THE EFFECT OF VARIABLE DEMAND AGGREGATION AREAS ON
MODELING PRIMARY HEALTH CARE ACCESSIBILITY IN IDAHO

A Thesis
Presented to the
Faculty of
San Diego State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Geography

by
Blake Oliver Brown
Fall 2010
The Undersigned Faculty Committee Approves the

Thesis of Blake Oliver Brown:

The Effect of Variable Demand Aggregation Areas on Modeling Primary Health Care Accessibility in Idaho

Piotr Jankowski, Chair
Department of Geography

John R. Weeks
Department of Geography

Eric G. Frost
Department of Geology

11-15-2010
Approval Date
Copyright © 2010

by

Blake Oliver Brown

All Rights Reserved
ABSTRACT OF THE THESIS

The Effect of Variable Demand Aggregation Areas on Modeling Primary Health Care Accessibility in Idaho
by Blake Oliver Brown
Master of Science in Geography
San Diego State University, 2010

This study examines the accessibility of primary health care services to rural populations. Specifically, the data used in this study are from the state of Idaho, which provides a good example of a largely rural area striving to provide adequate primary health care accessibility for many of its residents.

In this study, demand aggregation was incorporated into an a Maximal Covering Location Problem (MCLP) accessibility model in order, first, to determine the unmet demand for primary health care services in Idaho, based on the allocation of Idaho residents’ demand for primary health care to the state’s existing primary health care facilities (PHCF). This procedure identified areas without adequate accessibility to PHCF (locations where demand for Primary Health Care Services (PHCS) is unmet). Secondly, the study used this model to determine potential locations (from a set of feasible sites) that could meet unmet demand determined by the initial models. In addition, the study explored the consequences of different demand locations and different driving time constraints on all of these models’ results.

The representation of demand in location/allocation accessibility modeling is an important consideration in making accurate determinations of the location of demand since different representations of demand have different consequences with regard to model results.

Four different demand representations (tract centroids, block group centroids, aggregated block cluster centroids, and weighted aggregated block cluster centroids) and three different driving time constraints (30, 45, and 60 minutes) were used to run the models. This resulted in a total of 12 deterministic models which indicate where demand for PHCS is unmet. These same demand representations and driving time constraints were then used to run 12 predictive models which evaluate the fitness of 20 candidate sites to meet the unmet demand (as determined by the results of the deterministic models). Lastly, one final predictive model was run using only the most robust candidate site(s) from the 20 potential candidate sites. This final model used the weighted aggregated block demand representation at the 30 minute driving time constraint to make a final determination of the candidate site(s) that would maximally alleviate the unmet demand for PHCS.

In general, as the models moved from demand representations using tract to block group to aggregated block, the amount of unmet demand for PHCS decreased. The tract model yielded the greatest amount of unmet demand; The aggregated block model yielded the lowest amount (except at the furthest driving time constraint); and the weighted aggregated block model yielded slightly greater amounts of unmet demand than the non-weighted aggregated block model and yielded the least amount of unmet demand at the
furthest driving time constraint. The predictive models showed that only nine of the 20 potential candidates sites could alleviate any of the unmet demand for PHCS. Furthermore, they indicated that of these nine sites, only two could alleviate a substantial amount of unmet demand.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xiii</td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 RESEARCH CONTEXT</td>
<td>4</td>
</tr>
<tr>
<td>Location/Allocation Accessibility Modeling</td>
<td>4</td>
</tr>
<tr>
<td>The Maximal Covering Location Problem (MCLP) Model</td>
<td>5</td>
</tr>
<tr>
<td>Demand Representation</td>
<td>6</td>
</tr>
<tr>
<td>3 RESEARCH QUESTIONS AND HYPOTHESES</td>
<td>11</td>
</tr>
<tr>
<td>Problem Statement</td>
<td>11</td>
</tr>
<tr>
<td>Research Questions</td>
<td>11</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>12</td>
</tr>
<tr>
<td>4 METHODOLOGY</td>
<td>13</td>
</tr>
<tr>
<td>Models Used in Study</td>
<td>15</td>
</tr>
<tr>
<td>Data</td>
<td>16</td>
</tr>
<tr>
<td>Pre-Existing Data</td>
<td>16</td>
</tr>
<tr>
<td>Derived Data</td>
<td>17</td>
</tr>
<tr>
<td>Data Preprocessing</td>
<td>18</td>
</tr>
<tr>
<td>Demand Preprocessing</td>
<td>18</td>
</tr>
<tr>
<td>Track Demand</td>
<td>20</td>
</tr>
<tr>
<td>Block Group Demand</td>
<td>20</td>
</tr>
<tr>
<td>Aggregated Block Demand</td>
<td>21</td>
</tr>
<tr>
<td>Weighted Aggregated Block Demand</td>
<td>21</td>
</tr>
<tr>
<td>Supply Preprocessing</td>
<td>22</td>
</tr>
<tr>
<td>MCLP Model Execution</td>
<td>22</td>
</tr>
<tr>
<td>Initial Road Network Construction</td>
<td>22</td>
</tr>
</tbody>
</table>
Connecting Demand Points ............................................................22
Preparing the Supply Covers ..........................................................23
Preparing the Road Network to Run the MCLP Model .....................24
Running the MCLP Location-Allocation Models ............................24
Network Environment Setup .........................................................24
Location-Allocation Environment Setup .........................................25
Solving the Location-Allocation Problem .......................................25
Manual Road Network Correction/Connection ..................................26
Predictive Models ..........................................................................27
Candidate Primary Health Care Facility Selection ............................27
Final Candidate Primary Health Care Facility Selection ....................28

5 RESULTS ..................................................................................29

Deterministic Models (First 12 Models) ..........................................29
Tract Models ..................................................................................29
Block Group Models .....................................................................32
Aggregated Block Model ...............................................................33
Weighted Aggregated Block Model ................................................37

Predictive Models (Models 12 – 24) ...............................................41
Tract Predictive Models ..................................................................45
Tract Predictive Model 24,140m .....................................................45
Tract Predictive Model 36210m ......................................................45
Tract Predictive Model 48280m ......................................................46
Block Group Predictive Models ......................................................46
Block Group Predictive Model 24140m ..........................................47
Block Group Predictive Model 36210m ..........................................47
Block Group Predictive Model 48280m ..........................................48
Aggregated Block Predictive Models ..............................................48
Aggregated Block Predictive Models 24140m ..................................48
Aggregated Block Predictive Models 36210m ..................................48
Aggregated Block Predictive Models 48280m ..................................49
Weighted Aggregated Block Predictive Models ............................50
Weighted Aggregated Block Predictive Model 24140m ....................50
Weighted Aggregated Block Predictive Model 36210m ....................50
Weighted Aggregated Block Predictive Model 48280m .......................51
Final Predictive Weighted Aggregated Block Model (Model 25) ..................51
Analysis of the Coverage Pattern .....................................................53
Moran’s I Analysis .............................................................................53
Hot Spot Analysis with Rendering .....................................................55

6 DISCUSSION ..................................................................................60
Key Findings .......................................................................................60
Limitations ..........................................................................................63

7 CONCLUSION ................................................................................64

8 FUTURE RESEARCH .......................................................................65

REFERENCES .....................................................................................67

APPENDICES
A  AML SCRIPT USED TO EXECUTE MODEL ........................................70
B  AML SCRIPT USED TO DISPLAY MODEL RESULTS ..........................73
C  EXISTING SUPPLY (PHCF) CAPACITANCE ANALYSIS PRIMARY
   HEALTH CARE FACILITY OVERCAPACITATION .................................76
D  DASYMETRIC TECHNIQUES ..........................................................81
LIST OF TABLES

Table 1. Deterministic Tract Model Results .................................................................30
Table 2. Deterministic Block Group Model Results .........................................................33
Table 3. Deterministic Aggregated Block Model Results .................................................35
Table 4. Deterministic Weighted Aggregated Block Model Results .................................41
Table 5. Candidate Site Evaluation for Tract Predictive Model at 24140 Meters Driving Distance ..............................................................................................................46
Table 6. Candidate Site Evaluation for Tract Predictive Model at 36210 Meters Driving Distance ..............................................................................................................46
Table 7. Candidate Site Evaluation for Tract Predictive Model at 48280 Meters Driving Distance ..............................................................................................................46
Table 8. Candidate Site Evaluation for Block Group Predictive Model at 24140 Meters Driving Distance ........................................................................................................47
Table 9. Candidate Site Evaluation for Block Group Predictive Model at 36210 Meters Driving Distance ........................................................................................................47
Table 10. Candidate Site Evaluation for Block Group Predictive Model at 48280 Meters Driving Distance .............................................................................................48
Table 11. Candidate Site Evaluation for Aggregated Block Predictive Model at 24140 Meters Driving Distance ..........................................................................................49
Table 12. Candidate Site Evaluation for Aggregated Block Predictive Model at 36210 Meters Driving Distance ..........................................................................................49
Table 13. Candidate Site Evaluation for Aggregated Block Predictive Model at 48280 Meters Driving Distance ..........................................................................................49
Table 14. Candidate Site Evaluation for Weighted Aggregated Block Predictive Model at 24140 Meters Driving Distance .................................................................50
Table 15. Candidate Site Evaluation for Weighted Aggregated Block Predictive Model at 36210 Meters Driving Distance ..............................................................................51
Table 16. Candidate Site Evaluation for Weighted Aggregated Block Predictive Model at 48280 Meters Driving Distance ..............................................................................51
Table 17. Peak Moran’s I Index Values and the Corresponding Distance Band Values for all Four Demand Representation Models ..........................................................54
Table 18. Weighted Aggregated Block Model at 24140 Meters ........................................79
Table 19. Weighted Aggregated Block Model at 36210 Meters ........................................79
Table 20. Weighted Aggregated Block Model at 48280 Meters
LIST OF FIGURES

PAGE

Figure 1. Demand representation techniques.................................................................7
Figure 2. Schematic showing the data required to run the MCLP model...........................15
Figure 3. Demand representations used in this study ......................................................19
Figure 4. Tract model: Distance vs. demand. .................................................................30
Figure 5. Deterministic tract model coverage.................................................................31
Figure 6. Block group model: Distance vs. demand. .......................................................32
Figure 7. Deterministic block group model coverage......................................................34
Figure 8. Aggregated block model: Distance vs. demand. .............................................35
Figure 9. Deterministic aggregated block model coverage............................................36
Figure 10. Centroid shift from non-weighted aggregated block model to weighted
aggregated block model. ..............................................................................................38
Figure 11. Coverage change from non-weighted aggregated block model to weighted
aggregated block model. .............................................................................................39
Figure 12. Plot showing centroid shift distance vs. the number of visits per year. ............40
Figure 13. Weighted aggregated block model: Distance vs. demand. ..............................40
Figure 14. Deterministic weighted aggregated block model coverage............................42
Figure 15. Location of three clusters of the 20 candidate supply sites (PHCFs). ...............43
Figure 16. Ten candidate sites located in northern Idaho shown in box A in Figure 15. ........................………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………….
Figure 25. Results of the deterministic models.................................................................61
Figure 26. Travel time vs demand (visits per year) for all four demand representation
models..............................................................................................................................62
Figure 27. Dasymetric mapping method.............................................................................84
ACKNOWLEDGEMENTS

I would like to thank Dr. Piotr Jankowski for his guidance and patience throughout this process. I would also like to thank Dr. Weeks and Dr. Frost for their support. Thanks also goes to my fellow students and friends for assistance and moral support throughout my time at SDSU. Finally, I would like to thank wife and my parents whose love and guidance over the years made this accomplishment possible.
CHAPTER 1

INTRODUCTION

In the U.S., access to medical facilities is considered vital not only to individuals needing medical care, but also to the communities in which these individuals participate. The federal guideline for adequate access to primary care services states that all individuals should reside within a 30-minute driving distance (roughly 15 miles) from a primary health care facility (PHCF) (U.S. Department of Health and Human Services 1993). While this is a laudable goal, it is often difficult to achieve.

Previous studies have indicated that access to primary health care, while generally available in urban areas, is not as available in many rural areas of the U.S. (Jankowski 1995). The low population density of most rural areas, along with a low patient to doctor ratio, results in a large percentage of the rural population residing further than 30 minutes from a health care facility. The resulting inequality in health care services between urban and rural residents has been a matter of concern to federal and state health officials.

The state of Idaho is a good example of the difficulties of providing access to health care equally to all residents. In 2008, Idaho had the second lowest active physician to population ratio of all (50) states and the third lowest active primary care physician to population ratio, with 181.8 active physicians per 100,000 people, and 65.8 active primary care physicians per 100,000 people. These ratios were far below the national rates of 254.5 active physicians per 100,000 people and 89.6 active primary care physicians per 100,000 people (Center for Workforce Studies 2009). The low physician to population ratios might be explained partially by the low population density within the state. In 2009, Idaho was estimated to have a population density of 18.68 persons per square mile, compared to the estimated national average of 86.8 persons per square mile (U.S. Census Bureau 2009). In order to improve health care accessibility in rural areas like Idaho, it is necessary to know where under-serviced areas exist, or in other words, where demand for healthcare is unmet.

Location-allocation accessibility modeling can be used to determine where demand for primary health care services (PHCS) is unmet. This modeling technique can also be used
to determine optimal locations for potential new health care facilities that would reduce unmet demand in rural areas. Among existing accessibility models is the location/allocation Maximal Covering Location Problem (MLCP) model (available as a tool in ArcInfo Workstation). This model can be run by providing locations of supply (PHCF) and locations of population demand centers both of which need to be interconnected on a road network.

A limitation of this modeling technique is that the locations of demand are usually required to be in the form of points. In these models demand points are typically positioned at the centroids of the census enumeration units used in the accessibility model. In the U.S., census information is available at the following levels of enumeration: from largest to smallest they are national, state, county, tract, block group, and block. In the accessibility models, the smaller the spatial enumeration unit used to derive demand, the more accurate the representation. However, algorithms employed to solve location/allocation models aimed at assessing a state’s accessibilities to various services such as health care, cannot solve the models with the number of demand locations that would result from locating demand at the center of the state’s census blocks (the smallest census enumeration unit); in the case of Idaho, there are 37,740 populated censes blocks.

In an effort to reduce the number of demand locations to a level that will allow the model to run, the centroids of block groups (the second smallest census enumeration unit) could be used as the location of demand. However, the drawback to this approach is that (especially in rural areas) many block groups include large areas that are unpopulated (unpopulated blocks within a block group). By simply using the centroids of the state’s block groups, information regarding unpopulated areas within the block group is ignored, thus, increasing the risk that a centroid demand point will be located in an unpopulated area of the block group. The incorporation of demand aggregation into accessibility models has two benefits. First, it reduces the number of demand locations to a level that the model can run. Second, it gives those models the potential to more accurately represent the locations of populated areas (demand) within a census enumeration unit (block groups) rather than using the centroid of the block group enumeration unit. The incorporation of demand aggregation, guided by the presence of population, into accessibility modeling will likely result in a more accurate modeling of accessibility.
In this study, demand aggregation was incorporated into an accessibility model in order, first, to estimate the unmet demand for primary health care services in Idaho, based on the allocation of Idaho residents’ demand for primary health care, to the state’s existing primary health care facilities. This method identified areas without adequate accessibility (given the driving distance constraint) to PHCF. Secondly, the study used this model to determine optimal locations (from a set of potential sites) for attracting new primary health care providers, in order to maximize the coverage of health services available to people in those locations, which (as determined by the model) have unmet demand for PHCS. In addition, the study explored the consequences of different demand representations (tract centroid, block group centroid, etc.) on the amount of unmet demand. Finally this study explored the sensitivity of model results to different driving distances, other than 30 minutes driving distance, by rerunning all the models at 45 and 60 minutes travel times.
CHAPTER 2

RESEARCH CONTEXT

Since location/allocation accessibility modeling is the methodology used in this research, the basic principles and techniques of these models are reviewed here. Special attention is given to the Maximal Covering Location Problem location/allocation model. In addition, the earlier conceptualizations, generation techniques, and uses of demand aggregation are also reviewed.

LOCATION/ALLOCATION ACCESSIBILITY MODELING

Location/allocation modeling deals with the problem of how to optimally locate facilities to provide services and/or amenities in the most efficient manner to entities demanding services and/or amenities. These models can have different goals; for example, it may be desirable to select service supply locations that minimize the total distance traveled between a given supply and demand point. Or, it may be desirable to select supply locations that will maximize services available to demand locations within some specified distance. Location/allocation modeling allows researchers to answer certain spatial questions: What are the fewest number of facilities needed to service 100% of a region’s demand, while ensuring that no area of demand is located further from the nearest service facility than a user specified distance? Or, where should a facility be sited to minimize the overall distance traveled by populations accessing the facility? Location/allocation modeling in a GIS is often used to investigate and monitor access needs (Daskin and Dean 2004). It can also be used to model the effects on accessibility levels when services are relocated or when service delivery is changed (Langford and Higgs 2006). For more information on location-allocation covering problems, Schilling, Jayaraman, and Barkhi (1993) provides a detailed overview of covering problems by reviewing 96 papers from 37 different journals.
In this study, the allocation of demand for health care to PHCF locations was performed by implementing a location/allocation model run in Arc/Info 9.2 GIS Software (ESRI 1996 ARC/INFO). Specifically, the Maximal Covering Location Problem (MCLP) model (Church and ReVelle 1974) was utilized. This model is designed to maximize the demand assigned to a selected number of supply sites within a specified distance. Traditionally, the model has been used to find an optimal subset of supply sites from the set of all possible supply sites (Oppong and Hodgson 1994; Gerrard et al. 1997). Since, for this study, it is necessary to incorporate all existing PHCF into the allocation of demand, the model treated all the existing PHCF as fixed, thus “forcing the model algorithm to locate all centers and then assign the demand to centers such that the total assigned demand is maximized” (Jankowski 1995).

In the MCLP model, the network nodes represent locations of demand for services. Some nodes can also be used to represent candidate locations for service centers. The model allows for a node to be both a demand location and candidate facility site. The network arcs represent linkages between the nodes. Each of the network arcs has some defined impedance (e.g. travel time or distance) that represents the separation between nodes. The total impedance between each demand node and each candidate facility node is defined as the total impedance encountered between the two nodes while traveling the shortest distance over the network. A demand node (location) is considered covered if it is within some user specified distance to a supply node.

ArcInfo Workstation is capable of solving the MCLP model with the use of heuristic techniques as opposed to the use of linear programming. A heuristic method starts with a random initial configuration of facilities (supply sites) and then finds desirable substitutions of candidate facilities in the solution, and this continues until there can be no further improvement. It is worth noting that, while the substitution heuristics are very successful at solving different optimal location problems, their use does not guarantee the optimal result (Church and Sorenson 1996). Two heuristics can be used to solve the MCLP model: the Teitz and Bart (1968) or the Global Regional Interchange Algorithm (GRIA) heuristic of Densham and Rushton (1992). It is worth noting that when candidate facilities’ nodes and
demand nodes are entirely different from each other, due to intricacies in heuristic procedures, the Teitz and Bart heuristic is expected to perform better than the GRIA (ESRI 1996b ARC/INFO).

The MCLP model has frequently been applied to the problem of optimally siting emergency facilities needed to cover population areas in cities (Gerrard et al. 1997). Oppong and Hodgson (1994) used the MCLP location/allocation model to develop a solution for improving accessibility to health facilities in rural Ghana. They found that the relocation of particular mobile facilities allowed for improved access without having to build new facilities. Further examples of MCLP model applications can be found in the survey by Schilling, Jayaraman, and Barkhi (1993) who identified 65 papers on the optimization of location coverage.

Because location/allocation modeling in a GIS is well equipped to model spatial accessibility, it has become “an important policy tool for managing health care provisions and reducing health inequality,” (Langford and Higgs 2006, p.294). The ability to make spatially intelligent decisions with regard to providing equitable access to health care facilities is especially valued by planners and policy makers in rural areas where low population density and highly dispersed settlements generally exist.

Medical accessibility can be investigated from two different perspectives. First, potential accessibility measures (in the case of the proposed study, the number of people who could reach health care facilities within a 30-minute driving time) and second, realized accessibility measures (how many people actually received medical services). With regard to these studies (including this one), this detailed patient-level information is typically unavailable. As a result, most accessibility studies have developed potential measures of access based on straight-line distances or travel-time distances between health services and demand points. These measures are then used to identify areas where accessibility is poor and where additional health facilities are needed to improve accessibility (Langford and Higgs 2006).

**Demand Representation**

In order to utilize location/allocation modeling, the modeler must prepare data on where population (demand) is located and the amount of that population’s demand. There
are different approaches to representing the location of the population demand, but in most location/allocation models, population demand must be represented as a point (Rushton 1989). Because the population data are almost always available only as the total population contained within an enumeration unit, most accessibility studies use the centroid of the enumeration unit to represent that enumeration unit’s population (Figure 1a). It is also possible to distribute the population evenly across an enumeration unit as a series of evenly spaced points dispersed throughout the entire enumeration unit (Figure 1b) (Langford and Higgs 2006). If uniform distribution is a valid assumption, the pro rata method may offer more accurate estimates of population location than the centroid method. This assumption may be valid in small dense urban environments but is less likely to be valid in rural areas where population tends to concentrate into small settlements separated by areas of unoccupied/unpopulated land (Langford et al. 2008). In addition, this pro-rata technique for representing population locations typically results in a vast number of demand points which exceed the maximum limit of demand points that most accessibility models can accommodate.

![Figure 1. Demand representation techniques.](image)

Daskin et al. (1989) identify two sources of aggregation errors. The first type occurs as a result of converting continuous demand distribution (zones) into discrete representations of demand (points). Hillsman and Rhoda (1978) discuss in detail three types of errors that
arise from this conversion process. The first type of error (type A error) is a representation error that arise when the distance between an aggregated point and a supply center is used as the ‘true’ average distance between the enumeration unit and the supply center. This eliminates overestimations of the true average distance and results in negative source A errors (an under-estimation of the true average distance between the spatial unit and the supply center). Source B errors are a specific type of source A errors and occur when a supply center exists at an aggregation point. Distance from the aggregated point to the supply will be zero, but realistically the distance value must be greater than zero. Source C errors result from the A and B representation errors. Source C errors occur when an entire enumeration unit is allocated to one supply center even though some portion of that enumeration unit is closer to another supply center. Hillsman and Rhoda (1978) note that these conversion errors can result in: inaccurate objective function values, misallocation of demand and or, an incorrect/non-optimal choice of facility locations in the location-allocation model.

The second source of aggregation errors identified by Daskin et al. (1989) results when a subset of demand nodes is used in the model instead of the full set. This subset would be created to reduce the number of demand nodes used in the model to a computationally manageable amount.

Some examples of different demand representation techniques are reviewed in the following paragraphs. In a study by Verter and Lapierre (2002), the centroids of 638 populated regions were used as the demand locations for an MCLP model locating preventative health care facilities that maximized participation in prevention programs with the rationale that distance is a major determinant of participation in preventative programs.

In a study by Mitropoulos et al. (2006), patient level data about the annual number of visits to existing health care facilities was obtained for all inhabitants of 228 population regions in semi-rural Achaia, Greece. The centroids of these regions were used as the locations of demand for health care facilities.

Brabyn and Skelly (2002) used the centroids of meshblocks in New Zealand (meshblocks are the most detailed census enumeration units available) as the locations of demand in an accessibility model. This study examined the accessibility of the population
centers not only to primary health care facilities but also secondary and tertiary health care facilities.

If demand points are located at nodes that do not already exist on the road network, then it is necessary to connect them to the road network. This is usually done by creating an arc/line that connects each demand centroid/node to the nearest location on the road network. These connection lines are then appended to the original road network. Each and every demand point must be capable of traveling to each and every supply site over the road network before the MCLP model is ready to run. But, in models that contain a road network with numerous well dispersed nodes, it may be desirable to assign the demand of an enumeration unit to an existing network node that is located near to where the enumeration unit’s centroid would exist. This approach saves the time and effort required to connect the demand centroids (nodes) to the road network. This “nearest node to the centroid” technique was used in the 2004 study by Brabyn and Barnett to assess access to general practitioners in rural New Zealand. In addition, the study was able to obtain data on the precise locations of all addresses in New Zealand. Using these data, the mean center of the addresses within each meshblock was found and used as the location of that meshblock’s demand.

In a study by Wang and Luo (2005), population-weighted centroids of census tracts (based on block-level population data) were used instead of simple geographic centroids, to represent population locations more accurately. This process resulted in a computationally manageable number of demand points (2952) and was particularly successful in refining the locations of population in rural areas where large areas of land are unpopulated.

Similar to the method used by Wang and Luo (2005), this study used the population-weighted centroids (based on block-level population data) of contiguous census block clusters within block groups. This technique represents population locations more accurately than can be achieved with the use of the simple geographic centroids of the census block groups. This technique was particularly desirable for this study area because more than half the land area of Idaho is unpopulated.

Although there are many novel approaches to representing the continuous demand within an enumeration unit, Goodchild (1979) notes that the most desirable aggregation scheme depends on the solution, which is unknown a priori. Hence, he concludes that there exists no uniformly desirable aggregation scheme.
The incorporation of demand aggregation in location/allocation modeling of health care demand provides the opportunity to more accurately determine areas where the population lacks adequate accessibility to medical facilities. The Maximal Covering Locational Problem model is particularly suited to identify the locations where demand for healthcare is unmet. Techniques that allow demand to be represented and positioned more accurately in MCLP models would be of particular interest to agencies that are involved in facilitating the delivery of health care services, in this case, the Idaho Department of Health and Welfare.
CHAPTER 3

RESEARCH QUESTIONS AND HYPOTHESES

PROBLEM STATEMENT

Accessibility modeling can be a valuable tool for decision makers and planners in charge of providing PHCS. An important aspect of accessibility modeling is the choice of how demand is both determined and located. Previous research has shown that the choice of demand representation in accessibility models can have a powerful influence on model results. The more accurately the demand is represented in the model, the more reliable are the results. Idaho, which is known to have poor accessibility to PHCF, especially in rural areas, will be used in this study to investigate this relationship between the demand representations and the accessibility models’ results.

The objective of this study is twofold. First, it is to observe the change in the amount of unmet demand computed by accessibility models that use different representations of demand locations. The second objective is to evaluate a number of potential candidate sites based on their ability to alleviate unmet demand found in the corresponding deterministic models. In addition, the sensitivity of model results to changes in the driving distance constraint will be explored.

RESEARCH QUESTIONS

This study addresses the following research questions:

1. What effects do changes in the representations of demand aggregation in the location/allocation MCLP model, aimed at examining the accessibility of primary health care services to Idaho residents, have on the model results?

2. What effects do changes in the representations of demand locations in the location/allocation MCLP model, aimed at selecting new PHCF (from a set of potential sites), have on the model results?
**Hypotheses**

**Null Hypothesis 1:** No change in the model results (amount of unmet demand) will be obtained from using different representations of demand aggregations when determining Idaho residents’ access to primary health care facilities using the MCLP location/allocation model. In particular:

1 a. No change in the model results will be obtained using the 30-minute driving distance constraint.

1 b. No change in the model results will be obtained using the 45-minute driving distance constraint.

1 c. No change in the model results will be obtained using the 60-minute driving distance constraint.

**Null Hypothesis 2:** No change in the model results (best candidates sites) will be obtained from using different representations of demand aggregations when determining new candidate sites that would maximally alleviate unmet demand for primary health care services experienced by Idaho residents. In particular:

2 a. No change in the model results will be obtained using the 30-minute driving distance constraint.

2 b. No change in the model results will be obtained using the 45-minute driving distance constraint.

2 c. No change in the model results will be obtained using the 60-minute driving distance constraint.
CHAPTER 4

METHODOLOGY

The Maximal Covering Location Problem Model was used for all the models run in this study. The maximal covering location problem aims to maximum the population that can be covered within a user chosen service distance or time given a finite number of facilities. A mathematical formulation of the maximal covering location problem can be stated as follows:

Maximize \[ z = \sum_{i=1}^{I} a_i \cdot y_i \]

Subject to:
1) \[ \sum_{j \in N_i} x_j \geq y_i \] for all \( i \in I \) { This constraint allows \( y_i \) to equal 1 (node \( i \) is covered) only when one or more facilities are established at sites in the set \( N_i \) (meaning, one or more facilities are located within \( S \) distance units of demand point \( i \)).

2) \[ \sum_{j \in J} x_j = P \] { This constraint restricts the number of facilities to \( P \).

3) \[ x_j = (0,1) \] for all \( j \in J \) { Coverage decision variable

4) \[ y_i = (0,1) \] for all \( i \in I \) { Coverage decision variable

Where:
\( I \) = the set of demand nodes;
\( J \) = the set of facility sites;
\( S \) = the distance beyond which a demand point is considered “uncovered” (the value of \( S \) can be chosen differently for each demand point if desired);
\( d_{ij} \) = the shortest distance from node \( i \) to node \( j \);
\( x_j \) = { 1 if a facility is allocated to site \( j \), 0 otherwise;

\( N_i \) = { \( j \in J \mid d_{ij} \leq S \} \); the set of facilities \( j \) that can reach node \( i \) within the maximal service distance \( S \)
\[ a_i \] = population to be served at demand node \( i \);

\( p \) = the number of facilities to be located.

The equations and explanation of the MCLP model are taken from Church and ReVelle (1974).

To run the MCLP models, first, the input data needed to be preprocessed. Most of the preprocessing involved the representation of demand. Next, the network environment was set up using the ArcInfo Workstation 9.2 command line. This was followed by setting up the location-allocation environment. Then, the location allocation problem was solved and the results were analyzed.

In this study four different MCLP models were used to address the first research question, each with a different representation of PHCS demand (tract, block group, aggregated block, and weighted aggregated block). Each of these four models was eventually run using three different driving distance constraints to determine the amount and location of unmet demand for health care services in Idaho. The 12 resulting model versions that used the 78 pre-existing PHCF sites will be referred to as the deterministic models. The 12 model versions that use the 78 pre-existing sites and additionally the 20 potential candidate sites will be referred to as the predictive models. One additional predictive model version was also run, for a total of 25 model versions.

First, the deterministic models were run with each of the four representations of demand and each of the three driving distance constraints. The amount of unmet demand was thus determined for each of the three driving constraints: 15 miles/24,140 meters/30 minutes; 22.5 miles/36,210 meters/45 minutes; and 30 miles/48,280 meters/60 minutes.

Next, each of these 12 deterministic models was rerun using a supply file that included not only the existing 78 PHCF, but also the 20 potential candidate PHCF sites. Rerunning the deterministic models with this new supply file made it possible to evaluate these 20 new candidate (predictive) sites’ ability to service the PHCS demand that was determined to be unmet by the corresponding deterministic model. Re-running the predictive models with different driving distance constraints was done with the rationale that there may be new potential PHCF sites that could service much of the unmet demand within, for example, 20 miles, but none of the unmet demand within 15 miles. If such were the case, it
might be determined preferable to utilize new PHCF at locations like these, rather than omitting them, even though they are further than the federal guideline of 30 minutes driving distance.

Finally, in order to gain a sense of the reduction in unmet demand for PHCS that could be obtained with the inclusion of a few new candidate sites, a final predictive model (the 25th model) was run that included only the most robust candidate PHCF site(s) in the supply file (not all 20 sites).

**MODELS USED IN STUDY**

All models were prepared in ArcMap 9.2 and run in ArcINFO/Workstation 9.2. The allocation of demand for health care to PHCF locations was performed using a location/allocation model run in Arc/Info 9.2 GIS Software (ESRI 1996 ARC/INFO). Specifically, the Maximal Covering Location Problem (Church and Revelle 1974) model was utilized. This model is designed to maximize the demand assigned to a selected number of supply sites within a specified distance. To run the MLCP model, four different coverage/data files are required: (1) a road network in an ArcInfo line coverage file, (2) an ArcInfo polygon coverage file representing the boundaries of service areas, (3) supply locations in an ArcInfo point coverage, and (4) demand locations in an ArcInfo point coverage (see Figure 2). For the models to run correctly, all demand locations must be connected to the road network. The first phase of the road network construction is achieved using a standard set of commands. The complete connection of the demand locations to the road network sometimes requires manual editing using ArcEdit and ArcTable. The amount of manual editing increases as the number and proximity of demand points present in the model increases. Because of the time required to manually edit the road network, the number of manual edits that were required to completely connect the road network was the most limiting factor in the ability to run models with many demand locations.

![Figure 2. Schematic showing the data required to run the MCLP model.](image-url)
Four different representations of demand were used in this study. First, the ‘tract models’ were run using the centroid of each census tract as the location of that tract’s demand for PHCS. Second, the ‘block group models’ were run using the centroid of each populated census block group as the location of that block group’s demand for PHCS. Third, the ‘Aggregated Block Models’ were run using the centroids of contiguous populated blocks within each census block group to represent the locations of demand for PHCS. Finally, the ‘Weighted Aggregated Block Models’ were run using the weighted (based on demand) centroids of contiguous populated blocks within each census block group to represent the locations of demand for PHCS.

**DATA**

The data used in this study consists of pre-existing data and derived data. Most of the derived data were created by combining pre-existing data.

**Pre-Existing Data**

a. *Block .shp (shape) file.* The shape file of all census blocks in Idaho from 2000. This file (obtained from the ESRI DataDisk2000) was used (in conjunction with other data) to determine demand point locations for the Aggregated Block Model, the Weighted Aggregated Block Model, and the Final Predictive Weighted Aggregated Block Model.

b. *Block age and sex data.* U.S. Census 2000 Summary File 1: Table P12 Sex by Age. The file consists of a table showing the population of every census block in Idaho based on sex and 23 different age groupings (U. S. Census Bureau 2000).

c. *Block Group .shp (shape) file.* The shape file of all populated census block groups in Idaho from 2000. This file (obtained from the ESRI DataDisk2000) was used (in conjunction with other data) to determine demand point locations for the Block Group Model.

d. *Block Group age and sex data.* U.S. Census 2000 Summary File 1: Table P12 Sex by Age. The file consists of a table showing the population of every census block group in Idaho based on sex and 23 different age groups (U. S. Census Bureau 2000).

e. *Tract .shp (shape) file.* The shape file of all census tracts in Idaho from 2000. This file (obtained from the ESRI DataDisk2000) was used (in conjunction with other data) to determine demand point locations for the Tract Model.

f. *Tract age and sex data.* U.S. Census 2000 Summary File 1: Table P12 Sex by Age. The file consists of a table showing the population of every census tract in Idaho based on sex and 23 different age groups (U. S. Census Bureau 2000).
g. *Rate of visits file.* Estimated rate of primary health care facility visits based on age and sex (Center for Disease Control 2003).

h. *Primary care service areas (PCSA) data.* Population distribution in Idaho while taking into account natural terrain barriers and travel patterns, obtained from an existing study (Jankowski 1995). This file was used in all models.

i. *Road network file.* ArcInfo line coverage file representing the major roads in Idaho. This coverage was created during the previous study conducted by Jankowski in 1995. This file was used in all models.

j. *Supply coverage file.* A point coverage file, found in Jankowski’s 1995 study, contains information regarding the location of all practitioners and the number of primary care hours they can provide. The estimated number of visits per year was then calculated for all facilities using the U.S. Public Health Service standard of 4200 visits per year for primary care physicians and 2100 visits per year for midlevel providers (Jankowski 1995). In this file is a *candidate* attribute that specifies how the supply site is treated when the maxcover model is run (not a candidate = 0, a possible candidate = 1, a fixed/existing site = 2). This file was used in all deterministic models. This file was also combined with the candidate supply sites for use in the predictive models.

k. *Candidate supply sites.* This file contains potential locations to site new PHCFs. These additional locations were selected from Idaho’s 203 census places because it was determined that these places meet minimum sufficient infrastructure requirements. Only locations that have at least one physician and at least one medical facility in the area were included from these 203 census places. Some of the candidate sites are located very near the same locations as some of the existing facilities. Each candidate location was represented as having one new full-time primary health care provider. Thus, all the candidate sites’ supply values were assigned 4,200 visits per year in accordance with the U.S. Public Health Service standard of 4,200 visits per year for a primary care provider (Jankowski 1995).

l. *aml file.* From Jankowski’s 1995 study. This .aml file is used to implement the commands needed to setup the network environment and the location-allocation environment and then to solve and display the results of the model. (Appendix A).

**Derived Data**

m. *Tract demand file.* To run the Tract Model, a point coverage representing the demand locations of Idaho residents was created. This coverage locates demand at the centroid of the census Tract. For detail on the creation of the Tract demand file see section 4.3 Tract Demand.

n. *Block Group demand file.* To run the Block Group Model, a point coverage representing the demand locations of Idaho residents was created. This coverage places demand at the centroid of the Census Block Group. For details on the construction of the Block Group demand file see section 4.3 Block Group Demand.
o. **Aggregated Block demand file.** An Aggregated Block demand file was needed to run the deterministic and predictive Aggregated Block Models. The Aggregated Block demand file locates demand at the centroids of contiguous regions within the Census Block Groups. This process resulted in a demand file that is able to capture some of the detail offered by the block demand without having so many demand points that the model becomes intractable. For details on the creation of the Aggregated Block demand file see section 4.3 Aggregated Block Demand.

p. **Weighted Aggregated Block demand file.** A Weighted Aggregated Block demand file was needed to run the deterministic and predictive Weighted Aggregated Block models, as well as the Final Predictive Weighted Aggregated Block Model. The creation of this file was similar to the creation of the Aggregated Block Demand file. The Weighted Aggregated Block demand file locates demand at the weighted (based on demand) centroids of contiguous regions within the Census Block Groups. This process resulted in a demand file that was able to capture some of the detail offered by the block demand without having so many demand points that the model becomes intractable. For details on the creation of the Weighted Aggregated Block Model see section 4.3 Weighted Aggregated Block Demand.

q. **Full supply file.** The candidate sites (data item \( k \)) needed to be incorporated into the supply coverage file (data item \( j \)). These potential sites were treated as candidate sites as opposed to fixed sites when running the predictive models. This allowed all the candidate sites a chance to be incorporated into the model solution but did not force them into the solution.

**DATA PREPROCESSING**

In order to run the models, both the demand supply and input data needed to be preprocessed. Most of the preprocessing involved the creation of the demand files.

**Demand Preprocessing**

This study used four different spatial units for representing the location of demand. Figure 3 gives a visual overview of how demand is determined and located: Figure 3a shows how demand would be located for the Tract Models; Figure 3b shows how demand would be located in the Block Group Models; Figure 3e shows how demand would be located in the Aggregated Demand Model; and Figure 3f shows how demand would be located in the Weighted Aggregated Block Model.

The shadings used in Figures 3a – 3f are used only to help visually distinguish different enumeration units (the shades do not reflect any other value).
A. Demand is located at the centroid of the tract.

B. Demand is located at the centroids of the block groups.

C. The demand for each of the blocks in known.

D. Much of the tract is unpopulated (shown in white).

E. Demand is located at the centroids of contiguous regions of blocks within a block.

F. The demand is located at the weighted centroid (asterisk) of the points of demand within each contiguous cluster of blocks. The size of the point symbols is proportional to the amount of demand in each of the clusters blocks.

Figure 3. Demand representations used in this study.
TRACK DEMAND

To run the ‘Tract Model’, a point coverage representing the demand locations of Idaho residents was created. To do this, it was necessary to multiply the age/sex data (data item f in section 4.2.1) by the appropriate PHCF visitation rates (data item g) for every Idaho census tract to obtain the total number of PHCF visits per year for each Idaho census tract. Then, the total yearly visit data (calculated above) was joined to the census tract .shp file (data item e). The Mean Center tool in ArcMap was then used to create a point .shp file containing the tract numbers and the x and y coordinates of all the centroids of all of Idaho’s census tracts. These centroid points were then assigned the demand of their respective tracts. This was done by creating a new field in the attribute table of the centroid .shp file and then joining the census tract .shp file and the centroid .shp file by their census tract number. Next, the field calculator was used to assign the demand of each tract to its respective centroid point. Finally, the ‘join’ was removed between the census tract .shp file and the centroid .shp file.

BLOCK GROUP DEMAND

To run the ‘Block Group Model,’ point coverage was created to represent the demand locations of Idaho residents. To do this, it was necessary to multiply the age/sex data (data item d) by the appropriate PHCF visitation rates (data item g) for every Idaho census block group. This produced a total number of PHCF visits per year for each Idaho census block group. Then, the total yearly visit data (calculated above) was joined to the census block group .shp file (data item c). The Mean Center tool in ArcMap was then used to create a point .shp file containing the block group numbers and the x and y coordinates of all the centroids of all of Idaho’s census block groups. These centroid points were then assigned the demand of their respective block group. This was done by creating a new field in the attribute table of the centroid .shp file and then joining the census block group .shp file and the centroid .shp file by their census block group number. Next, the field calculator was used to assign the demand of each block group to its respective centroid point. Finally, the ‘join’ between the census block groups .shp file and the centroid .shp file was removed.
**Aggregated Block Demand**

To create this file it was first necessary to export only populated blocks into a separate shape file (51,012 of the 88,452 of Idaho’s blocks are unpopulated; that is 52.9% of Idaho’s total area that is unpopulated). Next, the age/sex data (data item b) were multiplied by the appropriate PHCF visitation rates (data item g) for every Idaho census block. This produces a total number of PHCF visits per year for each Idaho census block. Then, the total yearly visit data (calculated above) was joined to the census block .shp file (data item a). Next, the Dissolve tool was used (in ArcMap) to dissolve the census block .shp file (data item a) based on census block groups. The option to create multipart polygons when dissolving was not utilized, and the sum of each census block’s total yearly visits was retained for each dissolved region. By not allowing multipart polygons, the dissolve tool creates clusters of contiguous blocks within each block group. Next, the Mean Center tool in ArcMap was used to find the centroids of each dissolved region, and the demand associated with that region was assigned to that region’s centroid point. This process resulted in a demand file that was able to capture some of the detail that is offered by the block demand but without having so many demand points that the model becomes intractable.

**Weighted Aggregated Block Demand**

The creation of the Weighted Aggregated Block demand file was similar to the creation of the Aggregated Block Demand file (see 4.3 Aggregated Block Demand). After dissolving the census block .shp file based on census block groups (using the Dissolve tool), the Mean Center tool in ArcMap was used to find the centroids of each census block from the census block .shp file. Then, the Intersect tool was used to give each block centroid point an FID that related them to a common ‘block cluster’ (the regions resulting from dissolving the blocks based on their shared block groups but without allowing the creation of multipart polygons). Next, the Mean Center tool was used on the file that resulted from the use of the intersect tool. The mean centers were calculated within common dissolved regions. The placement of each block cluster centroid point was weighed based on the demand of the block points within that particular block cluster, and the dimension field was utilized to retain the demand of the block cluster for each weighted block cluster point.
Supply Preprocessing

In order to run the predictive models it was necessary to create a file that consisted of all the 78 existing supply sites (data item \( j \)) and all 20 candidate sites (data item \( k \)), 98 sites total. This file was created using the Merge tool in ArcMap 9.2. The ‘candidate’ attribute specifies how the supply site is treated when the maxcover model is run (not a candidate = 0, a possible candidate = 1, a fixed/existing site = 2). These potential sites were treated as candidate sites as opposed to fixed sites when running the predictive models. This allowed all the candidate sites a chance to be incorporated into the model solution, but it did not force them into the solution if they were not part of the optimal solution.

MCLP Model Execution

In order to execute the MCLP Model, the initial roadwork needed to be constructed (4.4.1) and the MCLP Model had to be run (4.4.2).

Initial Road Network Construction

Constructing the road network involves three different steps: connecting demand to the road network, preparing the supply file, and preparing the road network to run the MCLP model.

Connecting Demand Points

The first step in the road network construction was to use the Near command in ArcInfo to find the closest point of the road network for each of the demand points. The x and y coordinates and the distance from the demand point to the nearest point on the road network were recorded in the attribute table of the demand file.

It was necessary to use a program called Conversion4generate.exe to reformat the data into the correct format for the Generate command. The Generate command was used to create lines connecting demand points to their nearest location/point on the road network (obtained from the near command). Then, the projection of this new ‘demand connections’ line cover was defined (in the case of this study, to the Lambert Conformal Projection) using ArcMap. The Append command was then used to combine the road network and the ‘demand connections’ line covers into a single coverage.
Then, in order to ensure that the road network and demand connections were appended correctly, the Build command had to be used (rather than the Clean command) for creating coverage topology. In trial runs of this study’s models, the use of the Clean command to create coverage topology resulted in the removal of some of the demand locations from the final road network. Even after reducing the fuzzy tolerance value to very low, demand locations were still being removed from the final road network. The Build command was used on the new road network cover (created by the use of the Append tool) with the line option to create an arc attribute table (AAT) for all the arcs in the coverage.

Next, using ArcEdit, it was necessary to edit the road network coverage by editing the arc features. The command, Sel Dangle, was used to select all dangling nodes. Then, the command, “extend 5,” was used to extend the selected dangling nodes five meters so that they could intersect existing arcs. The ideal value for use in the Extend command is one that is as small as possible without resulting in more than a few messages of “Arc not written out due to zero length”. Each message displayed after running the Extend command corresponded to a demand location that was not correctly connected to the road network. In a model with few or well dispersed demand locations, it might be possible to identify the unconnected demand connection lines by displaying the node errors in the draw environment of ArcEdit. In the case of this study, the model was run once with the incomplete road network in order to obtain the demand output results. Any demand location without a populated nearest supply site attribute was a demand node that needed to be manually connected in Arcedit (for details see section Manual Road Network Correction/Connection).

After the appropriate Extend value was determined (Extend 5m) and implemented, the coverage was saved. Next, the Build command with the line option was used on the newly saved road network. Then, the Build command was used once again, this time with the node option to create a node attribute table (NAT) for the nodes in the coverage. Finally, an extra copy of the road network was saved so that any manual connection of the road network that might be needed could be made.

**PREPARING THE SUPPLY COVERS**

First, the Near command was used to find node features in the road network near those in the supply point coverage. Then, an empty road network ID attribute was added to
the supply cover’s point attribute table (PAT) using the Add Item command. Next, the Relate Add command was used to create a relation between the road network node attribute table (NAT) and the supply point attribute table. Finally, the Calculate command was used to populate the newly created empty road network ID in the supply cover, based on its relation with the road network NAT.

**PREPARING THE ROAD NETWORK TO RUN THE MCLP MODEL**

First a new attribute (demand) was added to the road network NAT. Next, the Near command was used to find node features in the road network cover near those in the demand cover. Then, the Relate Add command was used to add a relation between the demand cover and the road network cover. This was done in order to populate the new attribute (demand) that was created in the road network NAT with the demand values from the demand cover (using the Calculate command). Next, using the Relate Add command, three additional relations were created that were necessary for the proper execution of the Network AMLs. The first two relate primary and secondary sites to a road network ID, and the final relation links a city name to a road network number (for details see Appendix B). Finally, all five of the relations were saved using the Relate Save command.

**Running the MCLP Location-Allocation Models**

The commands used for running the MCLP location-allocation models were implemented using an existing aml file (see Appendix A). In the aml file, the network environment was set up, then the location/allocation environment was set up, and finally, the location/allocation problem was solved.

**NETWORK ENVIRONMENT SETUP**

The NETWORK commands were used to set up the network environment for implementing location/allocation commands. The NETCOVER command was used to specify the network coverage as well as the output route system to be used by network commands. Next, the DEMAND command was used to specify the items in the network coverage attribute table that contained the demand values (PHCF visits per year). The IMPEDANCE command was then used to specify impedance items to be used by the
NETWORK command. The Length attribute (distance in network coverage) was used as the from-to and the to-from impedance item in all the models. No turn impedance was used. The CENTERS command was used to specify the INFO file that contained the attributes of supply nodes used by the location/ allocation commands.

**Location-Allocation Environment Setup**

The Location/Allocation commands were used to set up the location/ allocation environment and solve the MCLP problem, as well as to display the results. The LOCATECANDIDATES command was used (1 for a candidate site, 2 for an already existing, fixed site) to specify the features that are the candidate supply centers for location/ allocation by reading the items from the CENTERS file. The LOCATECRITERIA command specified the criteria or objective to use when solving the location/ allocation problem with the LOCATEALLOCATE command. The maxcover objective was chosen when running the LOCATEALLOCATE command and a max distance value of either 24,140 meters/15 miles, 36,210 meters/22.5 miles, or 48,280 meters/30 miles was used (depending on whether the particular model was adhering to the 30, 45, or 60-minute driving distance constraint, respectively).

**Solving the Location-Allocation Problem**

The LOCATEALLOCATE command solved the location/ allocation problem by computing the best location of centers (supply locations) to service given demand according to the criteria defined using the LOCATECRITERIA command. The allocation of demand nodes to these supply locations was also computed. When using the LOCATEALLOCATE command, the following had to be specified: an allocation output file, a centers output file, a global output file, the number of supply sites to be located, and finally, a heuristic to be used during computations.

The allocation output file contained information about the allocation of each demand node. Specifically, it contained, for each demand location, the nearest and second nearest supply location and the respective distances (along the road network) from the demand point to the supply point.
The centers’ (supply) output file listed the supply locations chosen in the location/allocation process, and it provided statistics on their performance, some of which were the total number of demand locations that could be serviced by the supply location; the amount of demand that could be serviced by the supply location; the total distance traveled by all of the demand locations that could be serviced by the supply location; the average distance traveled by all of the demand locations that could be serviced by the supply location; the distance of the furthest demand location that could be serviced by the supply location; etc. All of the above statistics were calculated only for demand locations that were within the driving-distance constraint of the given supply center.

The global output file consisted of one record containing various information about the model that was run. For example, the date and time the location/allocation model was run; the coverage used for the location/allocation; the name of the allocation and center output files; the number of supply centers to locate (previously specified). In addition, the global output file contained statistics about the entire location/allocation, most importantly, the total demand that was met and unmet within the distance constraint.

**MANUAL ROAD NETWORK CORRECTION/CONNECTION**

After solving the location/allocation problem, the output allocation file was inspected to ensure that all of the demand locations had a nearest facility listed (SITE1). If any of the demand locations in the output allocation file were missing a value in the SITE1 attribute column, it indicated that the demand site was not properly connected to the road network. Overshoots were the most common reason a demand point did not get connected to the road network. Demand locations that were not connected properly were highlighted and displayed in ArcMap, and to insure that all demand locations were considered in the model, it was necessary to manually connect these demand locations that were not connected during the initial road network construction.

To manually connect the demand locations, the following steps were performed. First, using ArcMap, all the demand locations from the output allocation file that did not get connected to the road network (those with an empty SITE1 value) were selected and displayed. Next, the extra copy of the road network, which was saved after its initial construction, was displayed using Arcedit’s draw environment.
Second, the demand locations that were not connected properly in the initial road network construction were connected one-by-one using the following technique. The demand connection arc of an unconnected demand point was selected (sel). Next the selected arc was Split near the location where the demand connection should have met the road network. Next the smaller section of the split arc (the portion closest to the road network) was removed (Delete). Then, the remaining portion of the split arc was selected, and the arc was manually extended enough to reach the road network using the Extend* (interactive Extend) command. This process properly connected the previously unconnected demand locations to the road network.

When a cluster of improperly connected demand point shared a common ‘near point’ on the road network, first the Select, Split and Delete commands were used, then each of the ‘clipped’ demand connections was selected and extended one-by-one. After all of the improperly connected demand points were manually connected, the road network coverage was saved.

Next, the Build command was used to build the new manually connected road network with, first, the line option and then, the node option. An extra copy of the road network was saved in case any demand locations were accidentally missed during the manual connection. Finally the model was rerun, now with a fully connected road network cover (no blank SITE1 values).

**Predictive Models**

The predictive models were run to evaluate the ability of the 20 candidate sites to meet unmet demand for primary health care services.

**Candidate Primary Health Care Facility Selection**

Candidate PHCF sites that could meet the demand which was found to be unmet in the corresponding deterministic model were considered to be potential sites for selection. Candidate PHCF site(s) that could alleviate the most unmet demand were considered the best selections.

Predictive models were run for four representations of demand at each of the three driving distance constraints (12 models). This was done in order to gain a better
understanding of the amount of unmet demand that could be eliminated by the incorporation of additional candidate PHCF sites. This also allowed for a better understanding of the robustness of the more suitable candidate sites.

Although the candidate site selection was performed on all four model representations of demand (Tract, Block Group, Aggregated Block, and Weighted Aggregated Block Model), it was assumed that the final candidate site selection should be based on what is believed to be the model that most accurately represents the population demand for primary health care service: the Weighted Aggregated Block Model.

It should be noted that all 20 candidate sites were being considered by the model at the same time, as opposed to running 20 models with each model considering each candidate site one at a time. Because of this, if two candidate sites are both capable of meeting previously unmet demand, the closer of the two sites would be reported as the primary site and the further as the secondary site even though both sites are capable of meeting the demand within the driving time constraint.

To insure that each of the candidate sites’ abilities to (by themselves) alleviate unmet demand is interpreted and reported correctly, it is important to determine if a candidate site is listed as a secondary site (within the driving distance constraint) for a previously unmet demand location. Sites like these were listed as secondary in the deterministic model but would be primary sites if the model had not incorporated all 20 candidate sites at once. This procedure insures that if a previously unmet demand point is within the driving time constraint of two candidate sites, both these candidate sites will be reported/considered capable of meeting the demand of that previously unmet point.

**Final Candidate Primary Health Care Facility Selection**

In order to gain a sense of the reduction in unmet demand for PHCS that could be obtained with the addition of a few new candidate sites, the final predictive model was run to include only the top performing candidate site(s) from the 20 new candidate sites. This model used the Weighted Aggregated Block demand because it was assumed to be the most accurate representation of PHCS demand locations.
CHAPTER 5

RESULTS

As the deterministic models moved from a low number of enumeration units (tract models) to a higher number of enumeration units (weighted and non-weighted aggregated block models), the amount of unmet demand for PHCS decreased.

Results of the deterministic weighted aggregated block model, which is assumed to be the most accurate model because of its potential to more accurately locate population/demand, indicate that 11.6% of Idaho’s population is farther than the federal guideline of 30 minutes driving distance to the nearest Primary Health Care Provider.

The results of the final predictive model indicate that, of the 20 candidate sites, the inclusion of a primary health care facility at the candidate site of Spirit Lake (located in north-western Idaho) would maximally increase services to the unmet demand that currently exists in the state.

DETERMINISTIC MODELS (FIRST 12 MODELS)

In this study four different types of demand representation were used, each incorporating three different driving-distance/time constraints to run the models, for a total of 12 (deterministic) models.

Tract Models

Idaho contains 280 census tracts, with an average census tract size of 772.0 km$^2$ and a standard deviation of 1,994.9 km$^2$. Each of the three deterministic tract models was run using a representation of demand at the centroid of the tract: resulting in 280 demand points. No manual editing of the road network connections was required to properly connect these tract demand points to the road network for the model.

In the tract models, the average distance from a tract demand point to the nearest primary health care facility is 18,415.8 meters (11.44 miles), and the average demand for primary health care services for a tract is 14,166.9 visits per year.
In Figure 4, all of the 280 tract demand points are plotted based on their respective distances to the nearest PHCF and their demand for PHCS.

![Figure 4. Tract model: Distance vs. demand.](image)

The tract model at the 30 minute driving distance constraint resulted in 19.9% of the state’s demand for PHCS unmet (80.1 of the demand was met). Using the 45 minute driving distance constraint resulted in 11.4% of the demand for PHCS unmet (88.6% of the demand was met). Running the model with the 60 minute driving constraint resulted in 4.8% of the demand for PHCS unmet (95.2% of the demand was met). (See Table 1.)

**Table 1. Deterministic Tract Model Results**

<table>
<thead>
<tr>
<th>Driving Time</th>
<th>Meters</th>
<th>Miles</th>
<th>Met Demand</th>
<th>% Met</th>
<th>Unmet Demand</th>
<th>% Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>24140</td>
<td>15</td>
<td>3179030.0</td>
<td>80.1</td>
<td>787706.0</td>
<td>19.9</td>
</tr>
<tr>
<td>45</td>
<td>36210</td>
<td>22.5</td>
<td>3513030.0</td>
<td>88.6</td>
<td>453702.0</td>
<td>11.4</td>
</tr>
<tr>
<td>60</td>
<td>48280</td>
<td>30</td>
<td>3778126.0</td>
<td>95.2</td>
<td>188608.0</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Figure 5 shows the results of the deterministic tract model. Tracts that are farther than 60 minutes driving distance from the nearest hospital are shown in black. Tracts that are
Figure 5. Deterministic tract model coverage.
between 45 and 60 minutes driving distance from the nearest hospital are shown in dark grey. Tracts that are between 30 and 45 minutes driving distance from the nearest hospital are shown in light grey and tracts that are within 30 minutes driving distance from the nearest hospital (tracts that are covered within the federal Guidelines) are shown in white.

**Block Group Models**

Idaho contains 952 block groups, with an average block group size of 226.7 km$^2$, with a standard deviation of 739.4 km$^2$. The block group models were run using a representation of demand at the centroid of each of the 952 block groups; this resulted in 3.4 times the number of demand points as the tract models. Manual connections were required to properly connect four of the block group centroid points to the road network.

The average distance from a block group demand location to the nearest primary health care facility is 14,592.4 meters (9.07 miles), and the average demand for primary health care services for a block group is 4,166.7 visits per year.

In Figure 6 all of the 952 block group demand points are plotted based on their respective distances to the nearest PHCF and their demand for PHCS.

*Distance from demand point to the nearest supply site.*

**Figure 6. Block group model: Distance vs. demand.**
The percentage of unmet demand for PHCS yielded by the block group models at a 30 minute driving constraint is 14.2% (met demand is 85.8%). At a 45 minute driving distance constraint, 7.6% of the state’s demand for PHCS is unmet (92.4% is met). At the 60 minute driving distance constraint 3.9% of the state’s demand for PHCS is unmet (96.1% is met). (See Table 2.)

Table 2. Deterministic Block Group Model Results

<table>
<thead>
<tr>
<th>Driving Time (min)</th>
<th>Meters</th>
<th>Miles</th>
<th>Met Demand</th>
<th>% Met</th>
<th>Unmet demand</th>
<th>% Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>24140</td>
<td>15</td>
<td>3401997.8</td>
<td>85.8</td>
<td>564736.3</td>
<td>14.2</td>
</tr>
<tr>
<td>45</td>
<td>36210</td>
<td>22.5</td>
<td>3664521.3</td>
<td>92.4</td>
<td>302212.8</td>
<td>7.6</td>
</tr>
<tr>
<td>60</td>
<td>48280</td>
<td>30</td>
<td>3812745.3</td>
<td>96.1</td>
<td>153988.8</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Figure 7 shows the results of the deterministic block group model. Block groups that are further than 60 minutes driving distance from the nearest hospital are shown in black. Block groups that are between 45 and 60 minutes driving distance from the nearest hospital are shown in dark grey. Block groups that are between 30 and 45 minutes driving distance from the nearest hospital are shown in light grey and block groups that are within 30 minutes driving distance from the nearest hospital (block groups that are covered within the federal Guidelines) are shown in white.

Aggregated Block Model

Idaho contains 1926 aggregated block clusters (groups of contiguous blocks within the existing block groups). Using the centroids of these aggregated block clusters results in roughly twice the number of demand points present in the block group model. The average size of these block clusters is 52.8 km², with a standard deviation of 188.4 km².

The average distance from an aggregated block cluster demand point to the nearest PHCF is 21,242.5 meters (13.2 miles), and the average demand for primary health care services for an aggregate block group is 2,059.6 visits per year. Manual corrections were required to properly connect 68 of the aggregated block group centroid points to the road network.

In Figure 8, all of the 1926 aggregated block group demand points are plotted based on their respective distances to the nearest PHCF and their demand for PHCS.
Figure 7. Deterministic block group model coverage.
The percentage of unmet demand for PHCS yielded by the aggregated block group models at a 30 minute driving constraint is 11.1% (88.9% of demand was met). At a 45 minute driving distance constraint, 5.9% of the state’s demand for PHCS is unmet (94.1% of the demand is met). At a 60 minute driving distance constraint, 3.3% of the state’s demand for PHCS is unmet (96.7% of demand is met). (See Table 3).

Table 3. Deterministic Aggregated Block Model Results

<table>
<thead>
<tr>
<th>Driving Time (min)</th>
<th>Meters</th>
<th>Miles</th>
<th>Met Demand</th>
<th>% Met</th>
<th>Unmet demand</th>
<th>% Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>24140</td>
<td>15</td>
<td>3527236.0</td>
<td>88.9</td>
<td>439497.8</td>
<td>11.1</td>
</tr>
<tr>
<td>45</td>
<td>36210</td>
<td>22.5</td>
<td>3734338.3</td>
<td>94.1</td>
<td>232395.8</td>
<td>5.9</td>
</tr>
<tr>
<td>60</td>
<td>48280</td>
<td>30</td>
<td>3836851.8</td>
<td>96.7</td>
<td>129882.2</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Figure 9 shows the results of the deterministic aggregated block model. Aggregated block clusters that are further than 60 minutes driving distance from the nearest hospital are shown in black. Aggregated block clusters that are between 45 and 60 minutes driving distance from the nearest hospital are shown in dark grey. Aggregated block clusters that are between 30 and 45 minutes driving distance from the nearest hospital are shown in light grey.
Figure 9. Deterministic aggregated block model coverage.
and aggregated block clusters that are within 30 minutes driving distance from the nearest hospital (clusters that are covered within the federal Guidelines) are shown in white.

**Weighted Aggregated Block Model**

All the weighted aggregated block models used the same number of demand points as the aggregated block models: 1,926. However, unlike the aggregate block models that placed demand at the centroids of the contiguous block clusters, in the weighted aggregated block model, the placement of those 1,926 demand points was influenced (weighted) by the amount of demand in each of the individual blocks that comprise the block group clusters. Figure 10 shows the distance between the placement of the non-weighted aggregated block demand centroids and the weighted aggregated block demand centroids (shown in order from the least amount of shift to the most). 544 of the 1,926 demand centroids had no change in their placement between the non-weighted and weighted aggregated block models. Of the 1,382 demand centroid that did have a shift in placement, the average shift was 2129.6 meters. It can be seen that there are two demand centroids that shifted far more that the others (around 45,000 meters). Figure 11 shows the change in coverage from the non-weighted aggregated block model to the aggregated block model. Block clusters with negative values showed a decrease in their degree of coverage from the non-weighted to the weighted block model, block clusters with positive values showed an increased in their degree of coverage. Figure 12 show that there is no relationship between the centroids’ shift in placement and its amount of demand for PHCS.

Manual corrections were required to properly connect 62 of the weighted aggregated block cluster centroid points to the road network.

The average distance from a weighted aggregated block cluster demand point to the nearest primary health care facility is 21,301.6 meters, and the average demand for primary health care services is 2,059.6 visits per year.

In Figure 13, all of the 1926 aggregated block group demand points are plotted based on their respective distances to the nearest PHCF and their demand for PHCS.

The percentage of unmet demand for PHCS yielded by the weighted aggregated block group models at a 30 minute driving constraint is 11.6% (88.4% of demand is met). At a 45 minute driving distance constraint, 6.1% of the state’s demand for PHCS is unmet (93.9% of
Figure 10. Centroid shift from non-weighted aggregated block model to weighted aggregated block model.
Figure 11. Coverage change from non-weighted aggregated block model to weighted aggregated block model.
Figure 12. Plot showing centroid shift distance vs. the number of visits per year.

Figure 13. Weighted aggregated block model: Distance vs. demand.

*Distance from demand point to the nearest supply site.*
the demand is met). At a 60 minute driving distance constraint, 2.0% of the state’s demand for PHCS is unmet (98.0% of demand is met). (See Table 4).

<table>
<thead>
<tr>
<th>Driving Time (min)</th>
<th>Meters</th>
<th>Miles</th>
<th>Met Demand</th>
<th>% Met</th>
<th>Unmet Demand</th>
<th>% Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>24140</td>
<td>15</td>
<td>3508209.8</td>
<td>88.4</td>
<td>458524.3</td>
<td>11.6</td>
</tr>
<tr>
<td>45</td>
<td>36210</td>
<td>22.5</td>
<td>3724111.5</td>
<td>93.9</td>
<td>242622.4</td>
<td>6.1</td>
</tr>
<tr>
<td>60</td>
<td>48280</td>
<td>30</td>
<td>3887181.3</td>
<td>98.0</td>
<td>79552.9</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Figure 14 shows the results of the deterministic weighted aggregated block model. Weighted aggregated block clusters that are further than 60 minutes driving distance from the nearest hospital are shown in black. Weighted aggregated block clusters that are between 45 and 60 minutes driving distance from the nearest hospital are shown in dark grey. Weighted aggregated block clusters that are between 30 and 45 minutes driving distance from the nearest hospital are shown in light grey and weighted aggregated block clusters that are within 30 minutes driving distance from the nearest hospital (weighted clusters that are covered within the federal Guidelines) are shown in white.

**PREDICTIVE MODELS (MODELS 12 – 24)**

The predictive models were run identically to the deterministic models except that the predictive models’ supply files, in addition to the existing 78 supply sites, contain 20 candidate sites (a total of 98 sites). Figure 15 shows the location of the 20 candidate sites. It can be seen that all the candidate sites are located in either northern Idaho (Figure 16) or west-central Idaho (Figure 17) except for the site in Ashton which is located in east-central Idaho (Figure 18).

The goal of the predictive models was to evaluate the ability of the 20 potential candidate sites to alleviate the unmet demand that was found in the deterministic models.

The overall results of the predictive models indicate that the candidate supply sites at Spirit Lake and Rathdrum each, individually, alleviate the highest percentage of previously unmet demand. It should be understood that because these sites are located relatively near to each other, much of the unmet demand that these two sites can service individually can be met by both or either of the sites.
Figure 14. Deterministic weighted aggregated block model coverage.
Figure 15. Location of three clusters of the 20 candidate supply sites (PHCFs).
Figure 16. Ten candidate sites located in northern Idaho shown in box A in Figure 15.

Figure 17. Nine candidate sites located in east-central Idaho shown in box B in Figure 15.
Figure 18. One candidate site located in west central Idaho shown in box C in Figure 15.

Tract Predictive Models

Based on the results of the corresponding deterministic models, it was revealed that only a small percentage of the unmet demand could be met by the incorporation of a candidate site; 1.4% - 13.81%.

Tract Predictive Model 24,140m

Five candidate sites are capable of servicing some of the unmet demand for PHCS within the 30 minute driving distance constraint. The candidate site at Rathdrum is capable of meeting more unmet demand than any of the other candidate sites: 40,558.1 visits per year (or 5.15% of the unmet demand for PHCS that exists in the state (See Table 5).

Tract Predictive Model 36210m

At the 45 minute driving distance constraint, only three candidate sites are capable of servicing any of the unmet demand for PHCS that exists in the state. It can be seen in Table 6 that the sites at Spirit Lake and Rathdrum are individually capable of servicing 11.78% of
Table 5. Candidate Site Evaluation for Tract Predictive Model at 24140 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>RATHDRUM</td>
<td>40558.1</td>
<td>5.15</td>
</tr>
<tr>
<td>SPIRIT LAKE</td>
<td>27394.9</td>
<td>3.48</td>
</tr>
<tr>
<td>PARMA</td>
<td>20948.6</td>
<td>2.66</td>
</tr>
<tr>
<td>MELBA</td>
<td>11633.2</td>
<td>1.48</td>
</tr>
<tr>
<td>ASHTON</td>
<td>11058.6</td>
<td>1.40</td>
</tr>
</tbody>
</table>

Table 6. Candidate Site Evaluation for Tract Predictive Model at 36210 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>53446.4</td>
<td>11.78</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>53446.4</td>
<td>11.78</td>
</tr>
<tr>
<td>ASHTON</td>
<td>14904.7</td>
<td>3.29</td>
</tr>
</tbody>
</table>

The unmet demand in the state. The candidate site at Ashton is capable of servicing 3.29% of the unmet demand.

**TRACT PREDICTIVE MODEL 48280M**

Again, the three candidate sites that are capable of servicing the greatest amount of unmet demand for PHCS within 60 minutes driving distance are Spirit Lake, Rathdrum, and Ashton (See Table 7).

Table 7. Candidate Site Evaluation for Tract Predictive Model at 48280 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>26051.5</td>
<td>13.81</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>26051.5</td>
<td>13.81</td>
</tr>
<tr>
<td>ASHTON</td>
<td>3846.2</td>
<td>2.04</td>
</tr>
</tbody>
</table>

**Block Group Predictive Models**

The results of the block group predictive models show a decrease in overall unmet demand for PHCS when compared to the results of the tract models.
**BLOCK GROUP PREDICTIVE MODEL 24140M**

Table 8 shows all the candidate sites that are capable of servicing the unmet demand for PHCS within the 30 minute driving distance constraint. The sites at Spirit Lake and Rathdrum are capable of servicing the most unmet demand for PHCS in the state, 7.34% and 6.4% respectively.

**Table 8. Candidate Site Evaluation for Block Group Predictive Model at 24140 Meters Driving Distance**

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>41473.5</td>
<td>7.34</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>36132.6</td>
<td>6.40</td>
</tr>
<tr>
<td>MELBA</td>
<td>12470.5</td>
<td>2.21</td>
</tr>
<tr>
<td>POST FALLS</td>
<td>12107.1</td>
<td>2.14</td>
</tr>
<tr>
<td>ASHTON</td>
<td>7591.2</td>
<td>1.34</td>
</tr>
<tr>
<td>PINEHURST</td>
<td>1833.1</td>
<td>0.32</td>
</tr>
</tbody>
</table>

**BLOCK GROUP PREDICTIVE MODEL 36210M**

At a 45 minute driving distance, five candidate sites are capable of servicing some unmet demand for PHCS in the state. Of these five candidate sites, Spirit Lake and Rathdrum are capable of servicing the most unmet demand, (each are capable of individually servicing 14.98%). (See Table 9).

**Table 9. Candidate Site Evaluation for Block Group Predictive Model at 36210 Meters Driving Distance**

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>RATHDRUM</td>
<td>45269.5</td>
<td>14.98</td>
</tr>
<tr>
<td>SPIRIT LAKE</td>
<td>45269.5</td>
<td>14.98</td>
</tr>
<tr>
<td>ASHTON</td>
<td>11437.4</td>
<td>3.78</td>
</tr>
<tr>
<td>PINEHURST</td>
<td>2848.9</td>
<td>0.94</td>
</tr>
<tr>
<td>SMELTERVILLE</td>
<td>2848.9</td>
<td>0.94</td>
</tr>
</tbody>
</table>
**Block Group Predictive Model 48280m**

Table 10 shows the three candidate sites that are capable of servicing some unmet demand for PHCS within a 60 minute driving distance constraint. Again, the sites at Rathdrum and Spirit Lake are capable of servicing the most unmet demand.

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>RATHDRUM</td>
<td>30762.9</td>
<td>19.98</td>
</tr>
<tr>
<td>SPIRIT LAKE</td>
<td>30762.9</td>
<td>19.98</td>
</tr>
<tr>
<td>ASHTON</td>
<td>8949.5</td>
<td>5.81</td>
</tr>
</tbody>
</table>

**Aggregated Block Predictive Models**

The results of the aggregated block predictive models indicated a further reduction of unmet demand for PHCS when compared to the tract and block predictive models.

**Aggregated Block Predictive Models 24140m**

At the 30 minute driving distance constraint, eight candidate sites are capable of servicing any unmet demand for PHCS in the state. Table 11 shows that half of those candidate sites that are able to service unmet demand for PHCS are only capable of servicing very small portions of the unmet demand (less than one percent of the existing unmet demand). As in most of the predictive models, the candidate sites at Spirit Lake and Rathdrum are able to service the largest percentage of unmet demand.

**Aggregated Block Predictive Models 36210m**

Table 12 shows the three candidate sites that can service some unmet demand for PHCS within 45 minutes driving distance. Both Spirit Lake and Rathdrum are individually capable of servicing 20.92% of the unmet demand in the state.
Table 11. Candidate Site Evaluation for Aggregated Block Predictive Model at 24140 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>47590.3</td>
<td>10.83</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>39487.7</td>
<td>8.98</td>
</tr>
<tr>
<td>MELBA</td>
<td>12459.7</td>
<td>2.83</td>
</tr>
<tr>
<td>ASHTON</td>
<td>7847.7</td>
<td>1.79</td>
</tr>
<tr>
<td>PINEHURST</td>
<td>2904.7</td>
<td>0.66</td>
</tr>
<tr>
<td>SMELTERVILLE</td>
<td>947.4</td>
<td>0.22</td>
</tr>
<tr>
<td>PARMA</td>
<td>85.4</td>
<td>0.02</td>
</tr>
<tr>
<td>WILDER</td>
<td>48.3</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 12. Candidate Site Evaluation for Aggregated Block Predictive Model at 36210 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>RATHDRUM</td>
<td>48624.6</td>
<td>20.92</td>
</tr>
<tr>
<td>SPIRIT LAKE</td>
<td>48624.6</td>
<td>20.92</td>
</tr>
<tr>
<td>ASHTON</td>
<td>8535.1</td>
<td>3.67</td>
</tr>
<tr>
<td>MELBA</td>
<td>23.3</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**AGGREGATED BLOCK PREDICTIVE MODELS 48280M**

Running the model using a driving distance of 60 minutes reveals that, four candidate sites can service some unmet demand for PHCS. Table 13 shows that Spirit Lake and Rathdrum are both individually capable of servicing 22.64% of the unmet demand for PHCS in the state.

Table 13. Candidate Site Evaluation for Aggregated Block Predictive Model at 48280 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>29406.7</td>
<td>22.64</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>29406.7</td>
<td>22.64</td>
</tr>
<tr>
<td>ASHTON</td>
<td>7607.8</td>
<td>5.86</td>
</tr>
<tr>
<td>MELBA</td>
<td>3616.0</td>
<td>2.78</td>
</tr>
</tbody>
</table>
**Weighted Aggregated Block Predictive Models**

The results of the three weighted aggregated block predictive models show a small increase in the amount of unmet demand for PHCS when compared to the non-weighted aggregated block predictive models. The results clearly indicate that the site at Spirit Lake would be the best choice of the 20 potential sites to maximally meet previously unmet demand for PHCS.

**Weighted Aggregated Block Predictive Model 24140M**

Table 14 shows that ten candidate sites are capable of servicing some of the demand for PHCS within the federal guideline of 30 minutes driving distance, but only Spirit Lake and Rathdrum are capable of servicing a substantial amount of unmet demand, 10.02% and 8.84% respectively.

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>45953.9</td>
<td>10.02</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>40538.4</td>
<td>8.84</td>
</tr>
<tr>
<td>MELBA</td>
<td>10495.5</td>
<td>2.29</td>
</tr>
<tr>
<td>ASHTON</td>
<td>8259.3</td>
<td>1.80</td>
</tr>
<tr>
<td>HAYDEN</td>
<td>4808.5</td>
<td>1.05</td>
</tr>
<tr>
<td>PINEHURST</td>
<td>1833.1</td>
<td>0.40</td>
</tr>
<tr>
<td>SMELTERVILLE</td>
<td>947.4</td>
<td>0.21</td>
</tr>
<tr>
<td>PARMA</td>
<td>75.2</td>
<td>0.02</td>
</tr>
<tr>
<td>WILDER</td>
<td>48.3</td>
<td>0.01</td>
</tr>
<tr>
<td>MARSING</td>
<td>19.7</td>
<td>0.006</td>
</tr>
</tbody>
</table>

**Weighted Aggregated Block Predictive Model 36210M**

Using the 45 minute driving distance constraint, six candidate sites are capable of servicing some unmet demand for PHCS. The candidate sites at Spirit Lake and Rathdrum are able to service a considerable amount of demand for PHCS; 20.76% and 18.65% respectively (see Table 15).
Table 15. Candidate Site Evaluation for Weighted Aggregated Block Predictive Model at 36210 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT LAKE</td>
<td>50359.0</td>
<td>20.76</td>
</tr>
<tr>
<td>RATHDRUM</td>
<td>45249.8</td>
<td>18.65</td>
</tr>
<tr>
<td>ASHTON</td>
<td>3181.5</td>
<td>1.31</td>
</tr>
<tr>
<td>PINEHURST</td>
<td>1936.2</td>
<td>0.80</td>
</tr>
<tr>
<td>SMELTERVILLE</td>
<td>1797.1</td>
<td>0.74</td>
</tr>
<tr>
<td>MELBA</td>
<td>23.3</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**WEIGHTED AGGREGATED BLOCK PREDICTIVE MODEL 48280M**

Table 16 shows that at the largest driving distance constraint (60 minute), four candidate sites are capable of servicing some unmet demand for PHCS that exist. In this scenario, the coverage capabilities of the sites at Rathdrum and Spirit Lake are identical (they cover the same demand points).

Table 16. Candidate Site Evaluation for Weighted Aggregated Block Predictive Model at 48280 Meters Driving Distance

<table>
<thead>
<tr>
<th>Candidate Site</th>
<th>Demand</th>
<th>% of Unmet</th>
</tr>
</thead>
<tbody>
<tr>
<td>RATHDRUM</td>
<td>30743.2</td>
<td>38.65</td>
</tr>
<tr>
<td>SPIRIT LAKE</td>
<td>30743.2</td>
<td>38.65</td>
</tr>
<tr>
<td>ASHTON</td>
<td>702.0</td>
<td>0.88</td>
</tr>
<tr>
<td>MELBA</td>
<td>470.1</td>
<td>0.59</td>
</tr>
</tbody>
</table>

**FINAL PREDICTIVE WEIGHTED AGGREGATED BLOCK MODEL (MODEL 25)**

Of the 20 candidate sites used in the predictive models, only 10 were able to alleviate any unmet demand for PHCS: Spirit Lake, Rathdrum, Melba, Ashton, Hayden, Pinehurst, Smelterville, Parma, Wilder, and Marsing.

Because the results of the weighted aggregated block model are assumed to be potentially the most accurate of the models, the results from the weighted aggregated block model were used to make a final recommendation and to determine a final coverage scenario that would exist with the inclusion of those potential site(s) that would maximally alleviate the unmet demand for PHCS that exist in Idaho.
Figure 19 graphs the ability of individual candidate sites to individually alleviate previously unmet demand for PHCS.

Initially, it might seem desirable to recommend all the candidate sites that are capable of servicing substantial amounts of demand. For the models in this study, Spirit Lake and Rathdrum are clearly the candidate sites that are most able to service unmet demand. However, the problem with simply choosing the top performing candidate sites from Table 14 is that it does not account for demand locations that can be serviced by multiple candidate sites. In the case of this study, the candidate sites of Spirit Lake and Rathdrum are located very close to each other and much of the demand that can be serviced within a given driving distance of Rathdrum can be serviced within that same driving distance of Spirit Lake (and vice versa).

Based on the results of what is assumed to be the model with the most accurate representation of the location of demand (the weighted aggregated block model), the addition of the candidate site at Spirit Lake would reduce the unmet demand for PHCS by 45,954 visits per year (10.02%).
ANALYSIS OF THE COVERAGE PATTERN

To better understand the coverage patterns of the deterministic models’ results, a Moran’s I and Hot Spot analysis was performed.

Moran’s I Analysis

Based on a visual analysis of the deterministic models (Figures 5, 7, 9, 14) it appears as though there is a clustering of covered and non-covered areas. To determine whether the clustering of enumeration units with similar coverage levels that was visually apparent was statistically significant, a Global Moran’s I analysis was conducted on all four predictive model outputs.

Moran’s I is a spatial statistic that measures spatial autocorrelation (feature similarity) based on feature location and attribute values. A Global Moran’s I analysis is often used to help identify appropriate neighborhood distance for other spatial analysis methods.

For the analysis, the degrees of coverage were divided into four categories: (1) coverage within 30 minutes, (2) coverage in between 30 and 45 minutes, (3) coverage in between 45 and 60 minutes, and (4) coverage in over 60 minutes. These four coverage divisions were used as the attribute input for the Moran’s I analysis.

Several different threshold distances were used to determine at which distance spatial autocorrelation (Moran’s I index value) was highest. Figure 20 plots the Moran’s I index value against different threshold distances for each of the four models (tract, block group, aggregated block, and weighted aggregated block). The plots span from the smallest threshold distance that could be used while ensuring that all the features have a nearest neighbor (a requirement for the statistical validity of the result output) to the largest threshold distance that could be used while ensuring that that distance is smaller than the feature class extent. For all four models spatial autocorrelation was highest when using the smallest threshold distance possible. All plots in Figure 20, regardless of the distance band value, yielded clustering at a significance level that indicates that there is less than a one percent chance that the clustered coverage patterns of the enumerations units used in the models could be a result of random chance.

Table 17 shows the peak Moran’s I index values and the corresponding distance band values as well as the z-scores and significance levels for all four model types.
Figure 20. Moran’s I index values for all four demand representation models at different threshold distances.

Table 17. Peak Moran’s I Index Values and the Corresponding Distance Band Values for all Four Demand Representation Models

<table>
<thead>
<tr>
<th>Enumeration Unit</th>
<th>Tract</th>
<th>Block Group</th>
<th>Agg Block</th>
<th>Wgt Agg Block</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morans I Index value</td>
<td>0.19</td>
<td>0.13</td>
<td>0.32</td>
<td>0.31</td>
</tr>
<tr>
<td>Z-score</td>
<td>9.35</td>
<td>16.97</td>
<td>50.39</td>
<td>48.27</td>
</tr>
<tr>
<td>Distance Band Value</td>
<td>79224</td>
<td>62104</td>
<td>35718</td>
<td>35718</td>
</tr>
<tr>
<td>(meters)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance Level</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>
The Moran’s I analysis confirmed that clustering of the enumeration units coverage values exists but it does not specify whether the clustering affects the high or low (or both) coverage levels. The analysis also indicates an appropriate threshold distance for use in other spatial analysis methods for each of the four model types.

**Hot Spot Analysis with Rendering**

Following the Moran’s I analysis to confirm that coverage level clustering exists among the enumeration units and to determine an appropriate distance band value, a Hot Spot analysis was run.

The hot spot analysis with rendering, implemented in ArcGIS software, calculates the Getis-Ord Gi* statistic for hot spot analysis and then applies a hot-to-cold type of rendering to the output Z scores.

Figures 21, 22, 23, and 24 show the results of the Hot Spot Analysis for the tract, block group, aggregated block, and weighted aggregated block model, respectively. Gi statistic z-scores of 2 or more indicate a statistically significant clustering of enumeration units with high levels of primary health care service coverage. Gi statistic z-scores of -2 or less indicate a statistically significant clustering of enumeration units with low levels of primary health care service coverage. Gi statistic z-scores between -1 and 1 indicate a lack of statistically significant clustering of enumeration units’ coverage values (high or low).
Figure 21. Hot spot analysis for tract deterministic model.
Figure 22. Hot spot analysis for block group deterministic model.
Figure 23. Hot spot analysis for aggregated block deterministic model.
Figure 24. Hot spot analysis for weighted aggregated block deterministic model.
CHAPTER 6

DISCUSSION

This study shows that as the representation of demand in the model moved from a low number of enumeration units (tract models) to a higher number of enumeration units (weighted and non-weighted aggregated block models), the amount of unmet demand for PHCS decreased. The tract model yielded the greatest amount of unmet demand. The aggregated block model yielded the lowest amount (except at the furthest driving time constraint), and the weighted aggregated block model yielded slightly greater amounts of unmet demand than the non-weighted aggregated block model, and it yielded the least amount of unmet demand at the furthest driving time constraint.

As the models moved from a low number of enumeration units (tract models) to a higher number enumeration units (weighted and non-weighted aggregated block models), the amount of unmet demand for PHCS decreased. Figure 25 shows the amount of unmet demand found in the 12 deterministic models (four demand representations each at the three driving distance constraints). Figure 26 plots the cumulative demand that can be met for each of the four different demand representations in the deterministic models, against the driving time constraint. It can be seen that one of the tract demand point is nearly 180 minutes for the nearest PHCF.

All models in the study had the same total demand for PHCS: 3,966,732 visits per year. It can be seen that the tract model results in greater amounts of unmet demand at all driving distance constraints than the other three models.

KEY FINDINGS

The results of the predictive models indicate that, of the 20 candidate sites, the inclusion of a primary health care facility at the candidate site of Spirit Lake would maximally increase services to the unmet demand that currently exists in the state (Idaho). Only nine other candidate sites (besides Spirit Lake) are able to alleviate any unmet demand for PHCS: Rathdrum, Melba, Ashton, Hayden, Pinehurst, Smelterville, Parma, Wilder, Marsing. The candidate site at Rathdrum performed almost as well (when considering all
Figure 25. Results of the deterministic models.
Figure 26. Travel time vs demand (visits per year) for all four demand representation models.
models) as the candidate site at Spirit Lake. This is due, primarily, to the fact that the candidate site at Rathdrum is close in location to the candidate site at Spirit Lake. At best, only 40.12% of the unmet demand for PHCS can be met by the incorporation of all 20 candidate sites (weighted aggregated block model at 48240 meter driving time constraint). This implies that additional candidate sites in other areas of the state would be required in order to service the rest of the unmet demand for PHCS. At worst (the tract model at the 24140 meter driving time constraint), only 12.3% of the unmet demand for PHCS can be met by the incorporation of all 20 candidate sites.

LIMITATIONS

This study makes the assumption that it takes 30 minutes to drive 15 miles. This assumes that in a rural, mountainous area, few freeways exist and roads are narrow, and often windy. These assumptions do not account for traffic congestion, difficult intersections, weather conditions, etc. This travel rate of 30 miles per hour adheres to the accepted estimations of the travel speed for major roads in Idaho.

The study assumes that all residents travel to their closest primary health care facility. This may be an acceptable assumption because in many rural areas a secondary PHCF would likely be quite far away.

This study also assumes that demand exists at the centroids of enumeration units only and that that demand travels to supply from that single point.

It is assumed that the Center for Disease Control’s estimated visits per year for the different age and sex groups are accurate.

It is also assumed that the conversion from PHCF visit hours per year into number of visits per year for each existing PHCF supply site is accurate.
CHAPTER 7

CONCLUSION

This study examined the accessibility of primary health care services to Idaho residents. Specifically, it used a MCLP model combined with four progressively more accurate representations of demand for PHCS to determine the amount and location of unmet demand for primary health care services by Idaho residents. In this study 12 models were run to identify areas where demand for primary health care services was unmet (four models with different representations of demand each run at three different driving distance constraints). Next, these 12 models were rerun using a supply file that contained 20 candidate sites in order to identify new candidate PHCF sites that would maximally alleviate the amount of unmet demand for health care services experienced by Idaho residents (as determined by the original 12 deterministic models).

As the models moved from a low number of enumeration units (tract models) to a higher number enumeration units (weighted and non-weighted aggregated block models), the amount of unmet demand for PHCS decreased.

The overall results of the predictive models indicate that the candidate supply sites at Spirit Lake and Rathdrum each, individually, alleviate the highest percentage of previously unmet demand (with the site at Spirit Lake performing slightly better).

The results of this study may be of particular interest to the Idaho Department of Health and Welfare which is charged with providing equitable access to primary health care services for all Idaho residents. The results of this (or a more current) study could be used by those in charge of recruiting primary health care practitioners. They could focus their efforts on enticing practitioners to practice in the locations determined by the model to maximally meet existing unmet demand for primary health care services.
CHAPTER 8

FUTURE RESEARCH

Although the MCLP models used in this study do not constrain capacity (meaning they do not account for facilities that are allocated more demand than they can service), it does not mean that the individual hospital supply information has to go unutilized. The information can be used to identify the locations and severity of a PHCF over-capacitation.

If it were possible to obtain, from the Idaho Department of Heath and Welfare, current data on the locations of primary health care practitioners, and their annual number of visit hours, it would be possible to reproduce this study with a focus on locations of PHCF that are under-capacitated with regard to the amount of PHCS the location can provide per year and the amount of demand that is within some driving distance threshold. This information would be valuable to those in charge of PHCF funding or those in charge of recruiting medical practitioners to existing sites. Appendix C shows an example of how an analysis of over-capacitated sites could be achieved. Besides simply identifying supply locations that are being allocated more demand than they can service, the analysis incorporated information about secondary supply sites that are within the driving time constraint and are under-capacitated (being allocated less demand than it is capable of supplying). This gives a more complete understanding of the potential extent of the over-capacitation among the supply sites.

Future research could also explore the capabilities of dasymetric mapping for improving the determination of population location for use in this location/allocation model. Dasymetric mapping is a technique used to refine information shown on choropleth maps by supplementing the data contained in choropleth maps with ancillary data. The ancillary data chosen should correspond to the information presented in the choropleth map. In this study, the block demand .shp file can be combined with an ancillary land-cover map that could distinguish developed locations from undeveloped locations. The use of dasymetric mapping techniques to determine locations of demand not only allows for a more accurate placement of demand points, it also ensures that the demand points will be placed within a populated
region. Unfortunately, initial attempts to incorporate dasymetric mapping techniques into this study, to aid in the placement of the demand points, yielded more demand locations than the MCLP model used in this study could accommodate. For more information about the incorporation of dasymetric mapping techniques into this accessibility model see Appendix D.

It should be noted that the techniques used in this study could be utilized in many contexts other than health care accessibility, such as accessibility to public schools, libraries, voting polls, postal offices, fire stations, etc.
REFERENCES


ESRI DataDisk2000, Redlands, CA: Environmental Systems Research Institute, Inc.


http://factfinder.census.gov/servlet/QTTable?_bm=y&-geo_id=04000US16&-qr_name=DEC_2000_SF1_U_QTP12&-ds_name=DEC_2000_SF1_U&-_lang=en&-redoLog=false&-CONTEXT=qt


APPENDIX A

AML SCRIPT USED TO EXECUTE MODEL
/* to run this aml you have to have output.aml in your workspace
/* used variable - .cand !!!this is global variable!!!
/*
/* COMMANDS FOR GENERATING SINGLE OUTPUT FILES AND DEMAND COVER
/*&type 'choose number between 78 and 121'
/*&sv .cand = [response 'how many candidate points would you like to locate']
/*
/* START COMMANDS FOR LOOP
&sv .cand = 78
&do &while %_.cand% le 78
/*
netcover clear
netcover roadfull route /* roadfull - line coverage;
    /* route - name of the route system
demand # totev
    /* totev - demand value item
    /* recalculated from demandfull cover!
impedance length length
    /* length - distance in roadfull cover
centers supplyfull.pat # # pev
    /* supplyfull.pat - file with candidates points
    /* route_id_item - #
    /* limit_item - no limits #
    /* pev - supply_item - how many visits the
    /* center can provide

locatecandidates clear
locatecandidates centers candidate
    /* it locates candidates from supplyfull.pat (see: centers)
    /* using candidate item
locatecriteria maxcover 24140
locate allocate out_demand%.cand% out_sup%.cand% out_stat%.cand% %.cand% /* statistic files /* number of candidate points &r output.aml clear mape pcساءfull arcs pcساءfull markersize 0.05 markersymbol 20 points demandfull spider roadfull node out_demand%.cand% roadfull# site1 2 /* /* END COMMAND FOR LOOP &sv .cand = %.cand% + 5 &end &return
APPENDIX B

AML SCRIPT USED TO DISPLAY MODEL RESULTS
/* output.aml
/* this aml is run by docs5.aml
/*
/* exiting arcplot
quit
/* starting tables
tables
/* adding items to selected files
/* see readme.txt file, points 6 and 7, for details
additem out_sup%.cand% city 25 25 C
additem out_sup%.cand% district 15 15 C
additem out_demand%.cand% site1_city 25 25 C # site1
additem out_demand%.cand% site2_city 25 25 C # site2
/* exiting tables
quit
/* starting arcplot
arcplot
disp 9999
/* restoring relations from info file named "relations"
relate restore relations
/*
mape demandfull
resel demandfull point mapextent
/* calculating city names and district numbers in out supply info file
calc out_sup%.cand% info city = cityname//city
calc out_sup%.cand% info district = cityname//district
/* calculating city names for selected sites as centers in out
demand info file
calc out_demand%.cand% info site1_city = site1//city
calc out_demand%.cand% info site2_city = site2//city
/* creating new output demand cover
&sys arc copy demandfull demand%.cand%
/* joining demand%.cand%.pat with out_demand%.cand% using road5# as join item
&sys arc joinitem demand%.cand%.pat out_demand%.cand% demand%.cand%.pat roadfull#
distance
&return
APPENDIX C

EXISTING SUPPLY (PHCF) CAPACITANCE
ANALYSIS PRIMARY HEALTH CARE FACILITY
OVERCAPACITATION
Valuable information about the capacitance of the existing supply sites can be derived from the models output data (i.e. the amount of demand available at each of the existing supply sites and the amount of demand that was allocated to those sites by the various models). An evaluation of the capacitance of the existing 78 PHCF was conducted for the deterministic weighted aggregated block models at each of the three driving time constraints. It should be noted that none of the location-allocation models available in ArcInfo solve capacitated location problems (meaning they do not account for facilities that are allocated more demand than they can service). Demand locations are assigned to only one supply site (the closest one) regardless of whether adequate supply of services exists at the site. Although ArcInfo contains tools to manually redistribute portions of demand for large models with many demand points, the time required to manually redistribute portions of demand until all supply sites are below capacity is prohibitive. Furthermore, in models that contain more demand than supply, an attempt to bring all supply facilities to capacitance by redirecting portions of demand would be impossible.

Based on the results of the deterministic models, it is evident that some of the facilities are allocated more demand than they can provide or, in other words, some of the facilities are within the driving distance constraint of more demand for PHCS than the facilities can provide. The simplest way to quantify the degree of a facilities overcapacitation is to calculate the difference between the total visits a facility can accommodate and the total visits the facility was assigned by the model. But this simple difference can not provide the best understanding of the degree of overcapacitation faced by the facility. This is because that simple difference does not account for demand locations that have secondary facilities that are undercapacitated and within the driving distance constraint that could meet the demand. Instead, the overcapacitation analysis incorporated information about secondary supply sites that were within the driving time constraint and were under-capacitated (being allocated less demand than it was capable of supplying). This technique gives a better understanding of the potential extent of the over-capacitation among the supply sites.

Demand that is assigned to an overcapacitated primary facility that also has a secondary facility that is undercapacitated and within the driving distance constraint can be subtracted from the overcapacitated primary site’s total assigned demand.
This technique provides a best-case/worse-case scenario with regard to the number of visits a PHCF can provide and the number of visits that PHCF might be asked to provide. The PHCF overcapacitation analysis was conducted (in Excel) for the deterministic weighted aggregated block model at all three driving time constraints.

This overcapacitation analysis determines the degree of overcapacitation of the site: whether the site is located near an alternate supply site that could meet demand some of the excess demand, or whether it is the only site in the near area.

Information derived from the analysis could be useful to those in charge of staffing facilities as it might be desirable to increase the number of physicians and/or the days and hours of operation at the existing facilities that are overcapacitated and located farther than the driving distance constraint from the nearest undercapacitated facility.

The overcapacitation analysis was performed in Excel.

It must be remembered that the supply values were estimated from data obtained in 1995 while the population and PHCF visitation rates are more current. Since the focus of the proposed study is methodological, rather than applicational (developed for the purpose of guiding a specific policy) these data limitations can be accepted.

Tables 18, 19, and 20 show the amount of demand that can be serviced by an undercapacitated secondary site (“Undercap Secondary”), as opposed to an overcapacitated primary site (“Overcap Primary”), with a driving distance constraint of 24140 meters, 36210 meters, and 48280 meters, respectively. The attribute “Overlapping Demand” shows the amount of demand that an overcapacitated primary supply site shares with an undercapacitated secondary site within the demand points’ driving distance constraints.
Table 18. Weighted Aggregated Block Model at 24140 Meters

<table>
<thead>
<tr>
<th>Overcap Primary</th>
<th>Undercap Secondary</th>
<th>Overlapping Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hayden Lake</td>
<td>Coeur d'alene</td>
<td>100960.96</td>
</tr>
<tr>
<td>Hailey</td>
<td>Sun Valley</td>
<td>21124.67</td>
</tr>
<tr>
<td>Sagle</td>
<td>Sandpoint</td>
<td>19217.96</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>American Falls</td>
<td>6802.33</td>
</tr>
<tr>
<td>New Meadows</td>
<td>McCall</td>
<td>3398.44</td>
</tr>
<tr>
<td>Plummer</td>
<td>St. Maries</td>
<td>682.82</td>
</tr>
<tr>
<td>Lapwai</td>
<td>Lewiston</td>
<td>591.65</td>
</tr>
<tr>
<td>Horseshoe Bend</td>
<td>Garden Valley</td>
<td>253.03</td>
</tr>
<tr>
<td>Craigmont</td>
<td>Cottonwood</td>
<td>96.29</td>
</tr>
</tbody>
</table>

Table 19. Weighted Aggregated Block Model at 36210 Meters

<table>
<thead>
<tr>
<th>Overcap Primary</th>
<th>Undercap Secondary</th>
<th>Overlapping Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hayden Lake</td>
<td>Coeur d'alene</td>
<td>115375.77</td>
</tr>
<tr>
<td>Hailey</td>
<td>Sun Valley</td>
<td>28460.76</td>
</tr>
<tr>
<td>Grangeville</td>
<td>Cottonwood</td>
<td>18011.19</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>American Falls</td>
<td>9933.30</td>
</tr>
<tr>
<td>Plummer</td>
<td>St. Maries</td>
<td>7887.57</td>
</tr>
<tr>
<td>Lapwai</td>
<td>Lewiston</td>
<td>7067.66</td>
</tr>
<tr>
<td>Kamiah</td>
<td>Orofino</td>
<td>5677.59</td>
</tr>
<tr>
<td>Pierce</td>
<td>Orofino</td>
<td>3946.44</td>
</tr>
<tr>
<td>New Meadows</td>
<td>McCall</td>
<td>3471.83</td>
</tr>
<tr>
<td>Craigmont</td>
<td>Cottonwood</td>
<td>2240.18</td>
</tr>
<tr>
<td>Plummer</td>
<td>Coeur d'alene</td>
<td>2136.03</td>
</tr>
<tr>
<td>Kellogg</td>
<td>Coeur d'alene</td>
<td>1646.40</td>
</tr>
<tr>
<td>Horseshoe Bend</td>
<td>Garden Valley</td>
<td>253.03</td>
</tr>
<tr>
<td>Mountain Home</td>
<td>Glenns Ferry</td>
<td>141.04</td>
</tr>
<tr>
<td>Pocatello</td>
<td>American Falls</td>
<td>14.99</td>
</tr>
<tr>
<td>Overcap Primary</td>
<td>Undercap Secondary</td>
<td>Overlapping Demand</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Hayden Lake</td>
<td>Coeur d'alene</td>
<td>115375.77</td>
</tr>
<tr>
<td>Mountain Home</td>
<td>Glens Ferry</td>
<td>36279.10</td>
</tr>
<tr>
<td>Hailey</td>
<td>Sun Valley</td>
<td>30805.45</td>
</tr>
<tr>
<td>Grangeville</td>
<td>Cottonwood</td>
<td>18081.16</td>
</tr>
<tr>
<td>Kamiah</td>
<td>Orofino</td>
<td>12124.87</td>
</tr>
<tr>
<td>Plummer</td>
<td>St. Maries</td>
<td>10342.62</td>
</tr>
<tr>
<td>Lapwai</td>
<td>Lewiston</td>
<td>9974.98</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>American Falls</td>
<td>9933.30</td>
</tr>
<tr>
<td>Pierce</td>
<td>Orofino</td>
<td>4103.90</td>
</tr>
<tr>
<td>New Meadows</td>
<td>McCall</td>
<td>3471.83</td>
</tr>
<tr>
<td>Craigmont</td>
<td>Cottonwood</td>
<td>2921.92</td>
</tr>
<tr>
<td>Kellogg</td>
<td>Coeur d'alene</td>
<td>2622.26</td>
</tr>
<tr>
<td>Plummer</td>
<td>Coeur d'alene</td>
<td>2162.96</td>
</tr>
<tr>
<td>Rupert</td>
<td>American Falls</td>
<td>60.30</td>
</tr>
<tr>
<td>Pocatello</td>
<td>American Falls</td>
<td>14.99</td>
</tr>
<tr>
<td>Fairfield</td>
<td>Sun Valley</td>
<td>10.26</td>
</tr>
</tbody>
</table>
APPENDIX D

DASYMETRIC TECHNIQUES
Initially, this study aimed to incorporate dasymetric mapping techniques to aid in the placement of the demand points. Unfortunately, the dasymetric techniques employed resulted in an intractably large number of demand points and the model was unable to run.

**LIMITATIONS OF CHOROPLETH MAPS**

The problem with using choropleth maps to aid in the placement of demand is that in choropleth maps the information regarding the population pertains realistically only to certain locations within the enumeration units. This is a well recognized problem with the use of choropleth maps in studies involving the location of population. Langford and Unwin (1994) found three problems associated with choropleth maps. First, the enumeration areas for data collection are arbitrarily determined. This creates large enumeration units for regions with low population density and small ones for high population density. Because the enumeration unit’s boundaries are arbitrary, they do not reflect densities of spatial distributions; as a result, more attention is focused on the size of the enumeration unit than on the data they convey. Second, the spatial intricacies of the mapped phenomena are generalized when aggregated to the enumeration unit; the enumeration unit acts much like a low-pass filter. Most importantly, these enumeration units do not distinguish populated areas from non-populated areas, such as agricultural areas, rivers, lakes, forests, mountains, etc. Therefore, the portrayal of population locations by choropleth maps is not accurate because non-populated areas appear as populated.

Openshaw (1983) correctly notes that when working with choropleth maps, it is often difficult to determine if the results of data analysis reveal useful information about the people living in the region being analyzed, or whether the results are just a function of the areal unit used in the study.

**DASYMETRIC MAPPING**

As described by Langford and Higgs 2006, “Dasymetric modeling utilizes ancillary information resources to internally redistribute variables within the limits of their tabulation zone so as to create subzones of relative homogeneity. By facilitating the spatial refinement of aggregated data, dasymetric maps can better show, in the case of population statistics, the places where people actually reside,” (p. 297). The use of dasymetric mapping, a method of
disaggregating standardized census data, can alleviate many of the problems associated with choropleth maps. Dasymetric mapping converts administrative units into smaller map units, which are more relevant to the study in which they are used (Langford and Unwin 1994). Dasymetric mapping technique would allow for a more accurate positioning of demand locations for use in location/allocation modeling.

When first developed by the Russian Cartographer Tian-Shansky in the 1920’s, dasymetric maps were far more difficult to produce than choropleth maps. However, with the development of Geographic Information Systems, the use of dasymetric maps as opposed to choropleth maps has increased because of the increased ease of implementing dasymetric mapping techniques and the known limitations associated with choropleth maps.

Generally, dasymetric mapping is a method of thematic mapping. It is often used to realistically redistribute population density within a Census enumeration unit using ancillary data relating to the distribution of population within the enumeration unit. The source of ancillary data to determine populated areas often comes from classified satellite imagery. The type of ancillary data used is unimportant as long as (in the case of population distribution) it can accurately separate occupied from unoccupied space. In this sense, dasymetric mapping can be thought of as an areal interpolation technique.

Dasymetric mapping techniques have been incorporated into studies that work with data that are aggregated into enumeration unit (e.g. crime mapping), (Bowers and Hirschfield 1999; Poulsen and Kennedy 2004).

There exists no standardized methodology with regard to the implementation of dasymetric mapping. Variations exist in terms of the choice of ancillary dated utilized or the degree of internal differentiation attempted e.g. two-tiered, three-tiered (Langford and Higgs 2006). Research has shown that using a simple two-tiered technique, a binary dasymetric method produced equivalent results to those using more complex multi-tiered dasymetric methods (Eicher 1999; Eicher and Brewer 2001; Mennis 2003; Langford 2006).

In the literature to date, it appears that few previous studies have utilized dasymetric approaches to aid in the computation of accessibility measures in health care accessibility studies. One such study, conducted by Langford and Higgs (2006), examines the results of using differing spatial models of population to determine the outcomes of healthcare accessibility modeling. Langford and Higgs (2006) found that using dasymetric models to
estimate demand population locations results in lower accessibility scores than using the generally accepted pro-rata method for representing the location of population (demand). Furthermore, they found that “the difference is spatially disproportionate,” implying that the degree of disadvantaged PHCF accessibility in rural areas may be greater than previously determined. However, their study was conducted in small areas in Wales, and their ancillary data for determining dasymetric demand included specific information about where buildings were located as opposed to more general land-cover data.

Figure 27 shows an example of the demand points obtained using dasymetric mapping.

**Figure 27. Dasymetric mapping method.**

**DERIVING DASYMETRIC DEMAND (IN THIS STUDY)**

Dasymetric mapping techniques were used to locate the demand associated with developed regions within populated census blocks. The dasymetric demand was obtained by using the block demand from the census block .shp file in combination with a Land Cover map as an ancillary data source. Using a combination of standard GIS tools (intersect, calculate area, dissolve, raster to vector, etc.), the total demand within an entire census block was redistributed to developed regions within that census block. The amount of demand redistributed was proportional to the size of the developed region relative to the size of all developed regions within that block. Finally, the mean center of all developed regions was found to produce a point coverage of demand where the points were located at the centroid of the developed regions and each of these points had an associated demand.

The land cover map of Idaho used for ancillary data was obtained from the National Land Cover Database 2001 and was used to identify developed areas in Idaho. More
specifically, it was combined with the block demand .shp file to identify developed areas within populated census blocks. The land cover map had a spatial resolution of 30 meters.

**Dasymetric Mapping Method: Assumptions/Limitations**

The dasymetric mapping technique used to redistribute the census blocks’ population densities assumes a uniform density within the areas considered to be populated. Therefore, within the same census block, larger ‘developed’ areas are attributed more population (demand) than smaller ‘developed’ areas. While it is fair to assume that larger developed areas will have larger populations, this will not always hold true in all locations, all circumstances, and all size scales. One positive aspect of using a binary dasymetric technique is not having to determine subjective population redistribution rates. The resolution of the ancillary Land Cover Map limits the size of the developed areas that can be utilized to redistribute the population density in the creation of the dasymetric demand map. While 30m resolution is fine enough to capture apartment complexes and large houses, it may be too coarse to capture smaller dwellings. The strength of the relationship between ‘developed’ areas and population distribution controls the amount of error introduced into the dasymetric population redistribution.