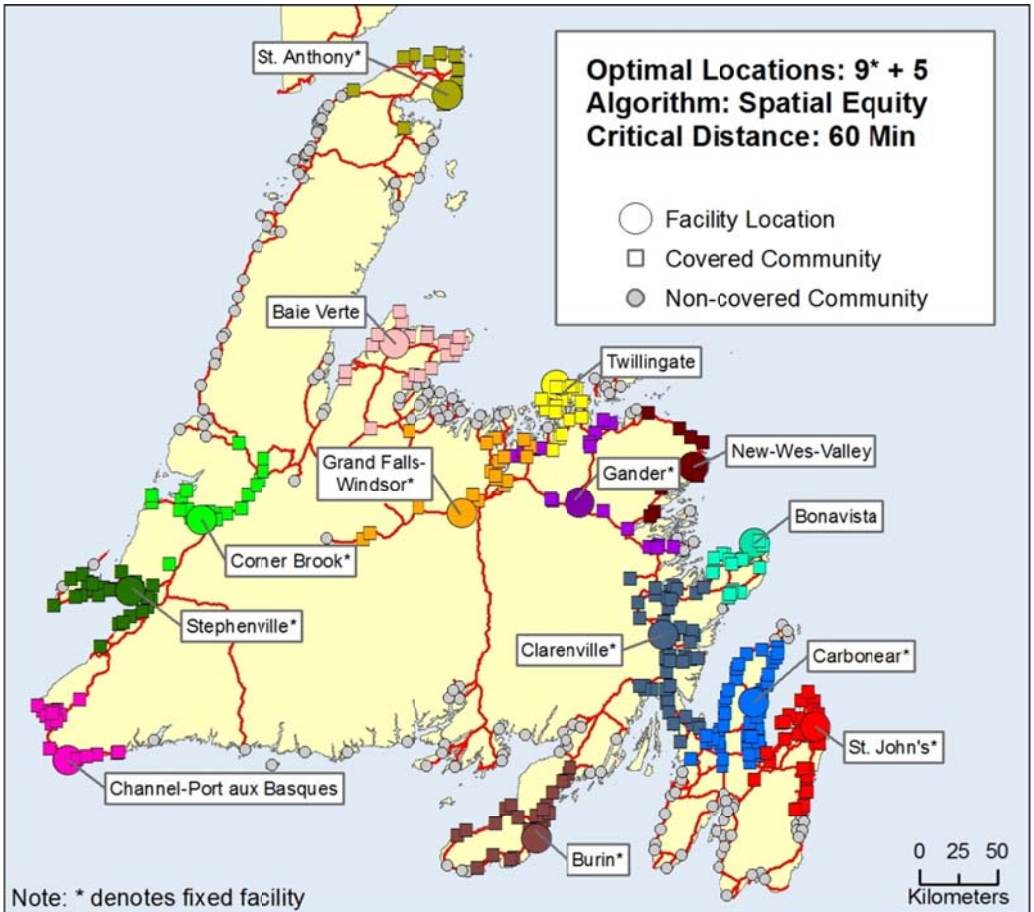
A tradeoff curve can be useful in these types of planning scenarios. Figure 5.7 illustrates the 2011 Census target population coverage for the island of Newfoundland at 30, 60, and 90 minute critical distances. It also demonstrates that as the number of facilities increases the gains in population covered by the service become less substantial.



Facility Number	Population Coverage %		
	30 Min	60 Min	90 Min
1	38.7%	49.2%	53.7%
2	46.5%	59.9%	66.8%
3	53.5%	69.9%	79.6%
4	59.2%	74.1%	84.1%
5	62.1%	77.7%	87.0%
6	64.9%	79.5%	89.6%
7	67.2%	81.3%	92.1%
8	69.3%	83.0%	93.7%
9	70.9%	84.5%	94.9%
10	72.5%	85.7%	95.6%
11	74.0%	87.0%	96.3%
12	75.3%	88.1%	96.9%
13	76.5%	89.2%	97.4%
14	77.6%	90.2%	97.9%
15	78.6%	91.2%	98.3%
16	79.5%	92.1%	98.7%
17	80.4%	92.9%	99.0%
18	81.4%	93.6%	99.2%
19	82.1%	94.3%	99.3%
20	82.8%	95.0%	99.4%

Figure 5.7: Tradeoff Curve of Spatial Equity Variant

A thematic map and summary table of the results of the analysis is shown in Figure 5.8. Note that fixed facilities are identified by an asterisk.



Score	Facility Name	PopCover	Cover%	PopTotal	Prov%
37586	St. John's*	43823	97.2	45063	39.8
7166	Corner Brook*	9225	85.2	10823	9.6
5497	Carbonear*	10296	85.5	12045	10.6
5354	Grand Falls-Windsor*	7266	64.8	11213	9.9
2940	Gander*	4888	88.9	5500	4.9
2709	Clarenville*	4473	95.3	4694	4.1
2572	Stephenville*	4516	85.9	5258	4.6
2208	Burin*	4370	89.3	4896	4.3
915	St. Anthony*	1222	38.3	3189	2.8
177009	Channel-Port aux Basques	2234	98.9	2259	2
132663	Bonavista	1966	97.6	2014	1.8
116191	Twillingate	2017	73.9	2728	2.4
95274	New-Wes-Valley	1986	100	1986	1.8
93998	Baie Verte	1455	100	1455	1.3
	Coverage Information	99737	88.2	113123	100

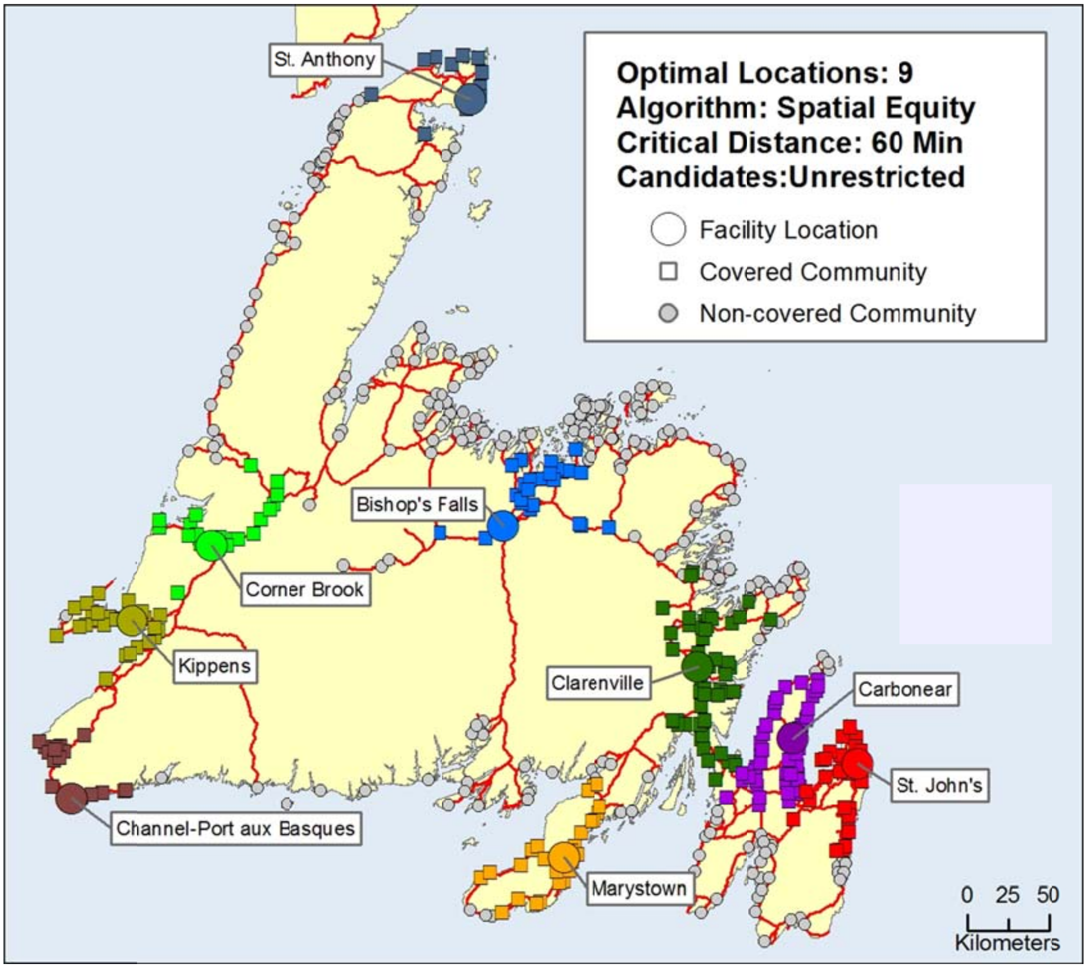
Figure 5.8: Spatial Equity Variant - Five New Facilities

The five new facilities are located in Channel-Port aux Basques, Bonavista, Twillingate, New-Wes-Valley, and Baie Verte. These selected locations are not unexpected as they are regional service centres for their respective areas. The total target population coverage has increased to 88.2% (99,774), an increase from the current 80.3% (90,837); similarly the total attenuated target population coverage has increased to 65.8% from 60.3% (68,480). There was also a substantial decrease in PCDS to 23.5 minutes, from the current 32.5 minutes. These new facilities help fill some noticeable service gaps through the island; however there are still large portions of the Northern Peninsula and Connaigre Peninsula without coverage.

Also shown in the table of Figure 5.8, each new facility will add between 1,455 and 2,234 to the total target population coverage. These are not substantial population gains in relation to coverage provided by the initial set of facilities, but it's notable that these proposed facilities would actually service larger populations than that of the St. Anthony facility within the critical coverage distance. It is likely that the placement of the St. Anthony mammography facility is intended to service the entire region, including the south coast of Labrador, not just those who live within the critical distance.

5.2.2 Unrestricted Candidates

In the primary analysis, the candidate set was restricted to the communities with health care centres, as these are the facilities that typically would support a mammography unit. In this scenario that restriction has been removed and the candidate set will be comprised of all the communities and the minimum total population constraint is removed. Removing the health care facility restriction on the candidate set accomplishes a couple of objectives. First, it ensures that the results of the model are not biased by the limited health care centre candidate set. Secondly, it also allows alternate locations to be considered that are possibly more suitable for service delivery. For this analysis the LA software will again implement the spatial equity algorithm with nine facilities. The results of the analysis are presented in Figure 5.9.



Score	Facility Name	PopCover	Cover%	PopTotal	Prov%
37586	St. John's	43823	97.2	45063	39.8
3061310	Corner Brook	9225	85.2	10823	9.6
992938	Bishop's Falls	10279	51.9	19822	17.5
594404	Marystown	4605	90	5117	4.5
390190	Carbonear	10296	85.5	12045	10.6
302338	St. Anthony	1222	38.3	3189	2.8
289448	Clarenville	4733	49.6	9547	8.4
221766	Channel-Port aux Basques	2234	98.9	2259	2
168806	Kippens	4540	86.3	5258	4.6
	Coverage Information	90957	80.4	113123	100

Figure 5.9: Spatial Equity – Unrestricted Candidates

Without restrictions on the candidate set, the spatial equity algorithm located nine facilities in very similar locations to the original spatial equity analysis. Six of the facilities (St. John’s, Carbonear, Clarenville, Corner Brook, Channel-Port aux Basques and St. Anthony) are in the exact same locations, while one location (Marystown) is in close proximity to the original location (Burin). The most notable difference within the results was the locating of a single facility in Bishop’s Falls to service the Gander / Grand Falls-Windsor area, as opposed to two separate facilities. The extra facility made available was located in Kippens, which is not surprising since nearby Stephenville scored well when the analysis was restricted to the health care centres.

A comparison between the unrestricted and restricted candidate sets, both using the spatial equity variant, shows that the total population coverage has increased to 80.4% (90,951) from 79.1% (89,480); similarly the attenuated population coverage has increased slightly to 59.7% (67,569) from 59.0% (66,763). The PCDS has risen slightly to 33.0 from 32.5 minutes. The removal of the restrictions on the candidate sites did not considerably change the overall results. The similarity in the locations indicates that the solution set was not overly biased when restricted to the sites of health care facilities. This is likely due to the fact that larger population centres have nearby health care centres. The selection of the Bishop’s Falls location may suggest that the placement of a single facility to service Gander and Grand Falls-Windsor will provide some improvement in coverage; however a more in-depth analysis would be required.

5.3 Equity Assessment

The goal of the spatial equity variant was to improve the balance between service equity and efficiency; two goals that often conflict in LA modelling (Truelove, 1993). Historically, it was generally assumed that efficient locations were distributed equitably;

however it has been demonstrated that this is often not the case (Morrill & Symons, 1977). To determine if the spatial equity variant actually improves equity, each model will be tested with an independent equity index, namely the Schutz index (1951).

The Schutz index will score each solution set to determine which model best demonstrates spatial equity. A Schutz index value of 0 would indicate equity, while higher numbers tend towards inequity. To ensure there are an adequate number of facilities in the solution set to test the index, each algorithm will now locate 25 facilities at a critical distance of 30 minutes. The target population will be set to the 2011 Census population and the candidate set will be unrestricted, but a minimum total population threshold of 100 will be set. The solution set is illustrated in Figure 5.10.

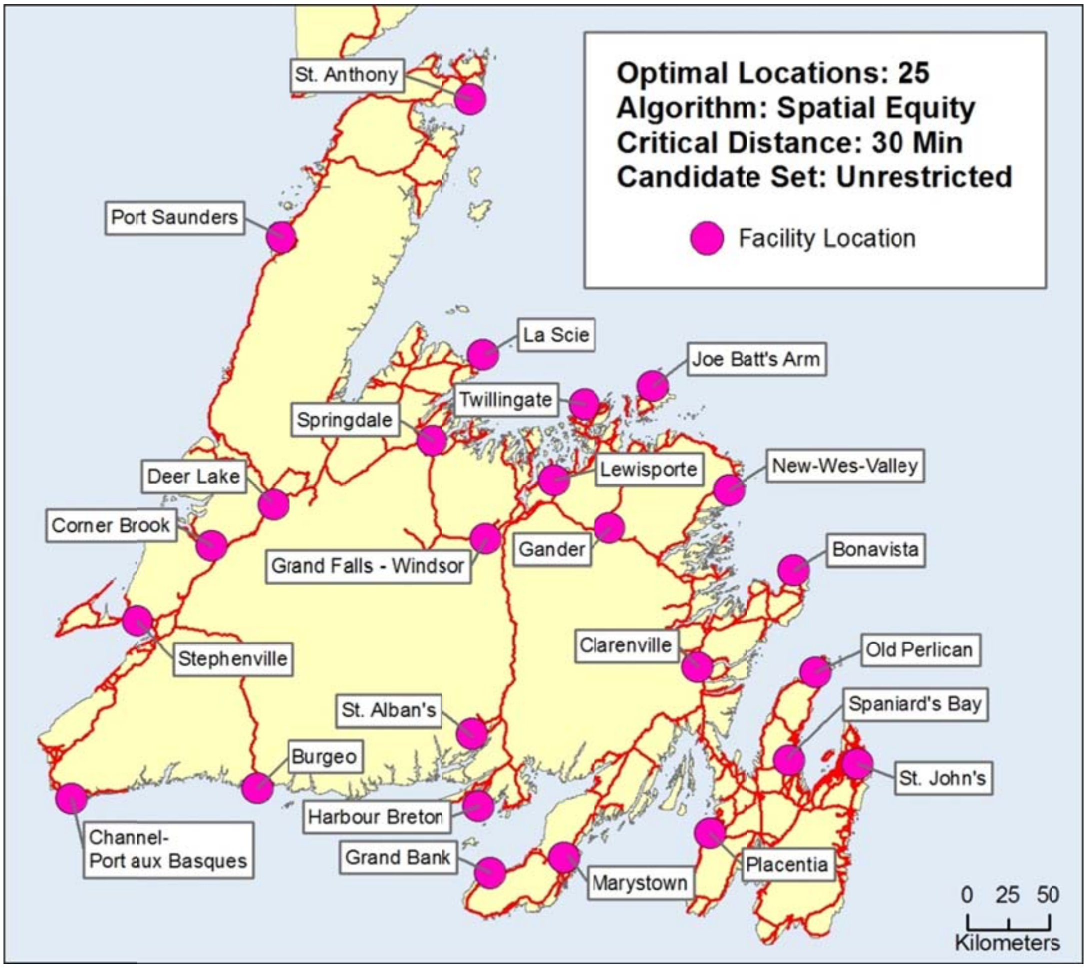


Figure 5.10: Solution Set for Schutz Index

As suggested by Truelove (1993) the index should be calculated using multiple geographies of various sizes, therefore testing will be completed using the three Regional Health Authorities, eight Rural Secretariat Regions, and the fifteen Economic Zones for the island of Newfoundland. The results of the index are displayed in Table 5.1.

Table 5.1: Schutz Index Results

	MCLP	Distance Decay Variant	Spatial Equity Variant	ArcGIS Maximize Coverage	ArcGIS Maximize Attendance
3 - Regional Health Authority (RHA)*	34.7	29.3	14.6	21.3	49.3
8 - Rural Secretariat Regions	57.3	52.0	42.7	46.6	56.0
15 - Economic Zones	104.0	64.0	45.3	57.3	74.7

***Note: the Western Health RHA boundary was extended to include the Newfoundland portion of Labrador-Grenfell Regional Health Authority**

In each of the three geographies tested, it is clear that the spatial equity variant provided more equitable results that the other models tested. The second most equitable model tested was the ArcGIS® maximum coverage model, followed by the distance decay variant and the ArcGIS® maximum attendance model. The traditional MCLP produced the most inequitable results overall and would not be recommended when considering spatial equity in service distribution.

5.4 Conclusion

As demonstrated through the analysis of various models and the Schutz index, the spatial equity variant produces the result that best incorporates equity into service delivery. The Schutz index was presented as verification of the improved spatial equity in comparison to the other models. This improvement over the currently available LA tools, such as those offered within ArcGIS®, indicates that the LA software developed for this research would be valuable in locating mammography facilities.

In terms of application research, this analysis suggests that the current mammography facilities on the island of Newfoundland are well positioned. The solution set generated by the spatial equity variant was nearly identical to the real-world locations. The only disagreement was the recommendation of the Channel-Port aux Basques facility, as opposed to the current Stephenville location. Additional analysis, however, indicates that the Stephenville location would improve the target population coverage with little change in PCDS. This suggests that future improvement of the spatial equity algorithm is possible.

The target population of mammography facilities was the 2011 Census population of females aged 40-69; however supplementary analysis using the total 2011 Census population (both sexes) shows that these facility locations remain the same. Therefore, it is anticipated that the spatial equity variant would work equally well to plan for similar types of preventive health services, and possibly general public services. Regardless, this research suggests that spatial equity should be considered in any public facility placement to ensure equity is considered in addition to efficiency. Expanding spatial equity into other LA models to locate other types of public facilities would be another potential area for future research.

In terms of the number of facilities, the nine current facility locations provide adequate coverage to the study area. Only slight improvement could be found from the addition of new facilities. With the exception of the St. Anthony location, all current locations provide service to more than 4,000 members of the target population within the 60 minute critical distance. The placement of St. Anthony can be interpreted as a conscious effort of government to incorporate spatial equity into service delivery.

Finally, this analysis suggests that limiting the placement of mammography units to the province's health care centres does not appear to substantially impair service delivery.

Chapter 6: Summary & Conclusion

The rapidly aging population of Newfoundland and Labrador presents a significant financial burden on the province's health care system. The growing number of senior citizens combined with the associated high per capita cost of providing health care is leading to dramatic growth in health care expenditures. As discussed previously in Chapter 1, it is anticipated that the trend will continue well into the next decade. One method that may help alleviate these costs is the promotion of preventive health care. Preventive health care is a wide-ranging area of medical care that includes commonly known services such as immunizations, blood testing, and cancer screening exams. Unlike primary health services, preventative health care is intended for healthy people, who are in general less willing to travel to access service. Therefore, an alternate location decision methodology focusing on accessibility is required when locating preventive health service facilities. This study presented a methodology for the optimal locating of preventive health facilities for the island of Newfoundland.

This research has outlined the value of using LA models in identifying the optimal locations of preventive health facilities. Unlike traditional models that focus on efficiency, preventive health models consider equity in service delivery. The goal has been to locate facilities efficiently while improving spatial equity in the distribution of services. This is of particular importance in Newfoundland with such a widely dispersed rural population, where equity would mean that individuals living in these rural communities have the same ease of access to a service as those living in urban areas. The traditional MCLP algorithm was shown to be not successful in meeting this goal.

Therefore, two variants were proposed to help improve spatial equity. The distance decay variant incorporated a linear distance decay function to more efficiently locate facilities, while the spatial equity variant improved equity by compensating the demand for the distance costs associated with accessing services.

To implement the proposed variants, a customized open-sourced LA software program was created. The software was developed in a loose-coupled approach with the ability to manipulate the data that was generated and compiled with the ArcGIS® 10 software. This software provides the user with the ability to adjust variables to evaluate the solutions sets of various planning scenarios. A summary table is presented within the software which provides details of the population coverage of each facility, as well as population totals and per capita distance to services (PCDS). For more in-depth analysis, individual community information is also stored in the underlying database, which is formatted in a manner that allows for easy import in to ArcGIS® for extended analysis and visualization.

Since different preventive health services have slightly different service standards and intended target populations, mammography facilities were used as a representative facility type. This study determined the optimal location of mammography facilities for the target group of females aged 40-69. The LA software was used to derive solution sets for the MCLP, distance decay, and spatial equity algorithms, while the ArcGIS® Network Analyst extension was used to solve the maximize coverage and maximize attendance problem types. During the analyses, each solution set was discussed and compared to the results of the other models and the real-world location of facilities.

Based on the Schutz index, the spatial equity variant was the best model investigated in terms of equity in service delivery.

The most notable outcome of this research was that the spatial equity variant produced the only solution set that prioritized the locating of a facility on the Northern Peninsula. This area is sparsely populated and would be considered an inefficient location in traditional LA modeling. The placement of the St. Anthony facility, however, improves equity in the overall access of services. Because a mammography unit currently operates in this community, the provincial breast screening program has likely considered spatial equity in service provision.

An unexpected outcome of the analysis was that the original real-world locations provided higher target population coverage than the spatial equity variant, with only a slight difference in the PCDS. The model suggested the placement of a facility in Channel-Port aux Basques versus the current Stephenville location. This result is due to the manner in which the software calculated the location score in combination with the augmented demand weight assigned to communities through the spatial equity variant algorithm. This was not discouraging as the results were still promising. It does indicate, however, that there is room for improvement in the algorithm. The results also indicate that the current mammography locations are well situated from a spatial equity perspective. With respect to optimizing the number of facility locations, it was shown that the addition of new facilities would improve the total population coverage and PCDS; however, the improvement was not substantial. It would ultimately be up to the judgment of the decision-makers to determine whether the improvement would be worth the financial investment. Otherwise, these locations could be target areas for alternative service delivery strategies, such as a mobile mammography program.

Through this research an increased understanding was gained of the impact of the model parameters on the analysis. As noted by Karasakal and Karasakal (2004), the optimal solution to the MCLP, as well as the variants proposed in this research, will be sensitive to the choice of critical distance. Moreover, Chung (1986) argued that justification of the MCLP is relatively weak in the presence of ambiguity in the critical distance. Similarly, the choice of the distance decay function could have a direct impact on the solution set. Therefore, it is important to discuss each parameter with the decision-makers before the analysis commences. If the parameters are not respected, LA models could be used as tools to create results that have been manipulated to suit a preconceived facility placement scenario.

When developing the MCLP variants, appreciation and insight was gained into the theory behind the algorithms. The translation from mathematical formulas to program code provided understanding of the weaknesses and strengths of the greedy adding approach. In retrospect, other solution methods, such as the greedy adding with substitution method, may have provided improved solution sets, so this will remain an area of potential future research. The strength of the simpler greedy adding approach, however, was that it allowed the spatial equity variant to be quickly and easily executed.

As noted statistician George Box once said, “Essentially, all models are wrong, but some are useful” (Box & Draper, 1987). Models are not the real world and people do not always act the way the models say they should (Patton & Sawicki, 1993). There are also many other factors that must be considered when locating any type of facility. Some of these factors, such as wait times, perceived service quality, personal barriers, or access to a vehicle, can be hard to quantify. For this reason, the results of these models should be used in conjunction with all other information to make the best decision possible. Regardless, it is anticipated that this research demonstrated that LA models can be part of an effective solution in preventive health facility planning, particularly when evidence-based decision making is considered important.

Bibliography

Anderson, G. F., Frogner, B. K., Johns, R. A., & Reinhardt, U. E. (2006). Health care spending and use of information technology in OECD countries. *Health Affairs, 25*(3), 819-831.

Anselin, L., & Getis, A. (1992). Spatial statistical analysis and geographic information systems. *The Annals of Regional Science, 26*, 19-33.

Bennett, D. A. (1997). A framework for the integration of geographical information systems and model-base management. *International Journal of Geographical Information Science, 11*(4), 337-357.

Bennett, R. J. (1980). *The Geography of Public Finance*. London: Methuen.

Bennett, R. J. (1983). Individual and territorial equity. *Geographical Analysis, 15*(1), 50-57.

Berke, E., & Shi, X. (2009). Computing travel time when the exact address is unknown: A comparison of point and polygon ZIP code approximation methods. *International Journal of Health Geographics, 8*(23)

Berman, O., & Krass, D. (2002). Generalized maximum cover location problem. *Computers & Operations Research, 29*, 563-581.

Berman, O., Krass, D., & Drezner, Z. (2003). The gradual covering decay location problem on a network. *European Journal of Operational Research, 151*(3), 474-480.

Berman, O., & Parkan, C. (1981). A facility location problem with distance-dependent demand. *Decision Sciences, 12*(4), 623-632.

Box, G. E. P., & Draper, N. R. (1987), *Empirical model building and response surfaces*. New York, NY: John Wiley & Sons.

Brandeau, M. L., & Chiu, S. S. (1989). An overview of representative problems in location research. *Management Science, 35*(6), 645-674.

Breast Cancer Society of Canada (2013), "Breast Screening Information for All Provinces and Territories"; (http://www.bcsc.ca/p/48/l/124/t/Breast-Cancer-Society-of-Canada---Regional-Breast-Cancer-Screening); Last Accessed: Jan 24/2014

Canadian Task Force on Preventive Health Care (2011). Recommendations on screening for breast cancer in average-risk women aged 40-74 years. *CMAJ Guildlines, 183* (17), 1991-2001.

Canadian Institute for Health Information. (2010). *Health Care in Canada 2010*. Ottawa, ON: CIHI

Canadian Institute for Health Information. (2012) *Breast Cancer Surgery in Canada, 2007-08 to 2009-10.* Ottawa, ON: CIHI

Chaudhry, S. S., Choi, I., & Smith, D. K. (1995). Facility location with and without maximum distance constraints through the *p*-median problem. *International Journal of Operations and Production Management, 15*(10), 76-81.

Christaller, W. (1966). *Central places in southern Germany* (C. W. Baskin Trans.). Englewood, NJ: Prentice Hall.

Chung, C. (1986). Recent applications of the maximal covering location planning (M.C.L.P.) model. *Journal of Operational Research Society, 37*(8), 735-746.

Church, R. (2002). Geographical information systems and location science. *Computers & Operations Research, 29*(6), 541-562.

Church, R., & ReVelle, C. S. (1974). The maximal covering location problem. *Papers in Regional Science, 32*(1), 101-118.

Cromley, R., & Mrozinski, R. (2002). Analyzing geographic representation error in capacitated location-allocation modeling. *Networks and Spatial Economics, 2*(1), 65-78.

Crompton, J., & Lamb, C. (1983). The importance of the equity concept in allocation of public services. *Jounral of Macromarketing, 3*(1), 28-39.

Current, J., Daskin, M. S., & Schilling, D. (2002). Discrete network location models. In Z. Drezner, & H. Hamacher (Eds.), *Facility location: Applications and theory* (pp. 81-118) Springer-Verlag.

Current, J., & Schilling, D. (1990). Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis, 22*(2), 116-126.

Daskin, M. S. (1995). Network and discrete location: models, algorithms and applications. New York: John Wiley and Sons, Inc.

Daskin, M. S. (2008). What you should know about location modeling. *Naval Research Logistics*, *55*, 283-294.

Daskin, M. S., & Dean, L. K. (2004). Location of health care facilities. In F. Sainfort, M. L. Brandeau & W. Pierskalla (Eds.), *Handbook of OR/MS in health care: A handbook of methods and applications* (pp. 43-76) Kluwer.

Daskin, M. S., Haghani, A. E., Khanal, M., & Malandraki, C. (1989). Aggregation effects in maximum covering models. *Annals of Operations Research*, *18*, 115-140.

Daskin, M. S., & Owen, S. H. (2002). Location models in transportation. In R. Hall (Ed.), *Handbook of Transportation Science* (2nd Ed.). (pp. 311-370). Kluwer Academic Publishers.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*, 269-271.

Drezner, Z., & Eiselt, H. A. (2002). Consumers in competitive location models. In Z. Drezner, & H. Hamacher (Eds.), *Facility location: Applications and Theory*. (pp. 151-178). Springer-Verlag.

ESRI (2010). *ArcGIS® resource center: Desktop 10: Location-allocation analysis* from http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Location_allocation_analysis/004700000050000000/, Last Accessed: June 17, 2013

Fallah, H., Sadigh, A. N., & Aslanzadeh, M. (2009). In Zanjirani Farahani R., Hekmatfar M. (Eds.), *Facility location: Concepts, models, algorithms and case studies* (VIII ed.) Physica-Verlag Heidelberg.

Farhan, B., & Murray, A. T. (2006). Distance decay and coverage in facility location planning. *The Annals of Regional Science*, *40*(2), 279-295.

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

Gaile, G. L. (1984). Measures of spatial equality. In G. L. Gaile, & C. J. Willmott (Eds.), *Spatial Statistics and Models.* (pp. 223-233). Dordrecht: D Reidel.

Garey, M. R., & Johnson, G. S. (1979). *Computers and intractibility: A guide to the theory of NP-completeness*. New York: W. H. Freeman and Company.

Goldsmith, J. (1989). A radical prescription for hospitals. *Havard Business Review, May/June 1989*, 104-110.

Goodchild, M. F. (1979). The aggregation problem in location-allocation. *Geographical Analysis*, *11*, 240-255.

Goodchild, M. F., Haining, R., & Wise, S. (1992). Integrating GIS and spatial data analysis: Problems and possibilities. *International Journal of Geographical Information Systems, 6*(5), 407-423.

Government of Newfoundland and Labrador. (2006). *Demographic change: Issues and Implications*. St. John's, NL: Economic & Statistics Branch, Department of Finance.

Government of Newfoundland and Labrador. (2012a). *Breast Screening Program Expanded.* [Press Release] April 27, 2012. Access: http://www.releases.gov.nl.ca/releases/2012/health/0427n07.htm

Government of Newfoundland and Labrador. (2012b). *Consolidated Census Subdivision (CCS) by Community, 2011*, *http://www.stats.gov.nl.ca/Statistics/Census2011/* (Last Retrieved: Jan 18, 2014)

Government of Newfoundland and Labrador. (2013). *Population Projections for Newfoundland and Labrador*, http://www.economics.gov.nl.ca/POP-default.asp (Last Retrieved: Jan 18, 2014)

Gu, W., Wang, X., & McGregor, S. E. (2010). Optimization of preventive health care facility locations. *International Journal of Health Geographics, 9,* 17.

Hakimi, S. L. (1964), Optimum locations of switching centers and the absolute centers and medians of a graph", *Operations Research, 12*(3), 450–459

Hillsman, E. L., & Rhoda, R. (1978). Errors in measuring distances from populations to service centres. *Annals of Regional Science, 12*, 74-88.

Hodgson, M. J., & Hewko, J. (2003). Aggregation and surrogation error in the p-median model. *Annals of Operations Research, 123,* 53-66.

Holmes, J., Williams, F. B., & Brown, L. A. (1972). Facility location under a maximum travel restriction: An example using day care facilities. *Geographical Analysis, 4*(3), 258-266.

Hurst, M. E. E. (1972). *A Geography of Economic Behavior: An Introduction.* North Scituate, Mass., Duxbury Press.

Johnston, R. J., Gregory, D., & Smith, D. M. (2000). In Malden M. A. (Ed.), *The Dictionary of Human Geography* (4th Ed.). Wiley-Blackwell.

Kalogirou, S., & Foley, R. (2006). Health, place & hanly: Modelling accessibility to hospitals in Ireland. *Irish Geography, 39*(1), 52-68.

Karasakel, O., & Karasakal, E. K. (2004). A maximal covering location model in the presence of partial coverage. *Computers & Operations Research, 31*(9), 1515-1526.

Longley, P. A., Goodchild, M., Maguire, D. J., & Rhind D. W. (2010). *Geographic Information Systems and Science.* (3rd Ed.). West Sussex, England: John Wiley & Sons.

Longley, P. A., Goodchild, M., Maguire, D. J., and Rhind D. W.

Losch, A. (1954). *The economics of location*. New Haven, CT: Yale University Press.

Love, D., & Lindquist, P. (1995). The grographic accessability of hospitals to the aged: A geographic information systems analysis within illinois. *Health Services Research, 29*(6), 629-651.

Lovett, A., Haynes, R., Sunnenberg, G., & Gale, S. (2002). Car travel time and accessibility by bus to general practitioner services: A study using patient registers and GIS. *Social Science & Medicine, 55*(1), 97-111.

Marianov, V., & Serra, D. (2002). Location problems in the public sector. In Z. Drezner, & H. Hamacher (Eds.), *Facility Location: Applications and Theory* (pp. 119-150) Springer-Verlag.

Martin, D., & Williams, H. C. (1992). Market area analysis and accessibility to primary health care centres. *Enviroment and Planning A, 24*(7), 1009-1019.

Maurer, D., & Martel S. (2002). Extreme Programming: Rapid Development for Web-Based Applications. *IEEE Internet Computing, 6*(1), 86-91.

Maxwell, A. J. (2000). Relocation of a static breast screening unit: A study of factors affecting attendance. *Journal of Medical Screening, 7*(2), 114-115.

Messina, J. P., Shortridge, A. M., Groop, R. E., Varnakovida, P., & Finn, M. J. (2006). Evaluating michigan's community hospital access: Spatial methods for decision support. *International Journal of Health Geographics, 5*(42)

Morrill, R. L., & Symons, J. (1977). Efficiency and equity aspects of optimum location. *Geographical Analysis, 9*(3), 215-225.

Patton, C. V., & Sawicki, D. S. (1993). *Basic Methods of Policy Analysis and Planning* (2nd Ed.). Prentice Hall.

Plastria, F. (1996). Optimal location of undesirable facilities: A selective overview. *Journal of Belgian Operational Research Society*, *36*, 109-127.

Rahman, S., & Smith, D. K. (2000). Use of location-allocation models in health service development planning in developing nations. *European Journal of Operational Research, 123*(3), 437-452.

ReVelle, C. S., Church, R., Schilling, D. (1975). A note on the location models of Holmes, Williams and Brown. *Geographic Analysis*, *7*, 457-459.

ReVelle, C. S., & Eiselt, H. A. (2005). Location analysis: A synthesis and survey. *European Journal of Operational Research, 165*(1), 1-19.

ReVelle, C. S., Eiselt, H. A., & Daskin, M. S. (2008). A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research, 184*(3), 817-848.

ReVelle, C. S., Marks, D., & Liebman, J. C. (1970). An analysis of private and public sector location models. *Management Science*, *15*(11), 692-707.

Rosing, K. E., & Hodgson, M. J. (1996). A systematic classification of applications of location-allocation models. *Belgian Journal of Operations Research, Statistics and Computer Science*, *36*, 77-108.

Schutz, R. R. (1951). On the measurement of income equity. *American Economic Review*, *41*, 107-122.

Sommerville, I. (1995). *Software Engineering* (3rd Ed). Essex, England: Addison-Wesley.

Statistics Canada. (2011). *Overview of the Census),* Statistics Canada Catalogue no. 98-302-XIE. Ottawa, ON. http://www12.statcan.gc.ca/census-recensement/2011/ref/overview-apercu/98-302-XIE.pdf (Last Retrieved: June 18, 2013)

Starfield, B., Shi, L., & Macinko, J. (2005). Contributions of Primary Care to Health Systems and Health. *The Millbank Quarterly*, *83 (3)*, 457-502.

Statistics Canada. (2012). *Table 051-0001 - Estimates of population, by age group and sex for July 1, Canada, provinces and territories, annual (persons unless otherwise noted),* CANSIM (database). http://cansim2.statcan.gc.ca/ (Last Retrieved: June 18, 2013)

Taylor, P. J. (1971). Distance Transformation and Distance Decay Functions. *Geographical Analysis, 3(3),* 221-238

Toregas, C., & ReVelle, C. S. (1972). Optimal location under time or distance restraints. *Papers in Regional Science*, *28*(1), 133-143.

Truelove, M. (1993). Measurement of spatial equity. *Environment and Planning C: Government and Policy*, *11*(1), 19-34.

Verter, V., & Lapierre, S. D. (2002). Location of preventive health care facilities. *Annals of Operations Research*, *110*(1-4), 123-132.

Walker, A. R. (1977). What is the outlook for health improvements. *South African Medical Journal*, *52*(15), 589-590.

Wang, F., McLafferty, S., Escamilla, V., & Luo, L. (2008). Late-stage breast cancer diagnosis and health care access in Illinois. *The Professional Geographer : The Journal of the Association of American Geographers*, *60*(1), 54-69.

Weber, A. (1929). *Theory of the location of industries* The University of Chicago Press.

Weiss, J., Greenlick, M., & Jones, J. F. (1971). Determinants of medical care utilization: The impact of spatial factors. *Inquiry*, *8*(4), 50-57.

Zhang, Y., Berman, O., & Verter, V. (2009). Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research*, *198*(3), 922-935.

Zhou, J., & Liu, B. (2003). New stochastic models for capacitated location-allocation problem. *Computers & Industrial Engineering*, *45*(1), 111-125.

Zimmerman, S. M. (1997). Factors influencing Hispanic participation in prostate cancer screening. *Oncology Nursing Forum*, *24*(3), 499-504.

Appendix: Source Code

```
Imports System
Imports System.IO

Public Class Form1
    '---VARIABLE DECLARATION---
    Dim DistanceMatrix(999999) As Integer
    Dim PopulationMatrix(999, 3) As 'Integer 'Position - 0: Population; 1: Covered; 2: Closest Facility; 3: Population Weight
    Dim SolutionSet(100, 2) As Integer     'Position - 0: Community ID; 1: Coverage Population; 2:Total Reliant Population
    Dim CandidateSet(999, 1) As Integer  'Position - 0: Community ID; 1: Covered
    Dim FixedSet(100, 1) As Integer 'Position - 0: Community ID; 1: Covered
    Dim CriticalDistance As Double 'Critical Distance
    Dim dtAvailComm As SampleDataSet.Community2011DataTable 'Community Data
    Dim dtMatrixData As SampleDataSet.Network2011DataTable 'Network Data

    '---INITIALIZE DATA FORM---
    Private Sub Form1_Load(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles MyBase.Load
        Me.Height = 360
        lblProgress.Text = "Ready"
        ComboBox1.SelectedIndex = 1
    End Sub

    '---HANDLE PROGRESS BAR---
    Private Sub ProgressBar(ByVal k As Integer, ByVal Tot As Integer)
        ProgressBar1.Value = (k / Tot) * 100
        Application.DoEvents()
        Me.Refresh()
    End Sub

    '---HANDLE CHANGE IN CRITICAL DISTANCE---
    Private Sub txtDist_TextChanged(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles
txtDist.TextChanged
        If Not IsNumeric(txtDist.Text) Then
            txtDist.Text = "0"
        End If
    End Sub

    '---HANDLE CHANGE IN FACILITY COUNT---
    Private Sub txtFacNum_TextChanged(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles
txtFacNum.TextChanged
        If txtFacNum.Text = "" Then
            txtFacNum.Text = "1"
        ElseIf Not IsNumeric(txtFacNum.Text) Or txtFacNum.Text < 1 Or txtFacNum.Text = "" Then
            txtFacNum.Text = "1"
        End If
    End Sub

    '---HANDLE CHANGE IN MINIMUM POPULATION---
    Private Sub txtMinPop_TextChanged(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles
txtMinPop.TextChanged
        If Not IsNumeric(txtMinPop.Text) Then
            txtMinPop.Text = "500"
        End If
    End Sub

    '---OPEN DATABASE FOR EDITS---
    Private Sub btnDbEdit_Click(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles btnDbEdit.Click
        Dim FixedCounter As Integer
        If DataGridView1.Visible = False Then
            Me.Height = 600
            DataGridView1.Visible = True
            btnDbEdit.Text = "Save / Exit"
            DataGridView1.ReadOnly = False
            Me.Community2011TableAdapter.Fill(SampleDataSet.Community2011)
        Else
```

```
        Me.Height = 360
        DataGridView1.Visible = False
        btnDbEdit.Text = "Edit Database"
        Me.Community2011TableAdapter.Update(SampleDataSet)
        Me.Community2011TableAdapter.Dispose()
        DataGridView1.ReadOnly = True
        If chkFixed.Checked = True Then 'Update Number of fixed facilities
            FixedCounter = Community2011TableAdapter.CountFixed
            chkFixed.Text = "Fixed Locations (" & FixedCounter & ")"
            Community2011TableAdapter.Dispose()
        End If
    End If

End Sub

'---HANDLE FIXED CHECKBOX---
Private Sub chkFixed_CheckChanged(ByVal sender As Object, ByVal e As System.EventArgs)
    Dim FixedCounter As Integer
    If chkFixed.Checked = True Then
        FixedCounter = Community2011TableAdapter.CountFixed
        chkFixed.Text = "Fixed Locations (" & FixedCounter & ")"
        Community2011TableAdapter.Dispose()
    Else
        chkFixed.Text = "Fixed Locations"
        FixedCounter = 0 'If no fixed locations, set counter to 0
    End If

End Sub

'---ALGORITHM GREEDY ADDING---
Private Sub btnGA_Click(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles btnGA.Click
    'Declare local variables
    Dim TopScore, TempScore, BestComm, MaxCommunity As Integer
    Dim FixedSet(100, 1), CommWeight As Integer
    Dim FixCount, CandCount, FacilityCount, k, j, u As Integer 'Counters
    Dim AlreadyFixed As Boolean
    Dim DistVar As Integer

    '---Initialize - Clear -------------
    lblProgress.Text = "Initializing Distance Matrix"
    GroupBox1.Enabled = False
    Me.SampleDataSet.ResultsTable.Clear()
    Array.Clear(SolutionSet, 100, 2)
    Array.Clear(PopulationMatrix, 999, 2)
    Me.Community2011TableAdapter.ClearAllCovered()
    Application.DoEvents()
    ReDim SolutionSet(100, 2), FixedSet(100, 1), PopulationMatrix(999, 3), CandidateSet(999, 1)

    'Reset variables for mulitple runs
    FixCount = 0
    CandCount = 0
    '---End initialize clear----------

    '---Initialize - Fill ----------------
    If txtDist.Text = "0" Then 'Critical Distance can't be 0
        CriticalDistance = 0.0001
    Else
        CriticalDistance = CInt(txtDist.Text)
    End If

    dtMatrixData = Network2011TableAdapter.GetData 'Copy Matrix to an Array
    For Each drMatrixData As SampleDataSet.Network2011Row In dtMatrixData
        DistanceMatrix(drMatrixData.OrigDest) = drMatrixData.T_Time
    Next

    lblProgress.Text = "Initializing Community Matrix"
    Application.DoEvents()
```

```
dtAvailComm = Community2011TableAdapter.GetData 'Copy Community to an Array
For Each drAvailComm As SampleDataSet.Community2011Row In dtAvailComm
    If ComboBox1.SelectedIndex = 0 Then ‘All Census
        PopulationMatrix(drAvailComm.CommID, 0) = drAvailComm.Pop2011 'Populate dataset all census
    Else
        PopulationMatrix(drAvailComm.CommID, 0) = drAvailComm.Pop2011F 'Populate dataset females 40-69
    End If
    ‘If community larger than minimum it’s a candidate
    If drAvailComm.Candidate = 1 And drAvailComm.Pop2011 > txtMinPop.Text Then
        CandCount = CandCount + 1
        CandidateSet(CandCount, 0) = drAvailComm.CommID
    End If
    If drAvailComm.Fixed = 1 And chkFixed.Checked = True Then
        FixCount = FixCount + 1
        FixedSet(FixCount, 0) = drAvailComm.CommID
    End If
Next
MaxCommunity = Community2011TableAdapter.MaxComm
FacilityCount = CInt(txtFacNum.Text)
'---End initialize fill--------------------

'---Fixed Facilities---------------------
If Community2011TableAdapter.CountFixed > 0 Then
    ProgressBar(0, 1)
    For i = 1 To FixCount ‘Do for each fixed facility
        lblProgress.Text = "Processing 'Fixed' facility " & i & " of " & FixCount
        Application.DoEvents()
        TopScore = -1
        For f = 1 To FixCount 'Find best fixed facility
            If FixedSet(f, 1) = 0 Then 'If not already used
                ProgressBar(f, FixCount)
                TempScore = 0
                For k = 100 To MaxCommunity 'For each demand Community "k"
                    If PopulationMatrix(k, 1) = 0 Then
                        DistVar = DistanceMatrix(CInt(FixedSet(f, 0) & k))
                        If DistVar < CriticalDistance + 1 Then 'If less than critical distance
                            If chkDistance.Checked = False Then 'No Distance Function
                                TempScore = PopulationMatrix(k, 0) + TempScore
                            Else 'Distance Function
                                TempScore = PopulationMatrix(k, 0) * (1 - (DistVar / CriticalDistance)) + TempScore
                            End If
                        End If
                    End If
                Next k
                If TempScore > TopScore Then
                    TopScore = TempScore 'Find best scoring fixed facility
                    BestComm = FixedSet(f, 0)
                End If
            End If
        Next f
        SolutionSet(i, 0) = BestComm

        For f = 1 To FixCount 'Mark covered in Fixed Set
            If FixedSet(f, 0) = BestComm Then
                FixedSet(f, 1) = 1
            End If
        Next f

        For k = 100 To MaxCommunity 'Mark covered communities
            AlreadyFixed = False
            DistVar = DistanceMatrix(CInt(BestComm & k))
            If DistVar <= CInt(TrackBar1.Value) Then
                For g = 1 To FixCount 'Don't mark a fixed community as covered
                    If FixedSet(g, 1) = 0 And FixedSet(g, 0) = k Then
                        AlreadyFixed = True
                    End If
                Next g
```

```
            If AlreadyFixed = False Then
               PopulationMatrix(k, 1) = 1
            End If
            For j = 1 To CandCount ' Remove Candidate from Candidate Set
               If CandidateSet(j, 0) = k Then
                  CandidateSet(j, 1) = 1
               End If
            Next
         End If
      Next

      UpdateOutputTable(BestComm, TopScore, True) 'Update Output Table

      If chkSpatialEquity.Checked = True Then 'If spatial equity variant then update population weights
         UpdatePopWeightMatrix(FacilityCount)
      End If
   Next i
End If

If chkSpatialEquity.Checked = True Then 'If spatial equity variant then update population weights
   UpdatePopWeightMatrix(FacilityCount)
End If
'---End Fixed Facilities----------------------


'---Non-Fixed Communities--------------
For u = (FixCount + 1) To FacilityCount
   ProgressBar(0, 1)
   lblProgress.Text = "Processing facility " & u & " of " & FacilityCount
   Application.DoEvents()
   TopScore = -1
   For j = 1 To CandCount 'For each potential facility candidate community "j"
      TempScore = 0
      If CandidateSet(j, 1) = 0 Then 'If a candidate
         ProgressBar(j, CandCount)
         For k = 100 To MaxCommunity 'To every other demand community
            If chkSpatialEquity.Checked = True Then
               CommWeight = PopulationMatrix(k, 3) 'Spatial equity variant population weight
            Else
               CommWeight = PopulationMatrix(k, 0)
            End If
            If PopulationMatrix(k, 1) = 0 And CommWeight > 0 Then 'If not covered and population not 0
               DistVar = DistanceMatrix(CInt(CandidateSet(j, 0) & k))
               If DistVar < CriticalDistance + 1 Then 'If less than critical distance
                  If chkDistance.Checked = False Then 'No Function
                     TempScore = CommWeight + TempScore
                  Else 'Distance Decay Function
                     TempScore = CommWeight * (1 - (DistVar / CriticalDistance)) + TempScore
                  End If
               End If
            End If
         Next k
         If TempScore > TopScore Then
            TopScore = TempScore 'Find best Community
            BestComm = CandidateSet(j, 0)
         End If
      End If
   Next j

   SolutionSet(u, 0) = BestComm 'Place best community in solution set

   For k = 100 To MaxCommunity 'Mark covered communities
      DistVar = DistanceMatrix(CInt(BestComm & k))
      If DistVar <= CriticalDistance Then
         PopulationMatrix(k, 1) = 1
         For j = 1 To CandCount ' Remove Candidate from Candidate Set
            If CandidateSet(j, 0) = k Then
```

```
                CandidateSet(j, 1) = 1
            End If
        Next
    End If
Next

UpdateOutputTable(BestComm, TopScore, False) 'Update Output Table

If chkSpatialEquity.Checked = True Then
    UpdatePopWeightMatrix(FacilityCount) 'If spatial equity variant then update population weights
End If
Application.DoEvents()
Next u
'---End Non-Fixed ------------------

'---ASSIGN CLOSEST FACILITY---------
Dim PerCapita As Double
Dim PerCapitaSum As Integer

PerCapitaSum = 0

ProgressBar(1, 1)
lblProgress.Text = "Assigning Communities"
Application.DoEvents()

Dim ClosestCode, ClosestFacilityNum, CurrDist, BestDist As Integer

For j = 100 To MaxCommunity 'For every Community
    ClosestCode = 0
    ClosestFacilityNum = 0
    BestDist = 2500
    For s = 1 To FacilityCount 'For each facility
        CurrDist = DistanceMatrix(CInt(j & SolutionSet(s, 0)))
        If CurrDist < BestDist Then
            BestDist = CurrDist
            ClosestCode = SolutionSet(s, 0)
            ClosestFacilityNum = s
        End If
    Next s
    PopulationMatrix(j, 2) = ClosestCode 'Put closest facility in dataset

    PerCapitaSum = PerCapitaSum + (BestDist * PopulationMatrix(j, 0))

    If DistanceMatrix(CInt(j & ClosestCode)) <= CInt(txtDist.Text) Then 'Put coverage total in Solution Set
        SolutionSet(ClosestFacilityNum, 2) = PopulationMatrix(j, 0) + SolutionSet(ClosestFacilityNum, 2)
        If j = ClosestCode Then
            Community2011TableAdapter.UpdateIsCovered(ClosestFacilityNum, j)
        Else
            Community2011TableAdapter.UpdateIsCovered(ClosestFacilityNum + 0.1, j)
        End If
    Else
        Community2011TableAdapter.UpdateIsCovered(ClosestFacilityNum + 0.2, j)
    End If
    SolutionSet(ClosestFacilityNum, 1) = PopulationMatrix(j, 0) + SolutionSet(ClosestFacilityNum, 1)
Next j
'---End Assign Community----------

'---Update Output Table-------------
lblProgress.Text = "Updating Population"
Application.DoEvents()
Dim PopTotal, PopCoverage As Integer

PopCoverage = 0
PopTotal = 0
For s = 1 To FacilityCount 'Calculate facility totals
    SampleDataSet.ResultsTable(s - 1).PopTotal = SolutionSet(s, 1)
    SampleDataSet.ResultsTable(s - 1).PopCoverage = SolutionSet(s, 2)
```

```
            PopTotal = PopTotal + SolutionSet(s, 1)
            PopCoverage = PopCoverage + SolutionSet(s, 2)
        Next s

        For s = 1 To FacilityCount 'Calculate facility statistics
            SampleDataSet.ResultsTable(s - 1).PercTotal = Math.Round((SolutionSet(s, 1) / PopTotal) * 100, 1)
            SampleDataSet.ResultsTable(s - 1).PercCoverage = Math.Round((SolutionSet(s, 2) / SolutionSet(s, 1)) * 100, 1)
        Next s

        'Populate summary table
        Dim newTotalsRow As SampleDataSet.ResultsTableRow
        newTotalsRow = Me.SampleDataSet.ResultsTable.NewResultsTableRow
        newTotalsRow.Code = "000"
        newTotalsRow.RddbName = "Coverage Information"
        newTotalsRow.PopTotal = PopTotal
        newTotalsRow.PopCoverage = PopCoverage
        newTotalsRow.PercTotal = "100"
        newTotalsRow.PercCoverage = Math.Round((PopCoverage / PopTotal) * 100, 1)
        Me.SampleDataSet.ResultsTable.Rows.Add(newTotalsRow)
        DataGridView2.Rows(FacilityCount).Selected = True

        'Calculate Per Capita Distance to Service (PCDS)
        PerCapita = PerCapitaSum / PopTotal
        Application.DoEvents()

        lblProgress.Text = "Complete - PCDS: " & Math.Round(PerCapita, 1) & " Min"
        GroupBox1.Enabled = True
        ProgressBar(0, 1)
        '---End Output Table----------------

    End Sub

    '---POPULATION UPDATE FOR SPATIAL EQUITY VARIANT----
    Public Sub UpdatePopWeightMatrix(ByVal FacCount As Integer)
        Dim ClosestFacility, ClosestDistance, TempDistance As Integer

        For pw = 1 To 999
            ClosestFacility = 0
            ClosestDistance = 9999
            TempDistance = 0
            'Find Distance to Closest Facility
            For ss = 1 To FacCount
                If SolutionSet(ss, 0) <> 0 Then
                    TempDistance = DistanceMatrix(CInt(pw & SolutionSet(ss, 0)))
                    If TempDistance < ClosestDistance Then
                        ClosestDistance = TempDistance
                        ClosestFacility = SolutionSet(ss, 0)
                    End If
                End If
            Next ss

            'Weight = Population * Distance ^ 1 – Equity Variable set to 1
            If ClosestFacility = 0 Then
                PopulationMatrix(pw, 3) = PopulationMatrix(pw, 0)
            Else
                PopulationMatrix(pw, 3) = PopulationMatrix(pw, 0) * (DistanceMatrix(CInt(pw & ClosestFacility)) ^ 1)
            End If
        Next pw
    End Sub


   '---UPDATE SUMMARY TABLE---
    Public Sub UpdateOutputTable(ByVal FacilityComm As Integer, ByVal FacilityScore As Integer, ByVal FacilityFixed As
Boolean)
        Dim newResultsRow As SampleDataSet.ResultsTableRow
        newResultsRow = Me.SampleDataSet.ResultsTable.NewResultsTableRow
        newResultsRow.Code = FacilityComm
```

```
        If FacilityFixed = False Then
            newResultsRow.RddbName = Community2011TableAdapter.GetCommName(FacilityComm)
        Else
            newResultsRow.RddbName = Community2011TableAdapter.GetCommName(FacilityComm) & "*"
        End If
        newResultsRow.Score = FacilityScore
        Me.SampleDataSet.ResultsTable.Rows.Add(newResultsRow)
        DataGridView2.Rows(0).Selected = False
        Application.DoEvents()
    End Sub

    '---HANDLE SPATIAL EQUITY CHECKBOX---
    Private Sub chkSpatialEquity_CheckedChanged(sender As Object, e As EventArgs) Handles
chkSpatialEquity.CheckedChanged
        If chkSpatialEquity.Checked = True Then
            chkDistance.Checked = True
        Else
            chkDistance.Checked = False
        End If
    End Sub

    '---HANDLE HOSPITAL CANDIDATES CHECKBOX---
    Private Sub chkCandidates_CheckedChanged(sender As Object, e As EventArgs) Handles
chkCandidates.CheckedChanged
        If chkCandidates.Checked = True Then
            Me.Community2011TableAdapter.SetCandidatesToAll(0)
            Community2011TableAdapter.SetCandidatesToHospitals()
        Else
            Me.Community2011TableAdapter.SetCandidatesToAll(1)
        End If
    End Sub

End Class
```