1-1-2013

A Geographic Modeling Framework for Assessing Critical Infrastructure Vulnerability: Energy Infrastructure Case Study

Leanne Sulewski
University of South Carolina

A GEOGRAPHIC MODELING FRAMEWORK FOR ASSESSING CRITICAL INFRASTRUCTURE VULNERABILITY:
ENERGY INFRASTRUCTURE CASE STUDY

By

Leanne Sulewski

Bachelors of Arts
University of North Carolina-Greensboro, 2008

Masters of Arts
University of North Carolina-Greensboro, 2010

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Geography

College of Arts and Sciences

University of South Carolina

2013

Accepted by:

Michael E. Hodgson, Major Professor

Susan Cutter, Committee Member

Steven J. Fernandez, Committee Member

Diansheng Guo, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

Dedication

This work is dedicated to my parents, without whom this document and journey would have never been feasible. Thank you for all of your unwavering love and support!

ACKNOWLEDGEMENTS

This has been an amazing journey and no acknowledgements section could possibly do justice to all of the people who have helped me complete this journey, but with that being said, I'd be remiss if I didn't try.

First and foremost, I'd like to thank my committee: Dr. Hodgson, Dr. Cutter, Dr. Fernandez, and Dr. Guo, whose guidance and support were unparalleled. A big thank you to Dr. Hodgson, for helping me through all the logistical issues and helping me develop research that combines the theoretic and practical aspects of Geography. I extend huge thank you to Dr. Cutter for all of her support and advice. I wouldn't have gotten very far without your guidance, and I truly appreciate it. I also must thank Dr. Guo for all his advice towards building a methodologically sound research project.

Dr. Fernandez, I cannot thank you enough for your support and mentorship throughout my time at Oak Ridge. This research would have been next to impossible without your support with data and overall knowledge of grid vulnerabilities. Thank you for guiding me and helping to improve this research to be better than I could have ever imagined on my own.

This research would have also been impossible without the help of my dear friends and colleagues. At USC I received so much support from my colleagues: Ian Kramer, Shufan Liu, Ronnie Schumann, Nancy Pazmino, Dan Morath, Caglar Koylu, and

Gregg Boswer. I'd like to extend special thanks to Chris Burton, who took time out to provide me with advice, even from all the way in Italy. I have also received amazing support from my colleagues at Oak Ridge National Lab. Thank you for laughter and support when times were rough, especially: Amy Rose (those dissertation writing parties were a blast!) and Harini Sridharan (food comas and coding anybody?). I would also like to extend special thanks to those friends who were simply there to listen and offer guidance, especially: Steven and Devon Weiler, Mike Swaim, Tracy Nickerson, Sunny Chudgar, Mary "Beebe" Hall-Brown, Paxton Neubert, Eric Hatcher, Danielle Breece, and Allison Tompkins. I would've been lost without you guys! Also, last but certainly not least, I am thankful for the advice and guidance of Dr. Mayur Gosai, who despite his busy schedule still was there to help.

ABSTRACT

Vulnerability of critical infrastructure systems is of the utmost importance to a nation's national security interests, especially the electric grid. Despite the importance of these systems and planning, disruptions continue to occur at an alarming rate, thus indicating a fundamental flaw in the way critical infrastructure systems are analyzed for vulnerability.

Critical infrastructure systems are typically analyzed using mathematical approaches such as graph theory, which strip systems of their important geographic information, and only look at connections between electrical stations (e.g. substations). While these relationships and metrics provide useful information, they cannot provide the entire picture. As such, the goal of this research was to develop a new, geographic framework for modeling infrastructure vulnerability, that not only takes into account the information uncovered by graph metrics, but information about the unique geography of the area that can impact these systems. Using Southeast Asia as a study region, the research questions probed and answered were:

1. What are differences that arise from analyzing energy network vulnerability using the new geographic framework versus a graph theoretic framework alone?
2. What types of evaluation methods are applicable for determining if the proposed framework is more effective than graph theory?

To answer these questions, this research developed a field-based model utilizing service areas as the unit of analysis. The factors in the model were betweenness, degree, closeness, land use, service area population, other critical infrastructure frequency, natural hazard frequency, and temperature extremes. The infrastructure nodes were ranked based on a weighted linear model of factor scores. These factors were then weighted, using the Analytic Hierarchy Process to determine the weights, and factor values summed to determine an overall vulnerability ranking for each node.

The results indicate that many of these factors provide modest insight into the vulnerability of the electric grid, when validated against real-world data from the 2012 Indian Blackout. The most important factors were betweenness, land use, natural hazard frequency, and temperature extremes. A measure of agreement between modeled substation vulnerability and the substation status (operation or non-operational) during the 2012 Blackout provided the basis for effective evaluation of each model.

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

List of Abbreviations

ADB..Asian Development Bank

AHP..Analytic Hierarchy Process

ANOVA .... Analysis of Variance

BEA ....Bhutan Electrical Authority

BPC ....Bhutan Power Corporation

CA .... Cellular Automata

CEA ....Central Electric Authority

CI ....Critical Infrastructure

CIP ....Critical Infrastructure Protection

CIESIN .... Center for International Earth Sciences Information Network

CR ....Consistency Ratio

DHS..Department of Homeland Security

DIME..Dual Incidence Map Encoding

DOE .... Department of Energy

DOP .... Department of Power

FOUO.. For Official Use Only

GEOINT.. Geospatial Intelligence

GIS .... Geographic Information Systems

GUI .... Graphical User Interfaces

IDW .... Inverse Distance Weight

IEISS .... Interdependent Energy Infrastructure Simulation System

IPCC .... Intergovernmental Panel on Climate Change

LANL .... Los Alamos National Laboratory

LDC .... Less Developed Countries

MAUP .... Modifiable Areal Unit Problem

MCDA .... Multi-criteria decision analysis

MW .... Megawatts

NCDC .... National Climatic Data Center

NGA .... National Geospatial-Intelligence Agency

NISAC .... National Infrastructure Simulation and Analysis Center

NRC .... National Resource Council

OD .... Origin-Desination

ORNL .... Oak Ridge National Laboratory

PAR .... Pressure and Release Model

PGP .... Power Generating Plants

PoDiuM .... Power Distribution Model

SA .... Sensitivity Analyses

SA/OA .... Service Area/Outage Area

SARI/E .... South Asia Regional Initiative for Energy

SDSS .... Spatial Decision Support System

SRLDC .... Southern Regional Load Dispatch Center

UNEP .... United Nations Environment Programme

US .... United States

WLC .... Weighted Linear Combination

CHAPTER I

INTRODUCTION

1.1. Overview

On a daily basis, critical infrastructure makes life easier for people around the globe. By definition, critical infrastructure is: the "foundation for national security, governance, economic vitality, and way of life" (United States (US) Government 2003). Critical infrastructures not only provide services that make life easier for a country's citizens, but these services also create a strong national defense and a sense of national identity and pride. Types of critical infrastructure may include, but are not limited to: energy, transportation, telecommunications, and even national monuments (US Government 2003). Often these infrastructures are interconnected, and damage to one network of critical infrastructure can have cascading effects upon other critical infrastructure networks, possibly causing major damage to a country's national security and identity. The interconnectedness of these infrastructures not only extends to other types of critical infrastructure, but can also be extended across political boundaries; in many cases critical infrastructures are dependent on international agreements and cross international borders (Schintler et al. 2007).

Industrialized and developing countries alike have identified the extreme importance of protecting critical infrastructure. Critical infrastructure disruptions, either intentional via an attack, or unintentional via natural disaster, are not infrequent occurrences, even for developed countries. For instance, ten major blackouts have occurred in industrialized countries between 2003 and 2006 (Pearson 2011, Koger 2008). Incidents have even occurred as recently as July 2012, where much of the world's population was without power due to a blackout in India (Memmott 2012). Typically, research has focused on transportation and energy networks and disruptions to these networks in industrialized countries. Little research on critical infrastructure in less developed countries has been conducted, likely because the data to support such research are scarce. Additionally, the sophisticated modeling techniques that are used to understand electrical grid (the network of transmission lines, substations, and power plants) vulnerabilities have not been able to isolate these occurrences. Blackouts are still frequent occurrences. It is clear that existing modeling techniques for determining vulnerabilities in the electrical grid are inadequate to project future vulnerabilities and incorporate into standard guidance.

1.2. Statement of Problem

There is a fundamental flaw in how critical infrastructure (e.g. utilities and transportation routes) is represented and analyzed in critical infrastructure protection models, as major disruptions to these systems are still occurring (Hines et al. 2010). Most critical infrastructures can be described and analyzed using graphs, and most research regarding infrastructure vulnerability has studied critical infrastructures using graphs (Hines et al. 2010, Arianos et al. 2009, Desmar et al. 2008, Holmgren 2007, Desmar et al. 2007, White and Smyth 2003). Hines et al. (2010) did, however, indicate that using only graph networks to describe, analyze, and assess networks can be misleading, especially in the case of vulnerability assessments.

Graphs are generally comprised of two components: nodes (or vertices) and links (or edges). In an energy network, a node might be a power station or a substation, while a link would be the transmission line between the power station and substation. Each link contains one or two nodes, called endpoints (Figure 1.1).

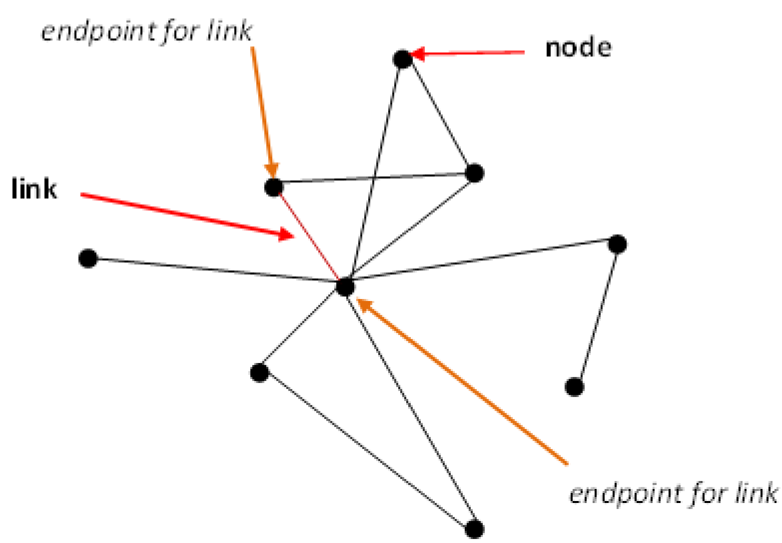


Figure 1.1: Representation of a graph.

These graphs also have a variety of properties that reveal information about the graph. Nodes, links, and endpoints are just a small subset of the properties that can describe a graph. A few other common properties of a graph or its elements might include are degree, geodesic path, diameter, and betweenness. The degree of a node refers to the number of links that are connected to the node. The geodesic path describes the shortest path through the network from one node to another. The diameter of the network is the number of links of the longest geodesic path between two nodes (Gross and Yellen 2003). Finally, betweenness indicates the number of links that pass through a node (Rocco et al. 2011).

Failures in the electrical network still occur, despite the abundance of graph-based techniques for analyzing critical infrastructure vulnerabilities, advances in techniques for analyzing vulnerabilities, and technologies analyzing the electrical grid. While graph approaches, such as the centrality metric betweenness, are useful and provide a great deal of information about the electrical grid and its interdependencies, there are shortcomings to only utilizing these approaches (Hines et al. 2010, Kim and Obah 2007, Holmgren 2006). This dissertation research utilized betweenness to refer to vulnerabilities as defined by graph metrics.

One major shortcoming of graph-based approaches is the lack of data availability and standardization to support such analysis. Even in developed countries, data about critical infrastructure, especially the energy grid, may be proprietary and difficult to access. Additionally, critical infrastructure models often suffer from a Data Death Spiral (Bhaduri 2013). Initially, data are only available in an aggregated form, and the critical infrastructure models were built to ingest these data. Eventually, models were built to perform simulations on parallel platforms and were able to answer new questions; however, these new models require finer spatial resolution data. Such fine resolution data are not available because the models that were created never required this level of granularity. It is a cycle that has limited the development of the granular infrastructure models that are really needed to understand critical infrastructure systems.

Another shortcoming of graph theoretic approaches was that they do not provide any great detail to geographic characteristics of the surrounding area that might contribute to a network's vulnerability. Hines et al. (2010) indicate that while graph metrics provide information about the general vulnerability of the network, these metrics are misleading when viewed alone and without ancillary information. First, graph and simulation approaches do not address infrastructure service areas, or the area with which the infrastructure serves. The attributes of the service area may make a node more or less vulnerable depending on the characteristics it encompasses. For example, the loss of nodes (substations) "a" and "b" may each cut off energy from two additional nodes on each side of an electrical network (Figure 1.2). If node "b" has a service area with very few clients (such as in a rural area), it may be less critical than node "a," which serves more people or contains critical facilities, such as a hospital (Figure 1.2). In this example, these two nodes might be ranked equally vulnerable by graph approaches that do not include characteristics of the population (absolute number, income, age, etc) within the service area. Alternatively, the node that serves more people and/or contains medical facilities (such as a hospital) may be regarded as more vulnerable and critical and should have a higher rank, indicative of higher vulnerability. If a natural disaster were to strike, and decision makers only looked at betweenness, substations "a" and "b" would have the same vulnerability, when in reality substation “a” is more vulnerable due to the higher population and higher frequency of critical assets.

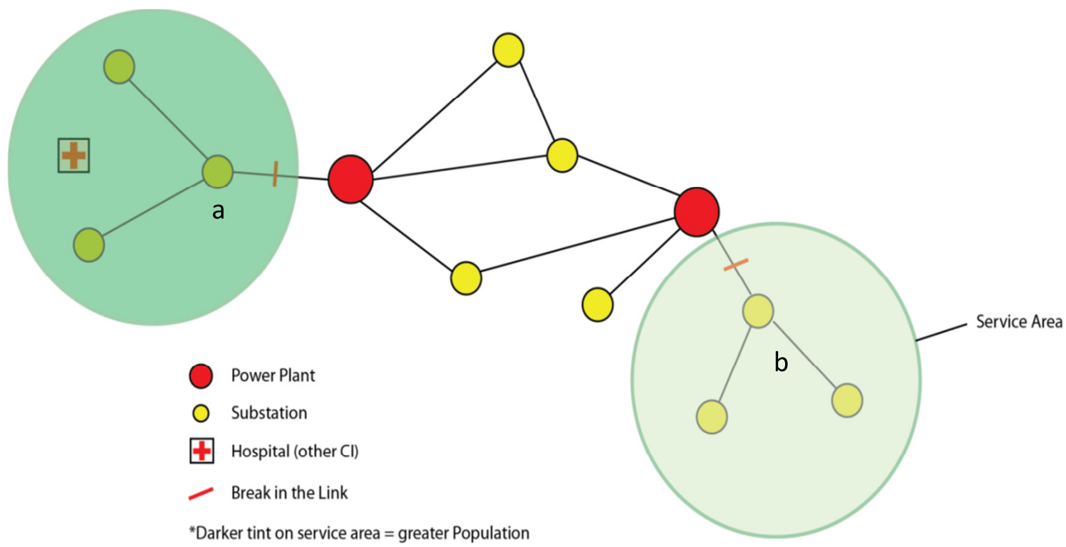


Figure 1.2: Example of characteristics affecting criticality and vulnerability, not just the topologic/structural characteristics.

Additional examples of the usefulness and importance of a geospatially integrated approach include the occurrence of a natural disaster, where a controlling authority (e.g. state government, private utility) may want to divert a node’s power to serve only the areas with the greatest amount of people and the most critical infrastructure, such as hospitals, fire departments, and shelters. In this case, the controlling authority may choose to shut down substations that serve fewer people and divert that power to a node with larger concentrations of critical infrastructure and population. Another operational example is the combinations critical infrastructure (CI) protection models with natural disaster reduction models. Present models are reactive in response to an event; however, vulnerability trends are leaning more towards risk reduction (Seck 2007). Combining CI protection models with the attributes of the geographic area, such as past frequencies of natural disasters, might identify areas where humanitarian assistance organizations can focus their funding efforts on reduction of risk to natural disasters in the critical infrastructure context (Seck 2007). A different, but also viable example of such an approach is from a military operations standpoint, where it may be desirable to identify the node that disrupts the greatest number of people or critical facilities. In such applications, the goal may be to cause the greatest harm rather than minimizing harm. (The converse use of such modeling is to predict likely targets on electrical infrastructure by terrorists.) The use of graph modeling alone would not appropriately rank the nodes utilizing these service area attributes. The modeling framework developed by this research can be applied to these example problems of predicting electrical nodes that serve infrastructure of greatest interest.

1.3. Research Goals

The goal of this research was to develop, demonstrate, and validate a combined graph/field-based modeling framework for modeling vulnerability from disruptions in a critical infrastructure network. This dissertation research addresses limitations in graph theoretic models with the addition of field-based (a 'field' in Geographic Information Science is a representation of geographic space as a continuum) modeling methods. This dissertation research provides a more effective method of analyzing critical infrastructure networks for vulnerabilities by incorporating spatial attribute information from a variety of sources to determine the most vulnerable nodes.

1.4. Research Objectives

This dissertation research created a modeling framework for analyzing energy infrastructure networks to determine the relative vulnerability of nodes. Nodes (substations) are defined as the central point at which energy is distributed. The modeling framework couples graph theoretic approaches with pertinent attribute and geographic attribute data to help reduce overall time spent analyzing networks (Pertinent attribute data is discussed more extensively in Chapter 2.). The combination of a graph-theoretic approach with an area field-based modeling approach creates a framework that, when implemented, takes into account the uniqueness (i.e. "place-based") of the area of interest.

The combination of betweenness with geographic data (i.e. 2 or 3-dimensional) creates a unique problem with the unit of analysis. Substations are typically represented as points; however, they serve geographic areas that vary in size depending on the capacity of the substation. Using geographic data only at the point location of the substation to determine its vulnerability is misleading. As such, the unit of analysis in this research was substation service area, which provides a more appropriate representation of the unique area the substation is servicing. Service areas, and the problem and process for determining service areas are described in greater detail in Chapter 3.

1.4.1. Research Objectives include:

1. Identification of representations of factors; and

2. Identification of the relationship between the joint combinations of factors when identifying vulnerabilities in the electric grid (e.g. population characteristics and node/link characteristics, or land use and population) using a raster-based Weighted Linear Combination (WLC) model.

After this analysis framework was created, the overall research questions were probed:

1. What are differences that arise from analyzing energy network vulnerability using the new integrated framework versus a graph theoretic approach? Are these differences substantial?
2. Where are the critical and vulnerable nodes in Southeastern Asia? Southeastern Asia for the purposes of this research comprises India, Bhutan, and Nepal. One reason for using Southeast Asia as the study area is this region has a unique electrical network, where certain areas have little western influence. Additionally, Indian experienced a widespread blackout in July 2012, making it of increasing interest to determine where vulnerabilities in their electrical grid lie. A more detailed explanation of the study area justification can be found in Chapter 3.
3. What types of evaluation methods are applicable for determining if the proposed framework is more effective than a graph theoretic approach? Effectiveness in this case refers to the performance of the approaches for

determining substation vulnerability (i.e. ability to correctly rank a substation in a historic blackout).

1.5. Dissertation Structure

This dissertation is separated into five chapters. These chapters include such information as a comprehensive literature review, methods, results, and conclusions. Chapter 2 presents a literature review of the previous research conducted on critical infrastructure vulnerability, and its applications in graph theoretic and geographic information science literature. Chapter 3 discusses the methods and data sources utilized; it also presents the research structure. Chapter 4 presents the results of the research. Finally, chapter 5 discusses the results and conclusions that can be drawn from this research and presents future directions that may be explored.

Chapter II

Literature Review

The issue of critical infrastructure vulnerability and criticality is an interdisciplinary issue that requires the synthesis of information from a variety of disciplines. As such, this dissertation research draws from literature in Geographic Information Science (GIS) and GIS-based modeling, graph networks, and vulnerability science. The sources of literature consisted of the Internet databases Web of Science and Google Scholar, and consisted of English language literature. This excludes literature that may have been written in other languages such as Hindi, Nepali, Dzongkha, and Chinese. However, there were collaborations with foreign counterparts in both India and Bhutan, who have indicated that literature on this topic is negligible. Also available and of interest is the methods and research conducted by the U.S. National Laboratories, such as Oak Ridge National Laboratory and Los Alamos National Laboratory. Through the support of the National Laboratories this "gray" literature was also reviewed.

2.1. GIS-Based Modeling

GIS-based modeling attempts to emulate the real world, in a simplified way, for the purposes of either understanding the biophysical processes or simply predict the outcomes of such processes. The application of GIS-based modeling extends to a variety of disciplines, including, but not limited to, environmental studies, epidemiological studies, and urban analyses. GIS-based models can take a variety of forms:

descriptive or predictive; verbal or graphical; static or dynamic. Verbal models are simply stated with words, while graphical models are depicted using graphs, charts, and figures. Static models are usually indicators, while dynamic models depict a process through time (Maguire et al. 2005).

Implementing models in a GIS environment can be accomplished in a number of different ways. The modes of implementation include loosely coupled, closely coupled, and embedded models, characterized by the degree of integration with a GIS system and its geospatial database. In loose coupling, a model exists outside of a GIS, where the model's output is not in a format that is read into a GIS. The data can be converted and then ingested into a GIS database for further analyses and visualization. One example of a loosely coupled model is running an Analytic Hierarchy Process (AHP) in a stand-alone program, and then bringing the results into ArcGIS for further processing. Close coupling includes a model that is outside of a GIS; however, the model's output is directly ingestible into a GIS and/or the model can directly use geospatial data. Closely coupled models, also called tightly coupled, include the Urban Flood Model for ArcGIS, which integrates the urban flood model into ArcGIS (Kang 2010). Embedded models are models that are implemented within a GIS, such as those developed using model builder. Each implementation technique has its own advantages and disadvantages. For example, loosely coupled models require little time to integrate, require more extensive data management methods and have a very low capability for executing tasks simultaneously. Embedded models require more time to integrate; however they may be much faster and require less data management (Maguire et al. 2005; Westervelt

2002; Goodchild et al. 1993). Inevitably there are tradeoffs with any integration technique; however, it is up to the modeler to determine which approach will work best for their research question.

2.1.1. GIS-Based Infrastructure Modeling

Much of the modeling with regards to CI protection is within the realm of simulation or graph metrics (graph metrics are discussed in section 2.3.3). One of the major missions within the simulation and CI protection community has been the interdependencies within CI. The National Infrastructure Simulation and Analysis Center (NISAC) developed the Interdependent Energy Infrastructure Simulation System (IEISS). IEISS is an actor-based model that utilizes not only the interactions within a critical infrastructure system (e.g. actors within electrical grid components, such as between transmission lines and substations) but also among them (e.g. corresponding parts of the electrical grid and natural gas system). The IEISS model analyzes the interdependencies among critical infrastructure networks by looking for potential cascading failures and the links between one CI network and another CI network to determine the further effects of failures in one system (Bush et al. 2003).

IEISS, like many other CI protection models, lacks a general data format for input and output, which makes integrating the IEISS model with complementary models difficult. Los Alamos National Lab (LANL) developed a Service Oriented Architecture called Hydra that integrates models such as IEISS and complementary models and associated data, and makes the models easily accessible to the user (Bent et al. 2009). The interface for Hydra is a web-based approach where the user accesses the models through web services. The major disadvantage to this approach is that if a web service is down, there is no way to access the models. The approach proposed by this dissertation research is an approach that is located within ArcMap™, which can either be part of a web service or stand alone, and its purpose can be adjusted to the situation. The approach in this dissertation can quickly be changed from a strategic military standpoint to natural disaster reduction. The dissertation model, like Hydra, can be integrated for a number of different missions, but may be adjusted by the researcher within the ArcGIS™ environment and does not need to (but can) access a web service.

2.1.2. Spatial Decision Support Systems

The GIS-modeling approach in this dissertation research is set in the context of a decision-making environment: deciding which service areas are most vulnerable. Jankowski (1995) identified two perspectives on using GIS for decision support. The perspectives include GIS as being the center of a Spatial Decision Support System (SDSS) and integrating GIS with specialized models. The SDSS perspective utilizes GIS to generate, evaluate, test, and provide recommendations to spatial decision problems. The integration perspective utilizes existing models and works to incorporate the models into a GIS (Jankowski 1995).

Multi-criteria decision analysis (MCDA) has been used with energy infrastructure; however, MCDA has been utilized in other applications besides determining network vulnerability and criticality. For example, MCDA was utilized by Wang et al. (2001) to determine the best method of restoring an electrical system in response to a failure. Additionally multi-criteria methods were also used for ranking potential failure of equipment in energy substations (Moreira et al. 2009). The framework in this dissertation research consists of ranking the most critical and vulnerable service areas in a network. Ranking each service area requires examination of various criteria. Each criterion is assigned associated weight, indicating the criterion's relative importance to the service area's rank. For example, a service area's access to resources may not be as important as the service area population, so the access to resources would be weighted less than service area population. MCDA easily allows for the weighting of such criteria.

GIS-based MCDA has been widely used and researched for the last 15 years, especially in land suitability and planning scenarios. Several methods have been developed for GIS-based multi-criteria decision analysis, including weighted summation and outranking methods. Of the MCDA methods, weighted summation is the most widely used method, as reflected in some 39% of the articles Malczewski (2006) surveyed. Malczewski (2000; 2006) cited that the wide use of weighted summation methods is due to the easy application of the methods in GIS with map algebra.

WLC involves several steps, many of which have been overlooked when applying WLC in a GIS environment (identified in Figure 2.1). Many researchers who utilize WLC in a GIS environment fail to have a complete understanding of the assignment of weights and deriving attribute maps. Despite the lack of complete understanding, WLC has been widely used in GIS-based decision rules (Malczewski 2000). Defining a set of attributes can often be the most difficult and controversial task. The attributes must be comprehensive, measureable, complete, operational, decomposable, non-redundant, and minimal. Creating maps of these defined attributes requires that each of the attributes be transformed into comparable units, so that they can be compared between and among each other. Defining factor weights is central, and the most controversial part of using WLC, to making sure that the method represents the relative importance of factors to decision makers. However, inappropriate use of weights (or the assumption of equal weights) is common in GIS applications of WLC. Several methods of defining weights have been developed to help reduce the bias associated with arbitrarily assigning weights. These weight definition methods include the swing weights technique and Analytic Hierarchy Process (AHP). The swing weights technique requires the decision maker to answer questions about their preferences. Alternatively, AHP requires pairwise comparisons of attributes and will be described more fully in the next section. Aggregating the attributes and weights utilizes Equation 2.1, where the summation of each attribute's weight and value are determined for each location (Malczewski 2000).

$$V(x_i) = \sum_i w_j r_{ij}$$ **Equation 2.1**

where: w_j= weight

r_{ij}=value for j-th factor

i = location

Finally, after the weights and the attribute values are combined, the alternatives are ranked. Typically, this includes a rank order where a value of 1 indicates the best alternative and 0.0 the worst alternative (Malczewski 2000).

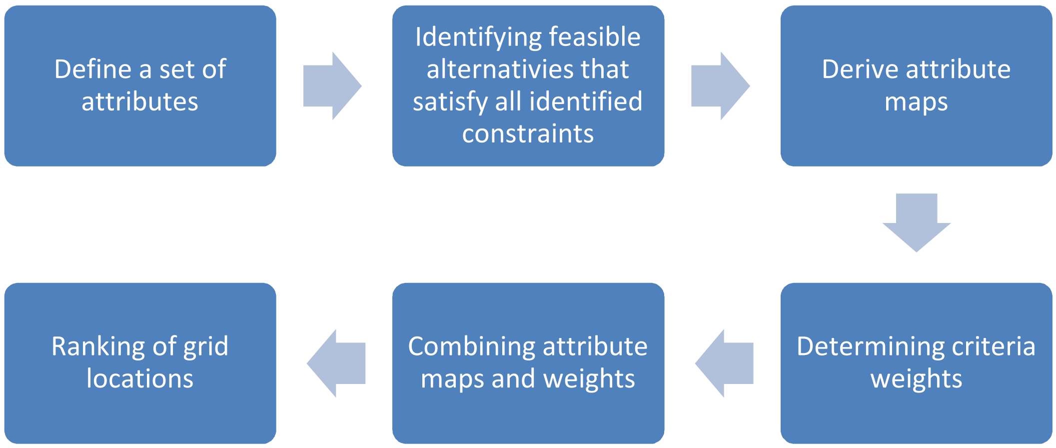


Figure 2.1: Weighted Linear Combination (WLC) steps in a GIS environment.

Little to no research has explored the use of attribute-based GIS suitability models in assessing the vulnerability and criticality of energy networks. One Master's Thesis conducted by Lemon (2004) utilizes AHP to help identify critical locations in electrical infrastructures; however, Lemon's method attempted to identify critical locations at the town level, which is much finer identification scale than required by this research. Additionally, Lemon's research did not involve national critical infrastructure networks, as required by this dissertation research. Most research involving vulnerability and criticality of energy networks focus on purely graph theoretic approaches, with few branching out to include the role of energy flow through the links (Arianos et al. 2009).

2.1.2.1. Analytic Hierarchy Process

A major controversy surrounding the use of WLC and many other methods for MCDA is the assignment of weights. The arbitrary assignment of weights and the bias associated with that assignment is a major limitation of the use of WLC; however, to help reduce this limitation, Saaty (1994, 2008) introduced the AHP.

Saaty (1994) maintains that a decision making approach should have five characteristics: the approach should be simple, adaptable to groups and individuals, natural, encourage compromise, and not be difficult to master or communicate. Saaty (1994) defines AHP as structuring a problem as a hierarchy, where the problem is decomposed into its most general factors and aggregated the solutions of the sub problems into a decision. Once the options are defined, a cost/benefit analysis must be conducted (Saaty 1994).

Once the criteria are defined, these criteria can be used in a pairwise comparison of relative importance to determine the relative weight of each criterion. The pairwise comparisons, termed "judgments" by Saaty (1994), are numerical representations of a relationship between two elements, representing the strength of the relationship (Table 2.1). The pairwise comparisons are set up like a matrix (Figure 2.2). If the element on left is less important than that on the top, enter the reciprocal value. If the element on the left is more important than that on the top, enter the whole number value (Table 2.1).

Table 2.1: Benchmark Pairwise comparison numbers for AHP as described by Saaty (1994). These, and in-between values, are always used.

Intensity of Importance	Definition	Explanation
1	Equal Importance	Contribute equally to the objective
3	Moderate Importance	Slightly favor one activity over another
5	Strong importance	Strongly favor one activity over another
7	Very Strong	Dominance is demonstrated in practice
9	Extreme importance	Evidence favoring one activity over another is highest possible

	Criteria 1	Criteria 2	Criteria 3
Criteria 1	1	1/9	3
Criteria 2	9	1	8
Criteria 3	1/3	1/8	1

The "3" indicates that Criteria 1 is more important than Criteria 3.

Figure 2.2: An example of an AHP matrix.

From the paired comparisons, priorities are calculated. The weights can be determined by either normalizing each column of priority values for each criterion for consistent matrices (see below) or by computing the eigenvalues for inconsistent matrices (Saaty 2008).

With any method that includes human input, there must be a method to validate the weights determined as a result of this method. The method utilized is the Consistency Ratio (CR). The CR is calculated by comparing the inconsistency of the set of judgments in a given matrix to the inconsistency of the set of judgments and the corresponding reciprocals taken at random scale (Equation 2.2).

$$CR = \frac{CI}{RI}$$ **Equation 2.2**

Where:

CI = Consistency Index

RI = Random Consistency Index

The consistency index indicates that the largest eigenvector (λ_{max}) should be equal to the size of the comparison matrix (Equation 2.3)

$$CI = \frac{\lambda_{max} - n}{n - 1}$$ **Equation 2.3**

The random consistency index (*RI*) offers a test statistics that compares the results from the RI to the number of items in the pairwise comparison (Table 2.2). Table 2.2 is based on the average random consistency index for a sample size of 500 matrices. For example, if there are four items in the comparison matrix, the *RI* is 0.9.

Table 2.2: Table for determining the random consistency index.

N	2	3	4	5	6	7	8	9	10
RI	0.00	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.51

If the consistency ratio is less than or equal to 0.10 the inconsistency is acceptable (Saaty 1977, 1994).

2.1.2.2. AHP in GIS and Alternatives

AHP has, in some applications, revolutionized the way weighting factors is conducted in GIS. The application of AHP has helped reduce the bias and subjectivity of weights in these types of analyses. Research that has utilized AHP includes land-use suitability analyses (focusing on site-suitability), land use management issues, and land-use planning (Sener et al. 2010, Ohta et al. 2007, Ma et al. 2005). To a lesser extent, AHP has also been integrated with GIS for hazard and environmental assessments (Rahman et al. 2009, Li et al. 2010).

The majority of the integration of AHP and GIS has been through close coupling, where the process of determining the weights through AHP is done outside of a GIS (Sener et al. 2010, Ma et al. 2005). In many cases, these analyses utilized AHP for the selection of weights for generating the suitability maps. For example, Ma et al. (2005) combined the Pass/Fail Screening, AHP, and WLC to derive suitability maps. Xu et al. (2012) combined AHP with WLC to evaluate the environmental suitability for living in 35 cities in China. Rahman et al. (2009) utilizing the same method for soil erosion hazard evaluation. While the method for combining AHP with GIS has considerable similarities among the various applications, few researchers have attempted to embed AHP into GIS (i.e. create an “embedded model”).

Research that has embedded AHP into GIS is sparse. Exceptions to this include the implementation of AHP developed in Visual Basic for Applications (VBA) for ArcGIS by Marinoni (2004) and the implementation of AHP and ordered weighted averaging (OWA) developed by Boroushaki and Malczewski (2008). Both tools are available for open download on the ArcGIS's ArcScripts developer's page. The AHP for ArcGIS extension developed by Marinoni (2004) utilizes integer raster datasets, ingests them, reclassifies them based on user input, requires input into the pairwise comparison matrix, and outputs a suitability raster dataset. Both approaches are easily integrated into legacy ArcGIS environments; however, both approaches require a fairly in depth understanding of AHP. The AHP-OWA extension developed by Boroushaki and Malczewski (2008) uses fuzzy quantifiers. The major difference between Boroushaki and Malczewski's (2008) approach and that of Marinoni (2004) is that instead of using a weighted linear combination to calculate the local scores for the raster cells, it uses a variety of local aggregations. Both approaches require that the user be able to fill in the comparison matrix, which may be difficult for a decision-maker who might want to utilize the methodology. Some closely linked models have used graphical user interfaces (GUIs) for the AHP calculations that utilize a slider bar (from most to least important) and populates the comparison matrix based on the user's selection along the slider bar. The matrix output was stored in a database that was then ingested into ArcGIS (Thirumalaivasan and Karmegam 2001).

While AHP is widely used for determining the weights for criteria in multi-criteria decision analyses, other methods have been utilized. One method is the ranking method, where the criteria are ranked by decision maker preferences, arbitrarily. Another method is the rating method, where criteria are rated on a pre-determined scale (Drobne and Lisec 2009). For example, if a decision maker is trying to decide which land would be most suitable to place their new store, they may have a rating system of 1 to 5, rating the importance of each criterion. Criteria may have the same rating in this method. While these methods are easily implemented, especially in a GIS, they are incredibly subjective.

2.1.3. Scale of Analysis

There are different observational scales available for analysis when discussing and analyzing electrical power networks. Holmgren (2007) discussed the general structure of an energy network, but not the specific components of the network. The observational scale may be at the transmission, sub-transmission, or distribution grid level. A typical transmission grid contains power stations (or power plants), transmission lines, substations, and transformers. Electrical energy typically flows from a power plant through transmission lines to substations, and finally to transformers that reduce the voltage to transmit to users (Brown and Sedano 2004). This dissertation research will only assess criticality and vulnerability of the network to the substation (service area) level due to data availability and ability to understand at the overall system (Figure 2.3).

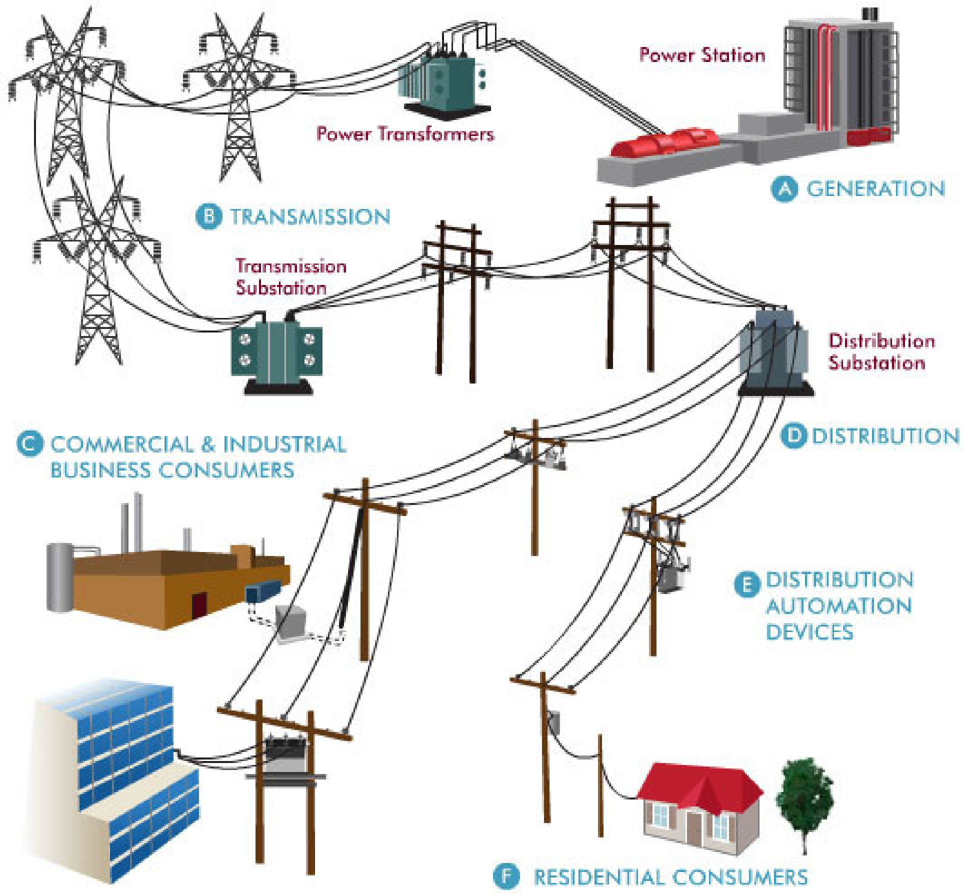


Figure 2.3: Power grid levels (source: http://venturebeat.files.wordpress.com/2010/10/grid.jpg, 2010).

2.1.4. Analysis Issues

This dissertation research needed to divide the network landscape into service areas. The characteristics of each service area, such as population, are not measured at a point (node) or at the service area level, so aggregation is necessary. Population censuses typically use historically formal units in the United States and other countries (e.g. block-group or census tract) or other non-electrical service area regions. This aspect of the analysis ignites the modifiable areal unit problem (MAUP). MAUP is often observed in two perspectives. The first scope is when the geographic scale of analysis is changed, the results also change. The second scope, which is exemplified by this research, is the "zonation problem." The second scope refers to the variation in results that is observed with different divisions of an area (Green and Flowerdew 1997).

MAUP is a geographic problem that has plagued geographic analysis since it was formally identified in 1934 by Gehlke and Biehl (Green and Flowerdew 1997). However, since MAUP's identification, several methods have been developed to minimize its effects. Langford and Higgs (2006) analyzed three different modes of representing population in catchment areas. The population distribution models they defined use (1) a weighted population centroid in the catchment area, (2) an evenly distributed population within the area, and (3) a dasymetrically distributed population. Langford and Higgs (2006) found that the method of population dispersal had a large impact on the modeling outcomes, but found that the dasymetric method gave the most realistic representation.

2.2. Model Evaluation

Any GIS model needs to undergo verification and validation to ensure that the model behaves as it was designed and that it mimics real life, respectively. Sensitivity analyses also evaluate model stability. Despite the importance of evaluating a model's performance, very little attention has been paid to these steps in multi-criteria decision analyses (Delgado and Sendra 2004).

2.2.1. Verification

Verifying a model requires the modeler to have an understanding of how the model is meant to operate and what the results should be. Often verifying requires the modeler to take a small subset of the data and test the algorithms by hand and compare them to the model output. Additionally, the repeatability of the model can be determined by running the model several times and subtracting the resulting rasters, which should yield pixel values of zero for the entire area (Demers 2002). In other words, for deterministic models the results should be consistently the same.

2.2.2. Validation

Determining a model's validity is often seen as more difficult than verifying it. In many cases, model validation is performed using field validation methods; however, field validation does have drawbacks: it is expensive and time consuming. An additional drawback is that many GIS models seek to predict future conditions, which make it impossible to use field validation methods (Demers 2002).

2.2.3. Sensitivity Analysis

Sensitivity analyses (SA) of models often include examining the variation of model outputs when the numerical input parameter values are varied. Very few studies in the GIS-based multi-criteria decision analysis (MCDA) realm have utilized SA to evaluate their models and results (Delgado and Sendra 2004). Of the greatest interest in evaluating GIS MCDAs is the uncertainty introduced when varying the factor weights (Chen et al. 2009, Tate 2013). The most widespread use of SA, albeit limited, is the changing of the weights of the criteria to determine changes to the overall outcome

(Delgado and Sendra 2004). In response to Delgado and Sendra's (2004) analysis, Chen et al. (2009) developed a method of SA for AHP multi-criteria decision analysis research. According to Chen et al. (2009), the best SA tool for MCDA would take into account several levels of uncertainty, including identifying which criteria are most sensitive to weight changes and visualizing the spatial changes. Chen et al. (2009) recommends a systematic approach to AHP spatial MCDA SA. This includes a series of simulations where each criterion is changed by a certain percentage to see how the overall results are affected and to see how the number of pixels for each criterion is affected per iteration of simulations (Chen et al. 2009). Tate (2013) also provided uncertainty analyses for social vulnerability indices, finding that the most uncertainty comes from the identification of weights in vulnerability indices.

2.3. Graph Theory Analyzing Critical Infrastructure

Several graph-theoretic approaches have been utilized in past research to determine the vulnerability of critical infrastructure networks of developed countries. Indices, such as average path length (the distance between two vertices), clustering coefficient (density of triangles in the network), and degrees (number of links connected to one node) are often used to determine vulnerability of a node. In these approaches, failures in the network are modeled by removing a node from the network. These node removal models remove nodes of decreasing degree order (Holmgren 2007, Hines et al. 2010). One of the negative aspects of the node-removal method way of determining critical infrastructure, especially energy infrastructure, is that removing nodes is computationally intensive, and may not allow for quick decision making.

Other research has focused on centrality measures within graph theory; nodes with high centrality have a larger impact on other nodes in the network (Demsar et al. 2007, Demsar et al. 2008, White and Smyth 2003). Demsar et al. (2007) compared the measures of degree, closeness, and betweenness. Degree (*d*) refers to the number of neighbors. Closeness refers to the shortest distance of a node to every other node, and is evaluated using (Equation 2.4).

$$C_C(v) = \sum_{t \epsilon V/v} 2^{-d_G(v,t)}$$ **Equation 2.4**

Betweenness refers to the number of links that pass through a node or more precisely the "proportion of shortest paths between every pair of vertices that pass through the given node." Calculations for betweenness are conducted using (Equation 2.5) (Rocco et al. 2011, pg 2).

$$C_B(v) = \sum_{s \neq v \neq t \epsilon V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$ **Equation 2.5**

Where: σ_{st}= total number of shortest paths from node s to node t

$\sigma_{st}(v)$ = number of paths between *s* and *t* passing through v

Desmar et al. (2007) found that betweenness is the best measure for determining vulnerable nodes in a network, because it is the only measure that was linked to the flow in the network. Disadvantages of other measures include 1) degree is only a local measure and 2) closeness describes only how well a node is embedded within the network (Desmar et al. 2007).

More complex graph theory concepts have also been developed to determine vulnerabilities in the power grid. Arianos et al. (2009) describe a variety of concepts related to power grid vulnerability. Arianos et al. (2009) describe efficiency and a critical component as one whose loss would cause the greatest loss of efficiency. Arianos et al. (2009) adapted the idea of minimum path length and global network efficiency to include the impedance of and the power that flows through each link, and used the resulting metric to determine the vulnerability of power grids.

Using the described graph theoretic techniques only assesses a network's topological characteristics. Graph theoretic approaches only take into account the nodes and links within the network, which disregards other external characteristics that might contribute to a network's vulnerability. Hines et al. (2010) indicate that topological graph theory solutions, while providing information about general vulnerability of a network, are misleading when viewed alone without ancillary information. Holmgren (2006) and Kim and Obah (2007) also indicates that a graph theoretic approach in electric power grid vulnerability studies has a variety of shortcomings; however, graph metrics have utility, especially in coarse scale studies (such as those at the country or continent level).

2.3.1. Graph Theory in GIS

Graph theory literature has had a place in the mathematics and engineering disciplines for many decades, but its research base in GIS literature is limited. There are two recognized data models for network data in GIS: topological data models and pure network data models (Curtin 2007). An example of topological data models is the Census Bureau's Dual Independent Map Encoding (DIME) data format. Then in 1980s the Census Bureau recognized the importance of maintaining topological properties for network data, especially roads, so the Census Bureau utilized the DIME data format, which was ultimately superseded by Topologically Integrated Geographic Encoding and Referencing (TIGER) (Curtin 2007). Curtin (2007) indicates that the "dual independent" refers to the identification of topological information between nodes and along links. Pure Network Data Models need to have the ability to support turns and directed links. Examples of pure network data structures have become available recently, such as Geometric Network and Network Data Set (ESRI) and Geographic Data Object Networks (Intergraph) (Curtin 2007).

Graph theory analysis in GIS is rather limited. Recent applications in GIS are linear referencing and routing. Routing often includes such measures as routing between locations, creating service areas, and creating origin-destination (OD) matrices (Curtin 2007). The limited applications of graph theory in GIS are likely due to network design and location problems. Newer, more recent advances, have tried to incorporate graph metrics into landscape connectivity and urban metric toolbars for ArcGIS (Goetz et al. 2009; Sevtsuk and Mekonnen 2012). Many tools for GIS have been developed looking at landscape networks and landscape connectivity with graph metrics, such as HabMod, ConnMod (Duke Marine Geospatial Ecology Tools), and FunConn (Space-Time Aquatic Resources Modeling and Analysis Program). Sevtsuk and Mekonnen (2012) introduced an open source approach, the Urban Network Analysis toolbox, which makes centrality metrics available to a wider consumer base. The Urban Network Analysis toolbox can calculate such centrality metrics as reach, gravity index, betweenness, closeness, and straightness; however, its main purpose is for urban networks.

2.4. Vulnerability Science

Understanding a system's vulnerability is imperative to its protection. Despite the importance of understanding vulnerability and the contributions that vulnerability science can make to the protecting critical infrastructure and other matters of national security, vulnerability is not well measured by many researchers and government entities. Different disciplines and even governments define vulnerability in different ways, which often may conflict and cause differences in research results. While the importance of consistent definitions across disciplines cannot be dismissed (National Research Council 2006), even more imperative is clearly defining what vulnerability science definitions are being implemented in this dissertation research.

2.4.1. Important Terms Defined

A variety of vulnerability science concepts will be utilized by this research and need to be clearly defined. Terms that need to be defined for this research include: vulnerability, risk, exposure, and criticality. Depending on which discipline the definition is derived from, these terms may be used interchangeably; however, they have fundamental differences between and among disciplines.

2.4.1.1. Risk

The definition of *risk* often includes, with varying degrees, components of all of the terms (vulnerability, exposure, and criticality). The Department of Homeland Security (DHS) defines *risk* as the "…potential for an unwanted outcome resulting from an incident, event, or occurrence, as determined by its likelihood and the associated consequences" (DHS 2008, pg. 24). This definition only tangentially mentions vulnerability and exposure; however, an extended definition specifically mentions vulnerability. Vulnerability experts during the Intergovernmental Panel on Climate Change (IPCC) define *risk* more robustly as the "…possibility for adverse effects in the future…, " but the IPCC indicates *that risk* has several components including *exposure* to a hazard and *vulnerability* (Cardona et al. 2012, pg. 69). This research will utilize Cardona et al.'s definition of *risk*, where risk is a function of *exposure* and *vulnerability*.

2.4.1.2. Vulnerability

Vulnerability is another determinant of risk, as described by Cardona et al. (2012). The definition of vulnerability is by far the most debated and least agreed upon. This research draws from a variety of different disciplines, including geography, vulnerability science, critical infrastructure, geographic information science, and national security, whose definitions of vulnerability are very similar, but are also different. In the hazards discipline, vulnerability is often seen as ". . . the propensity of exposed elements such as human beings, their livelihoods and assets to suffer adverse effects when impacted by hazard events" (Cardona et al. 2012, pg. 69). However, this definition has changed and evolved over time from vulnerability as an intrinsic risk to a more dynamic, multidimensional, and process-oriented definition (Birkmann 2006). Many frameworks for discussing vulnerability have been proposed including place-based frameworks such as the hazards of place vulnerability model. The hazards of place model is an exploratory framework, which explores a variety of elements, such as hazard potential, geographic context, social fabric, biophysical vulnerability, social vulnerability, risk, and mitigation, to ascertain the unique vulnerability of places (Cutter 1996). Risk and mitigation work together to identify hazard potential. The hazard potential is combined with the social fabric for an understanding of social vulnerability, and hazard potential is also combined with geographic context to ascertain biophysical vulnerability. The intersection of social and biophysical vulnerability yields the vulnerability of places (Cutter 1996).

Vulnerability in the national security realm has a less robust definition. DHS defines vulnerability as a "…physical feature or operational attribute that renders an entity open to exploitation or susceptible to a given hazard" (DHS 2008, pg. 34). The extended definition makes no reference to any other source other than the physical vulnerability. Social vulnerability or other population characteristics are not described in this context.

Vulnerability is also described in the critical infrastructure research; however, there is no widely accepted definition of vulnerability for technical applications (Holmgren 2007). Holmgren (2007) describes vulnerability as the susceptibility to threats and hazards that reduce the ability of a system to maintain its critical function. Gnansounou (2008) utilizes a similar definition, defining vulnerability as the inability "…of a system to cope with selected adverse events" (pg. 3735).

This dissertation research will adopt, in large part, the DHS definition of vulnerability. While DHS's definition of vulnerability is not as holistic as the definitions and frameworks provided by the hazards discipline, it is a functional definition that will be well-understood by the intended end-users of this dissertation research. Additionally, with the dissertation research focus being on less developed countries (LDCs), many of the data required by vulnerability frameworks from the hazards discipline either are not available or are extremely difficult to obtain.

2.4.1.3. Criticality

The criticality of infrastructure system components is important to identify and understand. Some research in electrical infrastructure use the terms criticality and vulnerability interchangeably; however, criticality and vulnerability are inherently different. Criticality is defined by the loss of a particular network component causing the loss of a critical function. Research often uses topologic characteristics of a graph to articulate criticality in a network system (Demsar et al. 2008; Zio and Sansavini 2011).

There is an essential difference between criticality and vulnerability that most critical infrastructure research does not enumerate. The difference between these terms is that is that each critical infrastructure system has a critical mission, which is often defined by the decision maker and may change at any given time (Quirk and Fernandez 2005). For example, for energy infrastructure, the critical mission may be to maintain power to the citizens, or even more simple than that: maintain power to a hospital. Just because a particular node is critical to this mission (maintaining power), thus, having a high criticality, does not necessarily mean it is vulnerable (Quirk and

Fernandez 2005). Criticality may be a component of vulnerability, but it is not synonymous.

2.4.2. Attributes Affecting Critical Infrastructure Vulnerability

One important aspect of modeling critical infrastructure is determining the substation service area. If a substation only serves a small population, the substation is not as vulnerable as a node that serves a large area (Bush et al. 2003). The magnitude aspect of vulnerability analysis directly ties into the population served or affected in the case of a disturbance (Schintler et al. 2007). If the substation serves a large population, then the service area's importance may be much greater than a substation that does not serve a large population. Additionally, the population served may fluctuate during the day and during the night, which may make certain substations more vulnerable during the day and others more vulnerable during the evening. In addition to population, the businesses affected may also be an important influence on the vulnerability of a substation (Schintler et al. 2007).

Koger (2008) also identifies several factors that can shape the vulnerability to critical infrastructure. These factors are broken categorized by: societal, system-related, technological, natural, and institutional. Societal factors include attractiveness for attack, public risk awareness, and demographics. System-related factors include the complexity and interconnectedness of the network. Technological factors include failure friendliness and infrastructure related operating principles. Natural factors include availability of resources and natural hazards. Finally, institutional factors include historic structures, legislation, and market organization (Koger 2008) (Table 2.3).

2.4.2.1. Climate and Energy Infrastructure

Research investigating the impacts of natural disasters and climate change to critical infrastructure, such as energy grid, is sparse. This is in part due to resolution differences between climate data and infrastructure data. In the past, most climate data and climate impact studies are conducted at a coarse spatial resolution not appropriate for combination with fine scale critical infrastructure data. Recent climate data, however, has been becoming available at finer spatial resolutions, making them easier to integrate with power grid modeling. While the recent advancements in climate data is extremely favorable for the electrical grid modeling community, little guidance for incorporating this new climate data has been established (Bhaduri 2013).

Table 2.3: Factors influencing vulnerability (after Koger 2008).

Category	**Factors**
Technological	Failure Friendliness
	Operational Principles
Natural	Resources Available
	Natural Hazards
Societal	Attractiveness for Attack
	Public Risk Awareness
	Demography
System Related	*Topology*
Institutional	Market Organization
	Government Policy
	Historical

2.4.2.2. Determining Service Areas for Electrical Infrastructure

Very little research has been conducted in estimating electrical substation service areas in the GIS. The fundamental research problem is based on work in location-allocation literature (i.e. the location problem of assigning customers to service locations). Allocation models are widely available, such as the Network Analyst module with the ESRI ArcMap GIS. Network Analyst in the ESRI suite does compute service areas; however, the major usage of this tool is for determining the service area of businesses based strictly on cost (e.g. Euclidean distance or weighted distance). The constraints for the service areas for businesses are tied to distance and time (drive time), while the constraints for electrical grid service areas are very different.

Research in service area definition for electrical infrastructure has utilized Voronoi and cellular automata (CA) approaches. The CA approach is an iterative approach, where substation cells gain service area by claiming neighboring cells. The area continues to expand iteratively until the area meets a neighboring substation service area, or other constraints, such as a substation's total capacity, have been reached (Fenwick and Dowell 1999). Fenwick and Dowell (1999) discuss methods of accuracy assessment for the CA model of determining substation service areas; however, they provide no accuracy assessment of their work. LANL has developed the Constrained Cellular Colonization (C^3) method, which can also be utilized for determining service areas and outage areas for electrical transmission grids (Bush et al. 2003). Oak Ridge National Laboratory (ORNL) developed a similar, but more robust, modified CA approach to determining service areas, which includes using population data from LandScan in addition to the load and location information of the substations. This approach uses the Moore neighborhood (square-shaped neighborhood) versus the von Neumann neighborhood (diamond-shaped). Neighbors are captured iteratively based on the supply sources within the specified radius, and the algorithm continues to capture cells until all non-zero cells are accounted for (Omitaomu et al. 2008). These methods have only been tested in the United States and have yet to be tested in less developed countries. Additionally, CA can be easily integrated into GIS, as GIS shapefiles can be exported from the service areas generated by the CA model (Bush et al. 2003).

The Voronoi diagram has also been used in research to approximate electrical service areas (Okabe et al. 2008, Akabane et al. 2002, Netwon and Schirmer 1997). Voronoi diagrams, also known as Thiessen Polygons, divide area into regions based on a set of points. In Voronoi diagrams, the polygons are constructed such that every location within the polygon is closer to the point it contains than any other point. Akabane et al. (2002) utilized the Voronoi diagram to approximate service areas when determining optimal locations of power quality control centers.

2.5. Summary

This dissertation research needed to draw literature from a variety of different disciplines; it needed to draw from literature in GIS, Graph Theory and Mathematics, and Vulnerability to develop the most informed research possible. In many cases, the literatures from these individual disciplines use different definitions of the same terms. Despite these discrepancies, the literature between the disciplines is often complementary and easily integrated.

CHAPTER III

METHODS

The methods chapter is broken up into several sections. The first section discusses the study area for this dissertation research, and the justification of its choice. The next section describes the data utilized for this dissertation's modeling framework. The third section discusses the data preparation and overall modeling framework. The final section discusses how the framework was evaluated.

3.1. Study Area

The study area for this research consisted of Southeastern Asia (India, Bhutan, and Nepal). This region was particularly interesting with regards to its electrical grid connections (discussed in section 3.1.1), especially because of the political instability that may ensue as a direct result of a loss of electrical power. Certain areas, such as Bhutan, Nepal, and parts of India have had little western influence and have large cultural differences from other parts of the world, which may have an impact on how the factors are weighted in this area versus another area. This modeling framework is meant to be applicable worldwide. Thus, a research area in Southeast Asia and a smaller study area in the western world (discussed more fully 3.4.4) demonstrated the design's versatility. The same factors may be used and weighted differently, or different factors can be added to further highlight the uniqueness of an area.

Much of the previous research conducted on the vulnerabilities of energy infrastructure focused on developed countries. This was not only true of energy infrastructure vulnerability studies, but studies involving vulnerabilities, hazards, and disasters in general. The previous research focus on developed countries is likely due to data availability and the lack of energy infrastructure data for developing countries (National Research Council, 2006). Review of available literature that used real-world data showed that the majority of research regarding critical infrastructure vulnerability and criticality has taken place in the developed world, especially North America and Europe. Africa, Asia, and South America have been under-represented in critical infrastructure vulnerability and criticality literature. With the lack of research in these particular regions, it was of key interest to use one of these areas as a test bed for this dissertation research.

Considering the lack of research in Asia and media interest in India's energy infrastructure sector, it was of interest to study Southeastern Asian electrical grid vulnerabilities. Southeast Asia is also extremely susceptible to environmental changes. These environmental changes are important for identifying areas in which the energy infrastructure will be impacted by impacts of climate change (Warner et al. 2009, Parry et al. 2007). With the region's susceptibility to climate change, the lack of literature on energy infrastructure in Asia, and unique impacts of critical infrastructure on the society, Southeast Asia was an ideal location to conduct this research. The dissertation Southeast Asia study area was defined as: Nepal, Bhutan, and India (Figure 3.1).

Figure 3.1: Dissertation study area of India, Bhutan, and Nepal.

3.1.1. Study Area Energy Grids

The study area only includes India, Bhutan, and Nepal, because these countries have the energy grid interconnections in the region. Currently, Nepal both imports and exports power to and from India. Bhutan, on the other hand, is a chief exporter of power to India. None of the other countries in the region, including Bangladesh and Pakistan, have currently exploited the potentialities of exporting or importing power regionally. However, in some cases, such as Bangladesh, there are discussions of regional power trade with a Bangladeshi-Indian interconnection (Mahmud 2012). Pakistan also has no interconnections with India. Discussions were raised between India and Pakistan about connecting their power grids; however, the price that Pakistan was offering to charge India in exchange for its surplus power was and is the major impediment to connecting these two country's power grids (USAID, 2009).

3.1.1.1. Indian Energy Grid

The Indian Energy Grid is extremely complex. Each individual Indian state has their own electricity board, which maintains most of the substations and transmission lines within a state. There are five regional grids in India: the Northern, Western, Eastern, North-Eastern, and Southern regions (Figure 3.2). The Northern Region covers the largest spatial area, and it is comprised of the states of Rajasthan, Uttar Pradesh, Uttaranchal, Himachal Pradesh, Punjab, and Jammu and Kashmir. The installed generating capacity of the Northern Region is approximately 56,000 Megawatt (MW), with one MW serving approximately 1,000 homes, utilizing both thermal and hydroelectric generation. The Western Region is comprised of the states of Gujarat, Madhya Pradesh, Goa, Chhattisgarh, and Maharashtra, with an installed generating capacity of approximately 67,000 MW. Energy generation in the Western Region is dominated by thermal generation, with smaller parts hydroelectric, renewable, and nuclear generation. The Eastern Region is comprised of the states of Bihar, Sikkim, West Bengal, Jharkhand, and Orissa. The Eastern Grid has an installed generating capacity of approximately 27,000 MW, with thermal, hydroelectric, and renewable generation sources. The North-Eastern Region consists of Arunachal Pradesh, Assam, Nagaland, Manipur, Meghalaya, Tripura, and Mizoram. The North- Eastern Region has an installed generating capacity of approximately 2,500 MW, coming from thermal, hydroelectric, and renewable energy sources (Ministry of Power 2012). The Southern Region includes the states of Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu, with an installed generating capacity of approximately 41,000 MW. Energy generation comes from a variety of sources including thermal, nuclear, hydroelectric, and renewable sources (Southern Regional Load Dispatch Center (SRLDC) n.d.). These five regions are interconnected through a limited number of transmission lines of varying voltages ranging from 220 kV to 765 kV. These interconnections help to distribute the generated energy from one region to another (Ministry of Power 2012).

Figure 3.2: Energy Regions of India

3.1.1.2. Bhutan Energy Grid

Bhutan's energy grid is not nearly as extensive or complicated as the Indian grid. The minimalistic Bhutanese grid is understandable considering that India is almost 70 times larger than Bhutan in terms of area, and over 1500 times (1,205,073,612 people in

India versus 716,896 people in Bhutan) greater in terms of population (Central Intelligence Agency 2012). Despite these vast differences in area and population, Bhutan is the only country in South Asia with surplus power production that contributes to its economy (Asia Development Bank (ADB) 2010). The export-import relationship between Bhutan and India began as early as 1988 with the commissioning of the Chhukha Hydropower Plant. The Chhukha power plant included not only a domestic network, but also connections to the Indian grid. After the commissioning of this power plant, the Department of Power (DOP) was split up into a variety of different entities including the Bhutan Power Corporation (BPC), Department of Energy (DOE), and the Bhutan Electricity Authority (BEA) (ADB 2010).

Most of Bhutan's power generation comes from a mix of hydropower and diesel generators. With Bhutan's steep slopes and abundant water resources, it is not difficult to understand that of the two, the majority of the country's power generation is from hydropower. Between 2004 and 2009, several hydropower plant projects were completed and commissioning increasing the installed capacity of Bhutan from 438 MW in 2004 to 1,498 MW in 2009 (ADB 2010). Currently, the only country that Bhutan exports its excess generation to is India (Kuensel 2012).

3.1.1.3. Nepal Energy Grid

Similar to Bhutan, Nepal is a landlocked country situated to the north of India. Also similar to Bhutan, Nepal has vast hydropower resources, with hydropower accounting for 88 percent of the installed generation capacity of the country, with the rest of the installed capacity coming from multi fuels and diesel plants (South Asia

Regional Initiative for Energy (SARI/E) 2012). Nepal and India also share energy through a bilateral trade treaty signed in 1996, though Nepal is a net importer of energy (Sarkar 2012; SARI/E 2012).

3.2. Data and Sources

This dissertation research required a great deal of data mining, preparation, and generation, especially when the study area was comprised of less developed countries (LDCs). The availability and quality of electrical infrastructure geospatial data in these regions is often sparse. The quality of the data is typically undocumented or not assessed. The paucity of good quality data has often precluded analyses of the electrical infrastructure; however, given abundant levels of time resources and current technology, the creation of these datasets was feasible.

3.2.1. Geospatial Data Model Design

This dissertation modeling framework had a variety of important geospatial components to take into consideration. First is the study area, which was defined as Southeast Asia. Secondly, the unit of analysis was determined to be the service area level. The unit of analysis is one of the most important components of this model, as it will determine a unit's vulnerability. If the unit of analysis was set to the state, very different vulnerability rankings would be produced. Service areas were chosen, as they are the most important unit when describing, geospatially, the area of importance, or influence, of a substation. For determining service areas, the most important attributes were the substation location and its capacity, as the service areas were delineated based on their geographic location. Service area delineation is discussed more fully later in Chapter 3.

The next most important part of this dissertation framework is the factors. One of the objectives of this dissertation research was to determine if graph metrics were more effective than geospatial attributes. With this in mind, the topology of the study area electrical infrastructure needed to be generated to obtain topological information. The location of the substations and power plants was very important in determining the topological characteristics of the grid. While the location of the transmission lines is also important, it is not as important as the location of the substation, because for many of the metrics the distance is calculated by Euclidean distance. Other factors, such as population and natural hazards, were important for assessing the importance of geographic variables to substations vulnerability.

One of the objectives of this dissertation research was to determine what characteristics, or attributes, was most likely to affect the vulnerability of a particular node. The following table and sections detail, based on an extensive literature review, the necessary data to determine the vulnerability of a node, and the sources that were used in this research (Table 3.1).

Table 3.1: Summary of the datasets utilized for analysis of Indian Energy Grids.

Data	Source	Format
Electrical Grid Network (Power stations, Substations, Transmission Lines, Capacities)	Generated by Researcher using Indian Maps and Google Maps™	Vector
Population	Oak Ridge National Lab – 2010 LandScan Global	Raster
Service Area	ArcGIS Thiessen Polygons/Oak Ridge National Lab (algorithm generated)/PoDiuM	Vector
Natural Hazard Physical Risk	United Nations Environmental Programme	Raster
Climate Extremes	WorldClim (Hijmans et al. 2005)	Raster
Additional Critical Infrastructure	Google Maps ™, Wikipedia™	Vector
Land Cover	Oak Ridge National Lab	Raster

3.2.2. Electrical Grid Networks

An electrical grid network structure and its associated attributes are often considered proprietary, and thus, are often not publically available. Datasets for developed countries, such as North America and Europe, have been analyzed fairly widely; however, as stated in previous sections, Asia has largely not been analyzed, partially due to data availability issues. Oak Ridge National Laboratory (ORNL), in conjunction with the National Geospatial-Intelligence Agency (NGA), is currently developing a Global Transmission Grid, utilizing only open source (e.g. available to the public) avenues of data collection to produce an open source international electric grid available to researchers.

Considerable time commitments were required for the development of a dataset of this magnitude. Bhutan and Nepal have relatively small transmission networks; however, India has a very large and elaborate network. In addition to developing the structure of the grid, it was also necessary to determine the attributes associated with each component. For substation and power plants the necessary attribute was capacity in megawatts (MW), and for transmission lines the necessary attribute was voltage in kilovolts.

The first step in creating the energy network layer was to find an existing map of the structure of the network. In some cases these were posted on the Internet (Figure 3.3). These maps were georegistered and digitized for incorporation into a GIS-based data model.

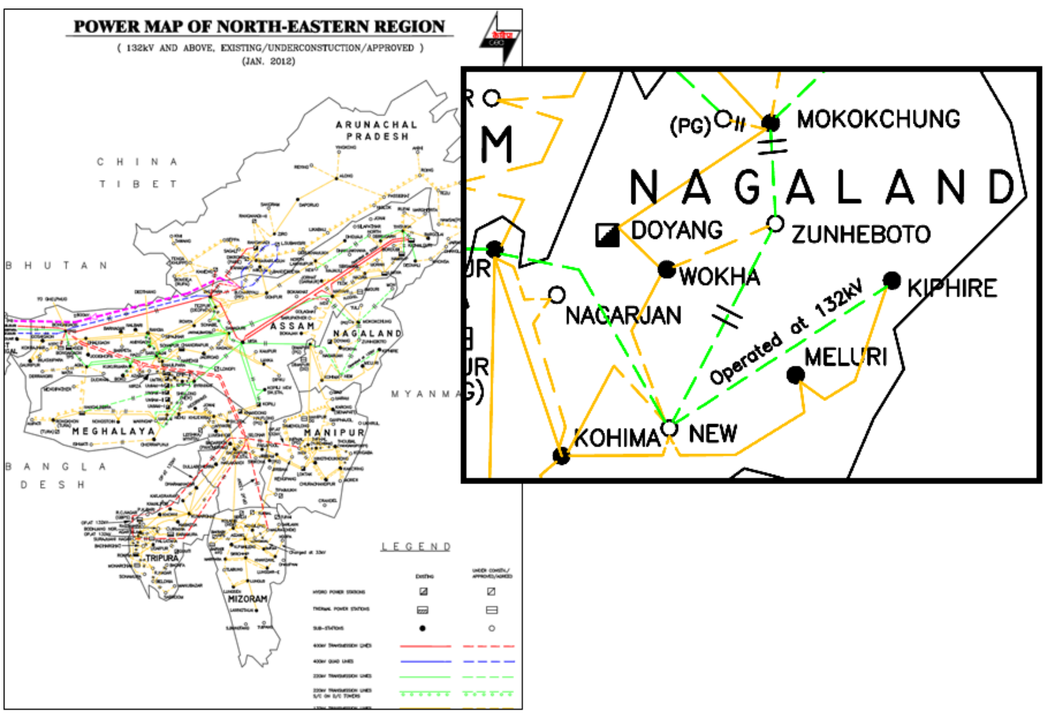


Figure 3.3: Example of map for digitizing (Source: Central Electric Authority (CEA) 2012).

Rarely were the digitized features accurate portrayals of the location of the elements in geographic space. These elements needed to be quality checked to ensure that they were in the accurate position. Often, high resolution imagery is used to ensure the accuracy of substation, power plant, and transmission line locations. Unfortunately, high resolution imagery is usually not available for this part of the world free of cost. Google Maps™ and Wikimapia™ were often adequate alternatives to purchasing high resolution imagery. Additionally, Google Maps™ and Wikimapia™ sometimes had descriptions that were useful (or at times misleading) in locating substations and power plants. These sources; however, had a limited usefulness for insuring the quality of transmission line locations.

Additionally, locating substations in developing countries was its own challenge. Generally, substations can be found on the outskirts of a city, and typically, at least in the developed world, are underlain with concrete or gravel (Hayes 2005). These substations characteristics should make them relatively easy to locate on imagery given the city names where they are located. In Southeast Asia determining substation locations were not easy. Sometimes the substations were underlain with gravel or concrete; however, many times the substations were not. The substations were positioned above a bed of grass or other natural environs, which made the substations especially difficult to locate (Figure 3.4).

The original geo-referenced maps often provided names of the substations that identified the closest city to the substation (Figure 3.4). However, when there was no name, or the name was ambiguous, such as "Southwest Station," it was often difficult or impossible to locate these stations without local knowledge. In cases such as this, the station was kept in the dataset to maintain the continuity of the grid, but it was noted in the attributes that the location was not ascertained and it was also noted in the overall accuracy of the dataset. No quality index was generated; however, the deficiencies in the data were noted for additional users. Additional problems included poor quality of imagery on Google Maps. In the case of poor image quality, the location of the substation's city was found and the substation was placed on the outskirts of the city (Hayes 2005).

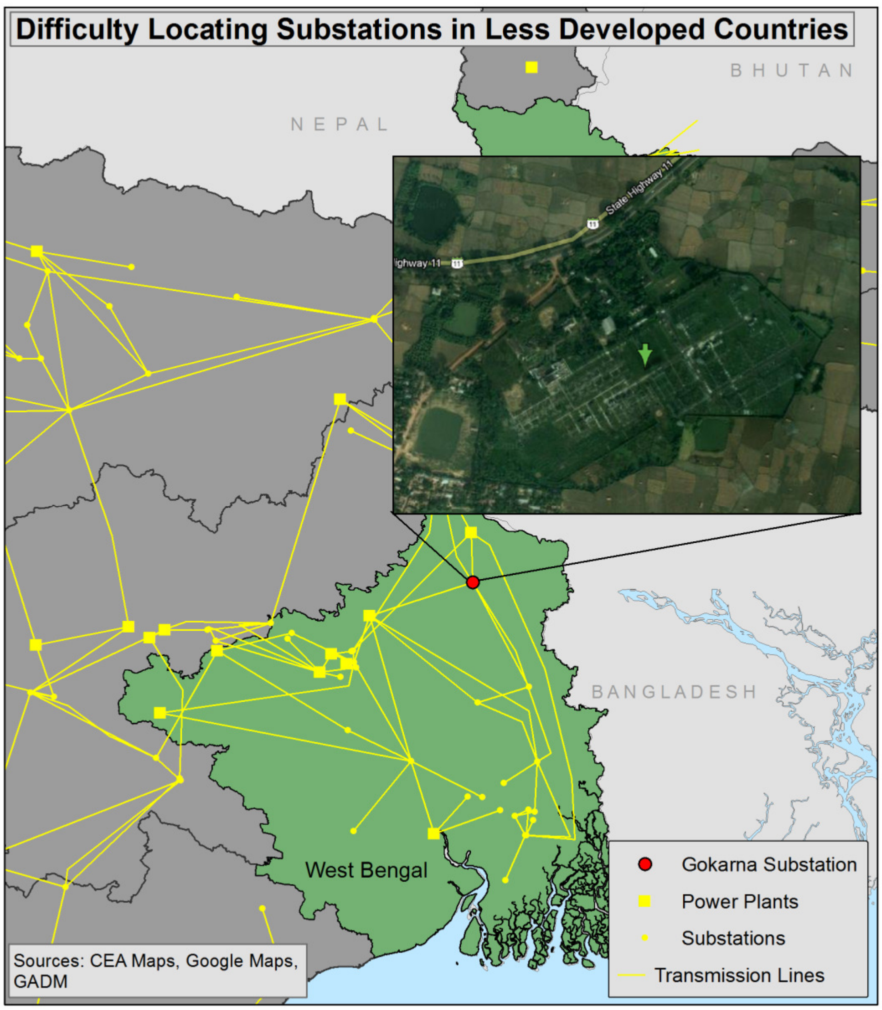


Figure 3.4: Challenges locating substations in less developed countries.

The attributes (such as capacity for substations and power plants and voltage for transmission lines) were ascertained from conducting a "web-crawl," where each substation and power plant was searched using search engines such as Google to determine its capacity. Only websites and documents from state, regional, and central government authorities were used to determine the approximate capacity of each substation and power plant. If possible, the capacities of the substations were verified on multiple websites to ensure accuracy; however, in most cases verifying capacity on multiple documents was not possible. If the capacity information for substations could not be obtained, it was estimated based on the attributes of similar substations in the area and noted in the attributes as estimation.

3.2.2.1. Description of Generated Grids

The grids for the three countries (India, Bhutan, and Nepal) of interest had varying levels of accuracy. Both Nepal and Bhutan had relatively small grids in comparison to the Indian Grid. Transmission lines in all cases indicated the Euclidean distance between connected nodes, because it was extremely difficult to locate the transmission lines using open source options, such as Google Maps. While the distance is important, the spatial location of the substations was more important. Also, it was impractical to accurately identify transmission line locations from open source imagery.

Power Generating Plants (PGPs) were located using a combination of the Global Energy Observatory and Google Maps ™. For all three countries, each PGP was geographically located and its generating capacity documented (Table 3.2).

Table 3.2: Located Power Plants.

Country	Power Plants Located	Power Plants Not Located
India	189 (100%)	0 (0.0%)*
Bhutan	4 (100%)	0 (0.0%)*
Nepal	25 (100%)	0 (0.0%)*

*All KNOWN PGPs, not necessarily all.

Substations were the most difficult to locate, as stated in earlier sections. India contained over 900 substations; Bhutan contained less than a dozen substations; and Nepal contained fewer than thirty substations. Substations for all three countries were identified at or above the 132 kV level. Locational accuracy was broken down by whether the substation was located, if only the city was located, or if no locational aspect of the substation was found (Table 3.3). As stated previously, the locational information of the substations is most important attribute of the electrical network. The capacity of the substations was separated by actual and estimated in Table 3.4. In the cases where the actual capacity of a substation could not be found, the capacity was estimated based on substations with known capacities of similar size, location, transmission line connectivity.

Table 3.3: Located Substations.

Country	Substation Located	Only City of Substation Located	Substation Not Located
India	856 (94%)	10 (1%)	41 (5%)
Bhutan	11 (100%)*	0 (0%)	0 (0%)*
Nepal	28 (97%)	0 (0%)	1 (3%)

*All KNOWN substations, not necessarily all.

Table 3.4: Capacity Data.

Country	Actual Capacity	Estimated Capacity
India	826 (91%)	81 (9%)
Bhutan	11 (100%)	0 (0%)
Nepal	27 (93%)	2 (7%)

Due to the lack of comparable data sources for this type of location and attribute data from these countries, there was no way of formally assessing the accuracy of attributes and location data. This description was a qualitative description of the general accuracy of the developed datasets.

3.2.3. Population

Population data was imperative to understanding the vulnerabilities and criticalities of any network. Critical infrastructure serves and benefits the people. Where there are more people, there are more people to be affected by the loss of a critical component, and thus poses greater vulnerability. Due to these facts, population distribution is very important, and ORNL boasts the world's finest spatial resolution dataset for global population. This dataset, called LandScan Global™, represents the 2010 ambient population (population averaged over 24 hours) per pixel and has a spatial resolution of 1 kilometer (km). It is distributed in an ESRI Grid format for easy utilization in GIS software, but it is also available in ASCII format for users of non-GIS compatible software (Bhaduri, n.d.).

3.2.4. Service Areas

Service areas were calculated using three methods. These methods included a Service Area/Outage Area (SA/OA) Calculation algorithm developed by ORNL, the

Voronoi method, and the Power Distribution Model (PoDiuM) developed for this research. The SA/OA algorithm requires three datasets: the supply (in MW), the geographic location of the substations, and the demand (population). The Voronoi method only requires the geographic location of the substations. PoDiuM requires the same data inputs as SA/OA (supply, geographic location, and demand), but in different forms described below.

The SA/OA algorithm uses a cellular automata (CA) approach in conjunction with substation location, capacity, and population counts. The algorithm required three inputs from Microsoft Excel™: a demand sheet, a supply sheet, and a combined worksheet. The demand and supply sheet mimicked a raster dataset, with the cells in the rows and columns representing pixels in a Cartesian surface. The supply sheet represented the locations and capacities (in MW) of the substations. The demand sheet represented the approximated demand (D) per cell. This was approximated using total supply (S), in Megawatts (MW), and population (P) (Equation 3.1).

$$D = \frac{P}{\sum P} * S \qquad \textbf{Equation 3.1}$$

The demand (in MW) for a cell was calculated by taking the proportion of regional population in the cell multiplied by the total energy supply for the region. The algorithm then assigned each cell to a service area for each substation by spiraling clockwise to the cells around it, continuing until the supply was exhausted by the service area demand (Omitaomu et al. 2009). The assumption with this allocation method was that the nearest locations (near in geographic distance) were strictly served by the nearest supply source. For the Southeast Asian study area, the results at times did not generate contiguous service areas. This was in part due to the algorithm being developed with United States engineering principles in mind, and may not translate well to another country, especially less developed countries. An additional disadvantage was that the SA/OA algorithm could only process small areas of the study area at a time. In many cases, entire states would have to be broken up, leading to discontinuous service areas that did not realistically resemble service areas that do not serve the entire population (Figure 3.5).

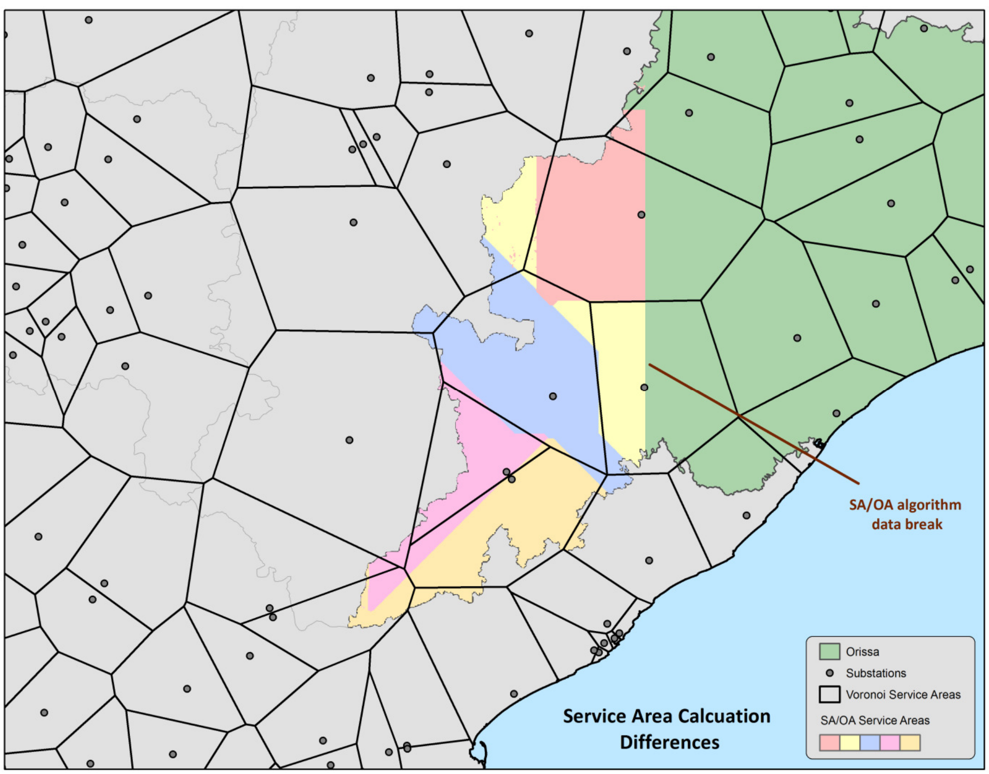


Figure 3.5: Data limitations of ORNL's Service Area/Outage Area (SA/OA) algorithm compared to the Voronoi Method.

The Voronoi diagram resulted in more contiguous service areas for the area of interest. The Voronoi diagram divided the area into zones based on substation location. The partitions were drawn such that the zones were "…full areas where any location within the zone is closer to its associated input point than any other input point" (ESRI 2012). The Voronoi method closely resembled those complete service areas produced by the SA/OA algorithm (Figure 3.5). There were several advantages to using the Voronoi diagram to the SA/OA algorithm. Firstly, the Voronoi diagram was easily implemented in a GIS environment. The SA/OA algorithm was processed outside of a GIS and must be imported after converting the output to an ASCII file. Additionally, the Voronoi method produced service areas for the entire study area, without needing to divide the study area, and provided results much more expeditiously. The largest benefit of this method was that there is no requirement for attribute information for the substations, such as capacity. This would increase the versatility of the model and allow it to be more widely implemented.

One of the limitations of the Voronoi method, however, was that it allocates the entire geographic area to a service area without taking into account areas that may not be served. Especially in Southeast Asia, not every area is served by a substation. There are areas and populations that are not served by the electrical grid, and the Voronoi method had no way of taking this fact into account. A new approach, developed by this dissertation research, called PoDiuM, was a cost distance based algorithm that applied distance in cost units versus geographical units. At a basic level, PoDiuM creates service areas by allocating cells of highest demand to nearby substations until the capacity of those substations are completely accounted for. This algorithm searches for cells of highest demand to allocate to substations in the form of service area. This approach took a similar approach to the SA/OA algorithm, but transformed it into a cost approach, where the lower the demand, the higher the cost, and the greater the transmission distance, the larger the cost. In this way, service area cost was the inverse of demand (Equation 3.2).

$$D_{inv} = \frac{1}{\frac{P}{\sum P} * S} \qquad \textbf{Equation 3.2}$$

For example, consider a given study region with a total population of 1000 people, and a total substation capacity of 100 MW. If the cell of interest has a population of 1, the Demand (denominator of Equation 3.2) would equal 0.1. Thus D_{inv} would equal 10. Similarly, if a cell of interest in the same study region had a population of 100, D_{inv} would equal 0.1. This indicated that cells with less demand were more costly to include in the service area. So, in a choice between the first cell of interest and the second, the algorithm would choose the second. Figure 3.6 describes the data flow for PoDiuM.

The inputs required a raster population distribution surface (raster), substation location data (vector), and capacity information (tabular). From the population, all areas with a population of 0 are excluded from the analysis, and the overall sum of the population in the area of interest is calculated and stored using summary statistics.

Similarly, the overall capacity of the area of interest was calculated and stored using summary statistics. These two values (population sum and capacity sum) were then used to calculate the demand, which was then inverted. A raster dataset was created based on the inverse demand field, which was used as cost raster for the cost allocation.

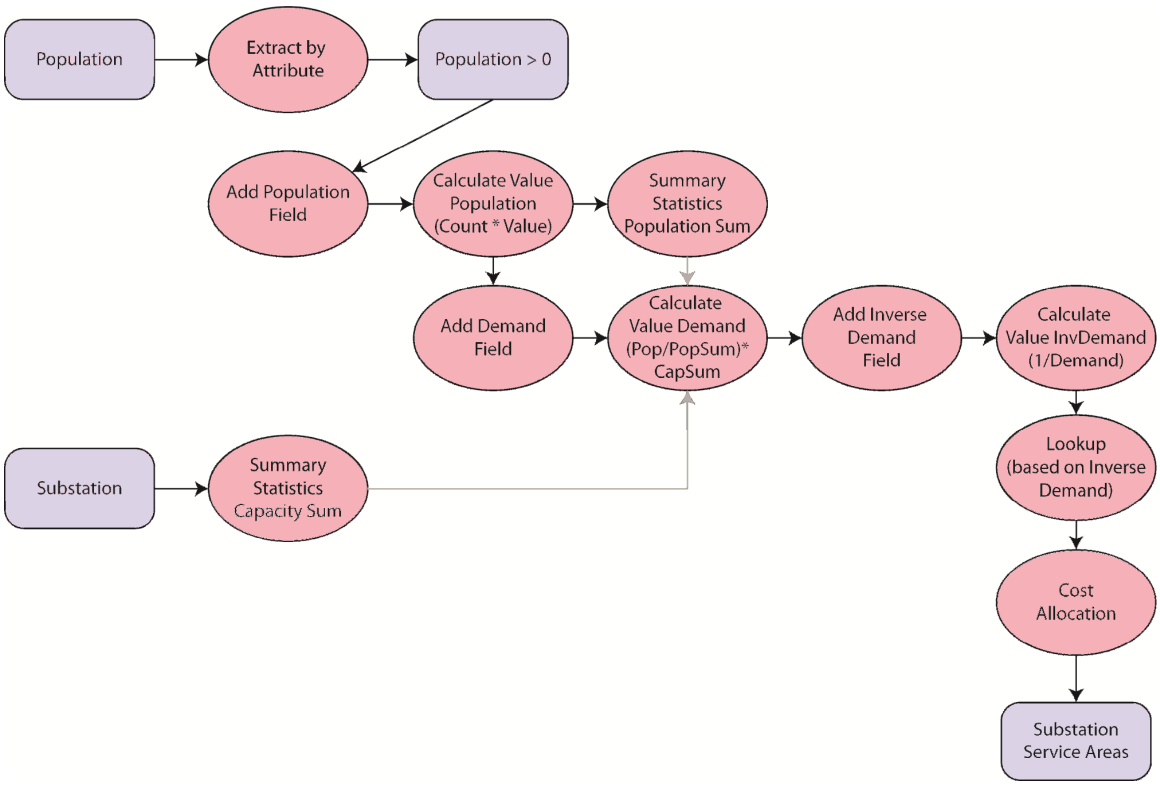


Figure 3.6: Data flow model for new service area algorithm.

The cost allocation algorithm was the crux of calculating service areas for PoDiUM. As stated earlier, cost can be equated to the inverse of demand. The SA/OA allocates cells with the highest demand, but in a cost allocation approach, high values (high demand) are avoided. By inverting demand, the highest demand values are of the lowest cost, and would be most attractive to the algorithm for allocation. The cost allocation required two inputs: a source (substations) and a cost raster (inverse of the demand). The cost allocation basically identified which cells were allocated to an associated source based on the lowest accumulated cost to reach the source. Though similar to Euclidean allocation, cost allocation does not use geographic units, it used cost units. The cost raster defined the impedance to movement through the cell, derived from the cost and the direction of movement. The cost per cell was calculated by taking the cost of the cell and multiplying it by the spatial resolution. When discussing movement, moving linearly from one cell to another was calculated by taking the cost of one cell, adding it to the cost of its neighboring cell, and dividing by two. If the movement was diagonal, the cost to travel is 1.414 times the cost of one cell added to the cost of its neighboring cell and divided by two. The accumulated cost was the total cost to move from one cell to another (Figure 3.7). In the case where the cost was over multiple cells, the costs were added to move between each pair of cells together to get the accumulated cost.

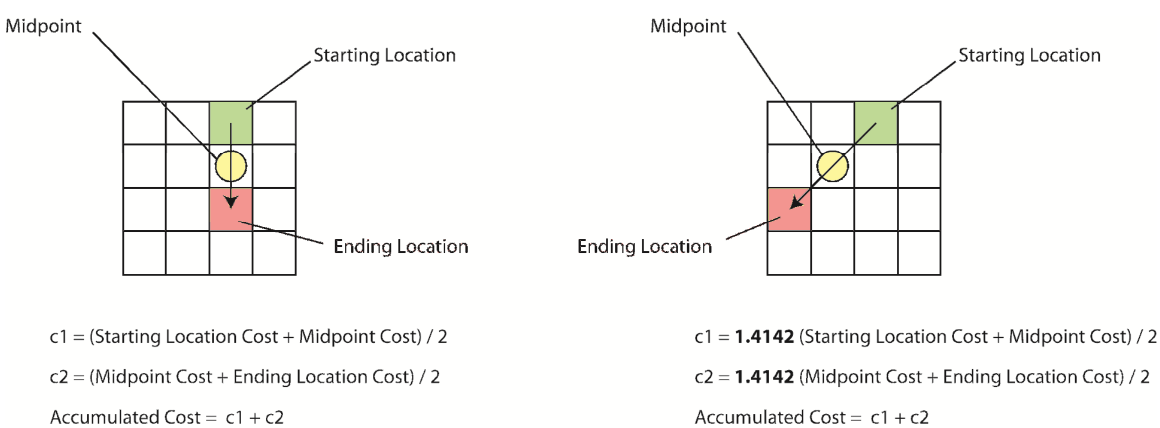


Figure 3.7: Calculating accumulative cost (after ESRI 2007).

The cost allocation algorithm was an iterative process (Figure 3.8). In the first iteration, the neighbors of the source cell were analyzed and the cell with the lowest accumulative cost was allocated to the source. Source cells were given a cost value of 0, as there was no cost for the source to return to itself. In the next iteration, the neighbors of the allocated cell were included in the analysis, as they now have a path to the source. Once again, the cell with the lowest accumulated cost was selected for allocation to the source. Changing of allocated cells was possible if the algorithm found a new cheaper route. The process continued until all of the cells were allocated, or an optional maximum distance threshold was met. The maximum distance threshold was the accumulated cost that could not be exceeded. This allowed for the realistic notion that there were areas in a study area that were not served by any service area.

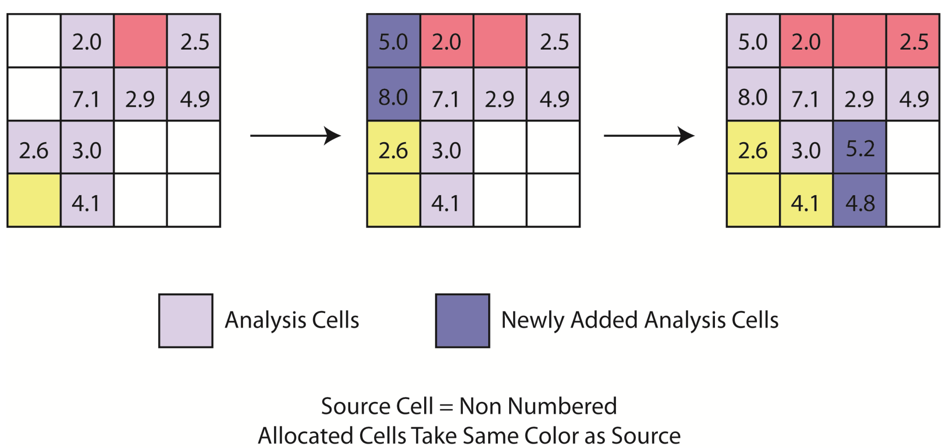


Figure 3.8: Cost allocation basic method (after ESRI 2007).

There were advantages and disadvantages to all three algorithms. The results of the newly developed PoDiuM approach were contiguous service areas that take into account population and capacity of the study area (Figure 3.9). The SA/OA algorithm was more computationally intensive, took more time to pre-process the data into the format ingestible into the algorithm, and took more transformations to ingest the outputs into a GIS. The Voronoi method, while very computationally efficient (calculating service areas in less than one minute versus 9 minutes for the entire study area for PoDiuM), allocated the entire geographic area, which was less accurate for areas of interest. Additionally, the Voronoi method was not appropriate for forecasting future changes in the grid. PoDiuM required less data pre-processing and less computational resources than the SA/OA algorithm, but it was more detailed and realistic than the Voronoi method. In terms of computational resources, PoDiuM was used to calculate service areas at the 1-km scale for the contiguous United States in approximately 20 minutes. On the same system, ORNL's SA/OA algorithm, implemented in MATLAB would take approximately one week. PoDiuM balanced the weaknesses and strengths of both methods (SA/OA and Voronoi) to produce efficient, realistic service areas. One additional benefit of PoDiuM was that the model can evolve to match the needs of grid modelers in the future. For instance, present service area models do not take into account the demand that is not served, but PoDiuM does. PoDiuM utilizes cost allocation, where the user can set a maximum, thus limiting how many cells are acquired by a given substation. Future iterations of PoDiuM will take into account the evolution of the grid, including how and where to build new assets in response to unallocated demand or future changes in demand.

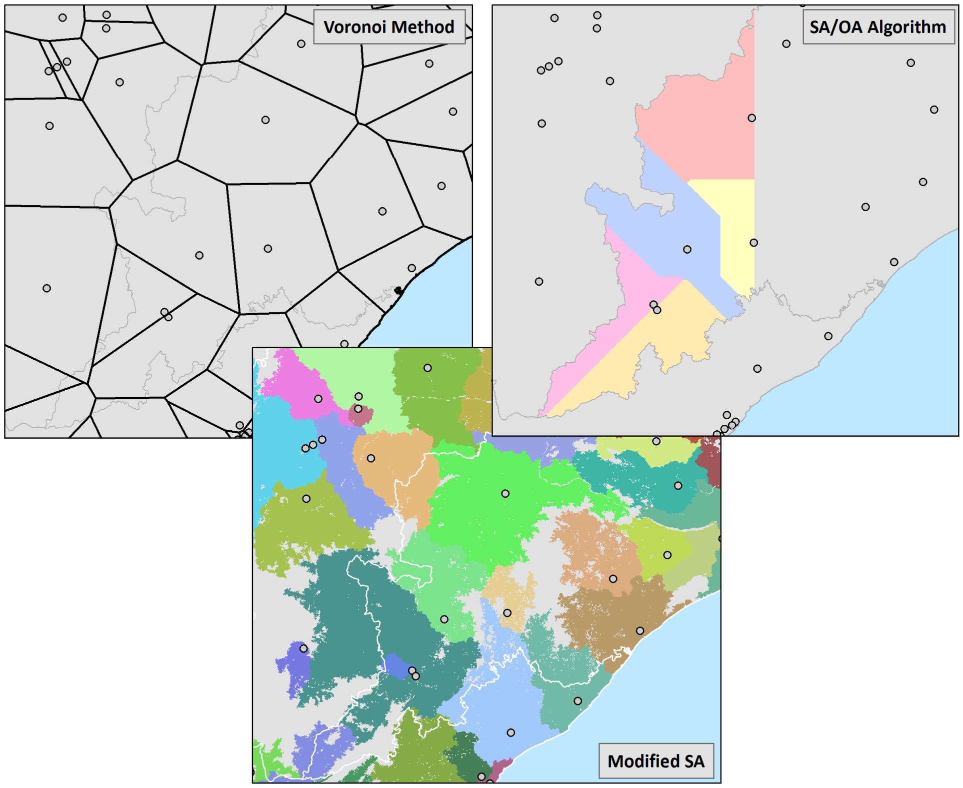


Figure 3.9: Varying outputs of the three algorithms.

Validating service area algorithms is difficult due to lack of reference data. Past validation attempts have used natural disaster power outages to validate service areas. For example, if two substations are known to be not operational, and the number of people impacted is known, it can be used as validation for the approximated service areas. LandScan Global can be used to find the number of people in the approximated service areas for the impacted substations, and these numbers can be compared with the number of known customer outages.

Calculating the service area was of the greatest importance to this research since service areas were used as the geographic unit of analysis. Points could not be used as the unit of analysis, as the attributes that influenced the vulnerability of that substation were not just located at the point, but allocated to the entire service area (Figure 3.10). Unfortunately, there was no way to validate the service areas, as these are proprietary boundaries that are not often shared by utility companies. While electrical company areas of influence boundaries may be available, actual substation service areas are not.

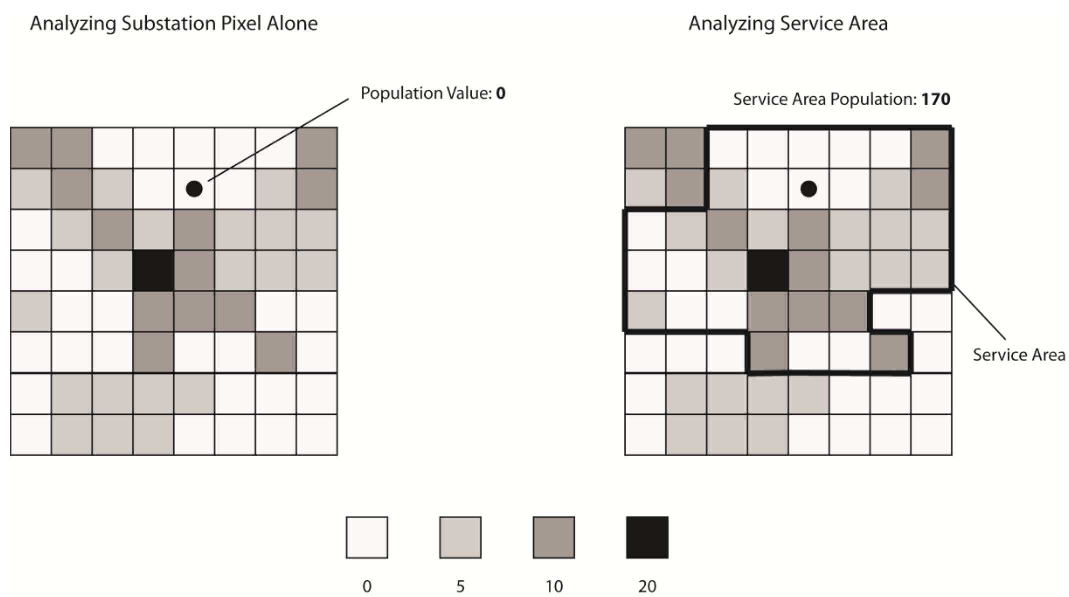


Figure 3.10: Unit of analysis justification.

The capacities of the substations in the study region ranged 1 to 1120 MW. Figure 3.11 represents a histogram of the frequencies of the capacities in the study region. The most frequent substation capacity was 70 MW which served a population of approximately 50,000 people.

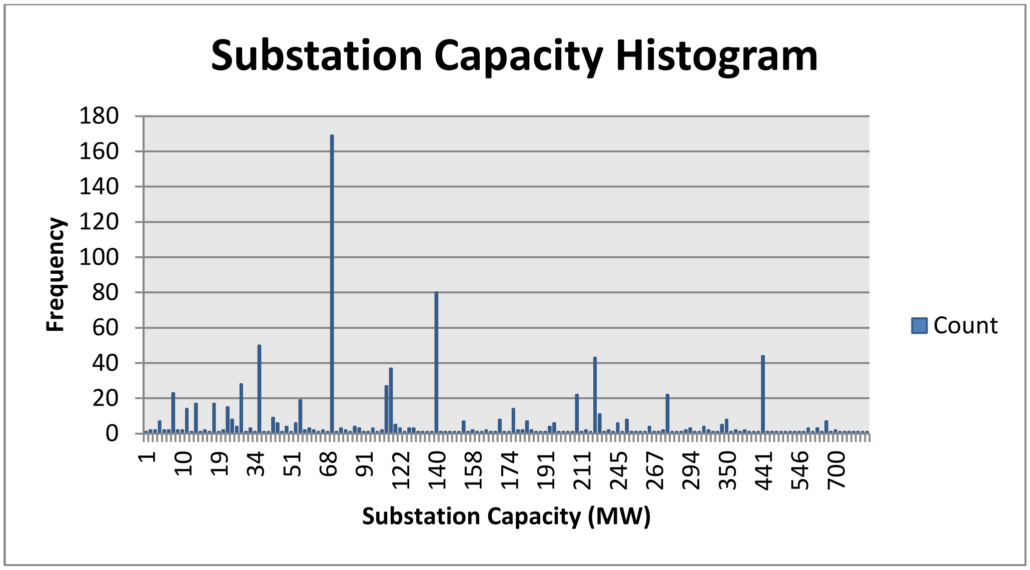


Figure 3.11: Histogram of substation capacity frequencies.

3.2.5. Climatic Extremes and Natural Hazards

3.2.5.1. Natural Hazard Frequencies

Natural Hazards and climatic extremes have an effect on both the physical condition of the critical infrastructure and its performance (Wilbanks et al. 2012). Natural disasters, such as earthquakes, landslides, and cyclones, are prevalent in the study area and can affect the physical condition of the power grid and thus its performance. Many datasets are available including global risk data from UNEP and from Center for International Earth Sciences Information Network (CIESIN) at Columbia University. The use of the Multi-hazard Risk from United Nations Environment Programme (UNEP) or Columbia University's Global Multihazard Frequency Distribution may cause other factors, such as population and infrastructure, to be duplicated in this model. Due to the duplication limitation, the multi-hazard datasets were not included in this research.

Both organizations have data disaggregated into individual hazard frequency, which were used to help identify the physical vulnerability of an area. The advantage of using the disaggregated data was that it did not include population or other additional data enhancements. The advantage of not having population was that it would not be included twice, as LandScan already accounted for population. Also, the user can mix and match the individual hazard frequency data to best serve their interests. This dissertation research focused on three of the major and most frequent hazards that affect this region: landslides, earthquakes, and cyclones. Table 3.5 documents what datasets were chosen for inclusion.

Table 3.5: Hazard datasets used in the analysis.

Hazard	Dataset Chosen	Justification
Earthquake	UNEP	• Higher Spatial Resolution (2 km)
Landslides	UNEP	• Higher Spatial Resolution (1 km) • More information (CIESIN only presents landslides with a frequency greater than 6)
Cyclones	UNEP	• Higher Resolution (2 km)

3.2.5.2. Climatic Extremes

Climatic extremes also cause vulnerabilities in the power grid (Wilbanks et al. 2012). Generally, warmer than average temperatures and colder than average temperatures affect the demand for power and thus add additional stress on the power grid (Wilbanks et al. 2012, McMorrow 2011). Hijmans et al. (2005) developed a global gridded dataset of monthly average maximum temperatures between 1950 and 2000 at a 1-km resolution, which was used to describe the magnitude of temperature divergence from average on a given day. One limitation of this dataset is that is lacks data from the latest 10 years. The omission of these 10 years of data may impact the results of the analysis. Raw maximum temperature for the two days of the Indian blackout were obtained from the National Climatic Data Center (NCDC) for the entire study area (approximately 90 weather stations) and converted into a raster surface for use in this model.

3.2.6. Additional Critical Infrastructure

Critical infrastructure (CI) comes in a variety of different types ranging from roads to monuments to hospitals. No data set was readily available for critical infrastructure in the study area. As such, CI data was manually generated for the three countries in the area of interest. Four major datasets were generated for each country: Major Airports, Major Hospitals, High Courts, Major Monuments, and Major Religious Places. There were several reasons for the selection of these categories; however, these categories are not exhaustive of the possibilities for additional critical infrastructure datasets. The major reason for the selection of these datasets was the data's ability to approximate important features. Hospitals were a proxy for the lifeline utilities critical infrastructure, whereas high courts were a proxy for the critical infrastructure known as continuity of government infrastructure. While many of the lists of CI certainly were not exhaustive, it provided a base for analysis. Additionally, due to time constraints and the necessity of manually creating each dataset, the amount of data was limited to those datasets that had readily available site name lists. Additionally, Bhutan is a very unique place, where government buildings (*Dzongs*) and religious facilities were many times one in the same. The lists for each dataset were generated from the resources outlined in Table 3.6. Once the lists were generated for each dataset, they were geolocated using Google Maps™, and imported into ArcGIS™ using the geolocated latitude and longitude coordinates (Figure 3.12).

Table 3.6: Sources of ancillary data.

	India	Bhutan	Nepal
High Courts	India Mapped http://www.indiamapped.com/high-courts-in-india/	Google Maps (Dzongs)	Google Maps
Hospitals	Wikipedia http://en.wikipedia.org/wiki/List_of_hospitals_in_India	Google Maps	Wikipedia http://en.wikipedia.org/wiki/List_of_hospitals_in_Nepal
Monuments	India Mapped http://www.indiamapped.com/monuments-in-india/	Google Maps	United Nations Educational Scientific and Cultural Organization Sites (UNESCO)
Places of Worship	India Mapped http://www.indiamapped.com/churches-in-india/	Bhutan 2008 http://www.bhutan2008.bt/en/node/325; Google Maps	Wikipedia http://en.wikipedia.org/wiki/Category:Places_of_worship_in_Nepal
Public Airports	VDS Technologies	VDS Technologies (Paro, Bhutan airport not in dataset)	VDS Technologies; Local Knowledge

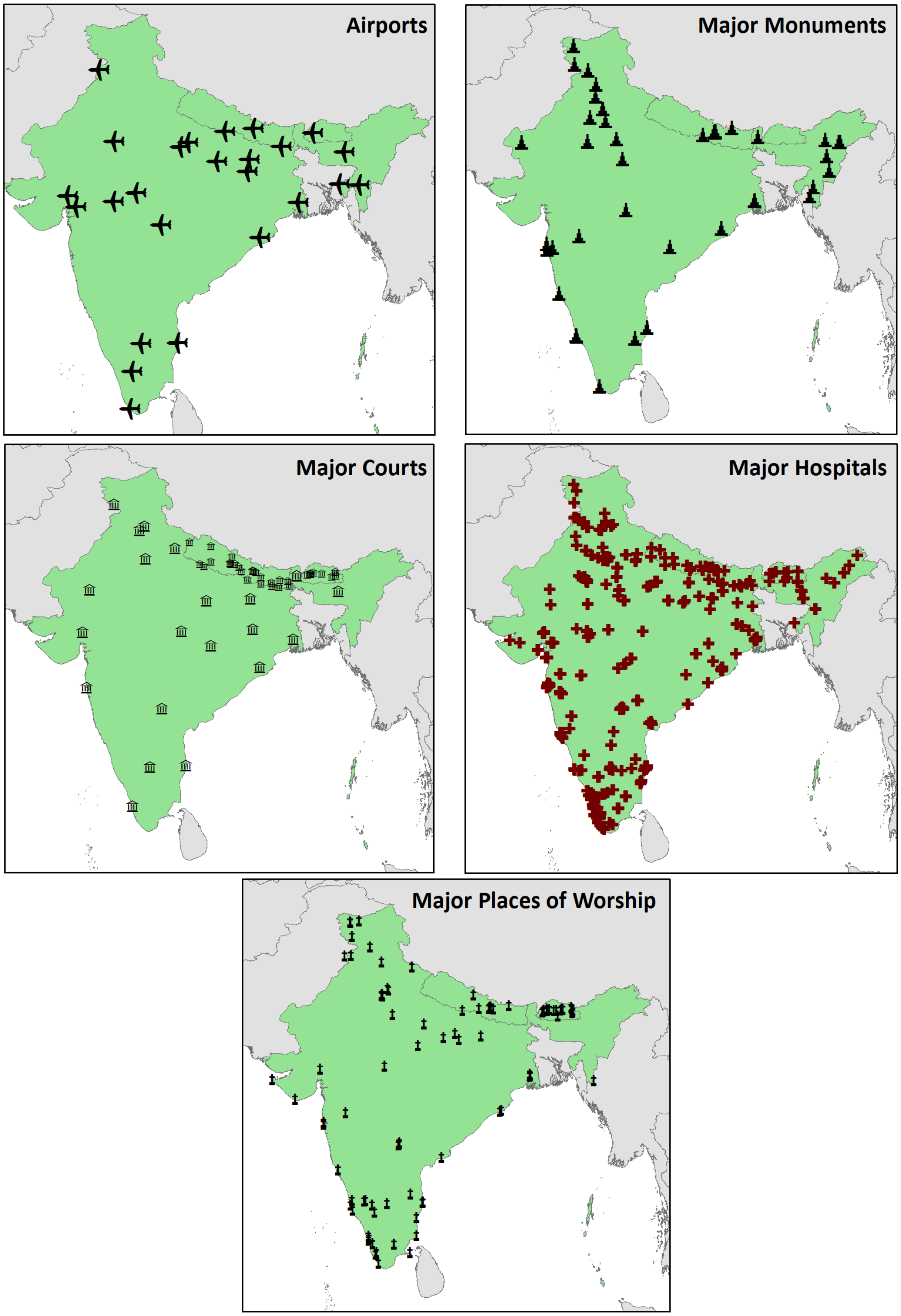


Figure 3.12: Ancillary data sets collected for this research.

3.2.7. Land Cover

2007 global land cover data, obtained from ORNL, was used with a spatial resolution of 1 km. Literature indicated that land cover can have an impact on the power needed for a particular location (Ministry of Power 2012). For example, in India, irrigated crop land requires a great deal of power, as does the urban areas, while forested area does not require as much power. Additionally the restoration of power to urban or irrigated cropland areas was of greater importance than to other land uses (Ministry of Power 2012). This data contained 25 land cover classes including Urban, Dry Cropland, Irrigated Crop Land, Cropland (Grassland), Cropland (Woodland), Grassland, Shrubland, Savanna, Deciduous Broadleaf, Deciduous Needle leaf, Evergreen Broadleaf, Evergreen Needle leaf, Mixed Forest, Water, Herbaceous Wetland, Barren, Herbaceous Tundra, Mixed Tundra, Bare Tundra, Snow or Ice, Partly Developed, and Coast Land.

3.3. Model Structure

Many considerations must be taken into account in the creation of a model. What type of data is appropriate to use? If raster, what spatial resolution should be used? What is the spatial unit of analysis? The answers to the questions are not just reflected in the research question, but also in the data availability. Multi-Criteria Decision Analysis (MCDA) specifically has important guidelines that researchers should follow. For example, if the model uses raster data, the data should all be at the same spatial resolution (aggregated, converted, or resampled), and it must cover the same spatial area. Additionally, before data can be utilized for MCDA, all data must be in the same units. If they are not in the same units, they should be reclassified into a comparable unit (such as low to high vulnerability) (Malczewski 2000).

3.3.1. Data Preparation

A variety of different types of data were utilized in this research, requiring varying degrees of preparation. This research used a variety of different data types: vector points, vector polylines, vector polygons, and raster data. This model was a raster-based model, so all vector data were converted into raster data. Additionally, the majority of the raster data that were being used for this analysis had a spatial resolution of 1 kilometer, so any raster data generated from vector data also had a spatial resolution of 1 kilometer. If raster data had a finer or coarser spatial resolution than 1 kilometer, the data were resampled to 1 kilometer. These transformations were performed with the knowledge that resampling finer resolution data to a coarser resolution would result in data loss and resampling coarser resolution data to finer resolution data would result in no additional information gain.

The scale of analysis for this research was the service area level. As stated earlier, the service areas were derived using the PoDiuM algorithm. All factors were aggregated to the service area level. The methods of preparing each individual dataset are describe in the following sections.

3.3.1.1. Betweenness Metric Data

Figure 3.13 describes the work flow for the betweenness data preparation. Degree and closeness were also determined using the same workflow process. The graph metrics were calculated using a modified version of the Urban Network Analysis

Toolbox for ArcGIS 10/10.1 out of the *City Form Lab* at Massachusetts Institute of Technology (Sevtsuk and Mekonnen 2012). This toolbox, created especially for analyzing urban terrain, calculates various graph metrics including reach, gravity, closeness, straightness, and, of most interest in this research, betweenness. The requirements to obtain the betweenness metric were a network dataset and ArcGIS's Network Analyst extension. The transmission grids for the three countries of interest were merged together, and a network dataset was generated. With the resulting network dataset, the Urban Network Analysis Toolbox calculates betweenness for the entire research area. Since the network data was processed as points, the points were then spatially joined to the Service Area file, and converted to a 1 km raster dataset, using the betweenness value as the raster value. The results were then classified into five classes, using natural breaks (Table 3.7). These classes were reclassified into values of 1 to 5 (1 being least critical and 5 being most critical). The reason this data was classified to was to make sure all of the factors were in the same units. For this dissertation research, the units were vulnerability units: 1 through 5. It also ensures that the data were in their natural classes, maximizing the differences between the classes. Maximizing the differences ensured the most agreement and similarities between members of the same class. Natural breaks were used to classify all of the factors utilized by this dissertation research.

3.3.1.2. Population

Figure 3.14 describes the data preparation process for the population data. 2010 LandScan Global™ was already a 1-km raster dataset, so it required no additional resampling. Zonal statistics were used to calculate the sum of the population present in the service area. The sum was used versus the density because the greatest interest to this dissertation research was how many total people the service area served, not the density. This raster was classified into five classes, using natural breaks to ensure like units between the factors (Table 3.8), and reclassified into values of 1 to 5 (1 being least vulnerable in terms of population (less people) and 5 being most vulnerable in terms of population (more people)) . The service areas tended to vary in size based on its capacity.

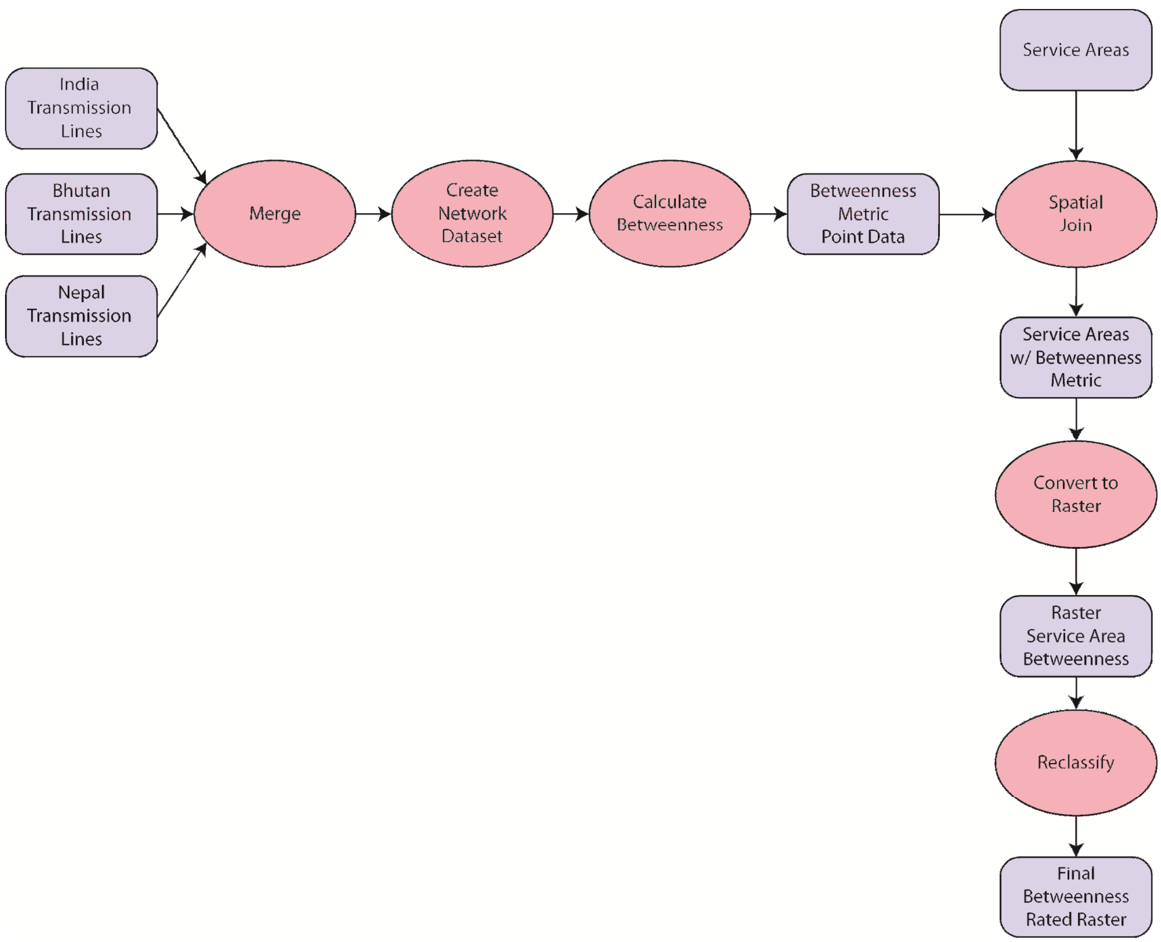


Figure 3.13: Data flow diagram for calculating the betweenness raster utilized in the analysis.

Table 3.7: Reclassified betweenness values.

Betweenness Range	Reclassified Value
0.0 – 10332.0	1
10332.1 – 31252.0	2
31252.1 – 63382.0	3
63382.1 – 119462.0	4
119462.1 – 235300.0	5

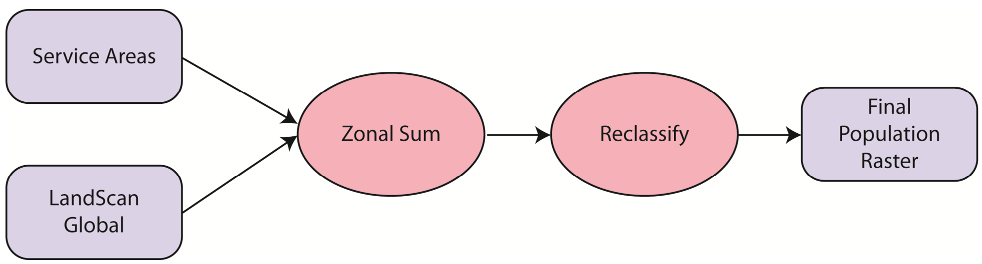


Figure 3.14: Data flow diagram for calculating the population raster utilized in the analysis.

Table 3.8: Reclassified population values.

Population Range	Reclassified Value
12,595 – 947,511	1
947,512 – 1,919,826	2
1,919,827 – 3,228,710	3
3,228,711 – 5,397,718	4
5,397,719 – 9,548,752	5

3.3.1.3. Additional Critical Infrastructure

Figure 3.15 represents the data preparation for the additional CI data. Additional CI was in vector point format. The individual CI shapefiles (Airports, Hospitals, Places of Worship, Monuments, and Courts) were merged into one shapefile, and then spatially joined to the Service Areas to obtain a count of how many CI locations were within the service area. The service areas were converted to a 1 km raster dataset based on the count. The count was classified into 5 classes using natural breaks, to ensure like units among all of the factors in this dissertation research, then reclassified into values from 1 to 5 (1 being least vulnerable and 5 being most vulnerable in terms of the number of structures located within the service area) (Table 3.9). A limitation of this approach is that it ignores in the individual criticality of each critical infrastructure type; however, this approach was taken because in this region many of these critical assets blend together (e.g. in Bhutan places of worship and courts are often the same).

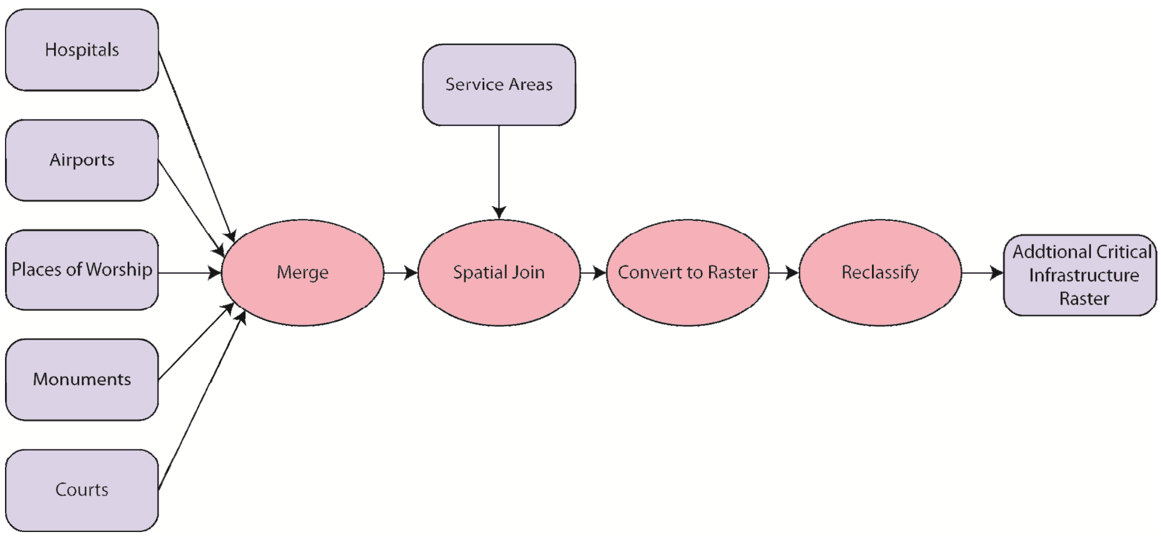


Figure 3.15 Data flow diagram for calculating the additional critical infrastructure raster utilized in the analysis.

Table 3.9: Critical infrastructure reclassified values.

CI Range	Reclassified Value
0 – 1	1
2 – 5	2
6 – 12	3
13 – 18	4
> 18	5

3.3.1.4. Land Cover

Figure 3.16 describes the data flow for preparing the 2007 land cover data. Land cover was also a 1 km raster dataset, so the land cover data did not need to be resampled in any capacity. The most vulnerable land covers (or land uses), as evidenced by the San Diego Blackout and the Indian Blackout, were urban areas and irrigated cropland that draw on electrical power. Part of the reason for India's 2012 back-out was irrigated crop land farmers were drawing more power than normal (Philpott and Jones 2012). Additionally, both blackouts had tremendous effects on major urban centers (San Diego and New Delhi, respectively). As such, those land cover categories were weighted as more vulnerable and more important to the grid. Additionally, there is an abundance of electrical power theft in India. The majority of the theft is in irrigated crop land areas and urban centers, which would put more stress on the electrical grid in these areas (Golden and Min 2012). The relative rankings used are summarized in Table 3.10. To obtain an overall ranking for the service area, zonal statistics were performed to obtain the zonal sum of the vulnerability rankings, which was then divided by the total number of cells in the service area for the average vulnerability ranking of the cells in that service area. The resulting rank was used for the service area as input into the WLC.

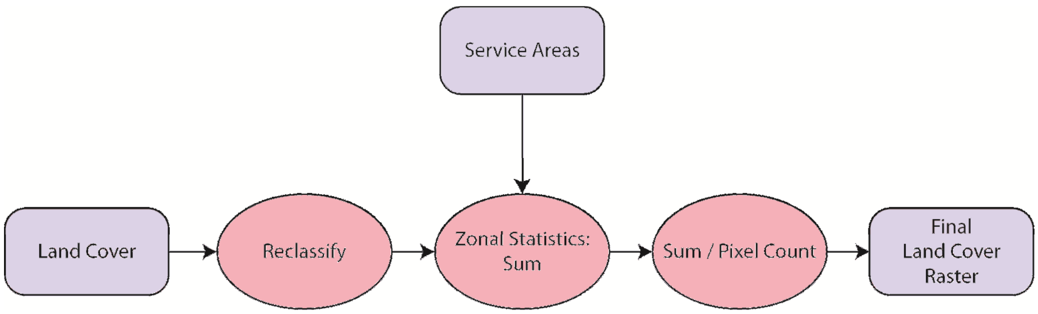


Figure 3.16: Data flow diagram for calculating the land cover raster utilized in the analysis.

Table 3.10: Land cover reclassified values.

Land Cover Class	Reclassified Value
All Others (Dry Cropland, Grassland, Shrubland, Savanna, Forests – all varieties, Water, Wetland, Barren, etc)	1
NA	2
Partially Developed	3
NA	4
Urban, Irrigated Cropland	5

3.3.1.5. Natural Hazards Risk

Figure 3.17 details the data preparation required for the natural disaster data. The data from UNEP came in a variety of different units, so to make sure the research was comparing like data forms, the data was transformed into similar units: average annual frequency. Cyclone, for instance, were in average number per 100 years times

100, while landslides were an index of the probability of a landslide per year. To transform the data into the same unit (frequency per year), simple map algebra functions were performed (Table 3.11).

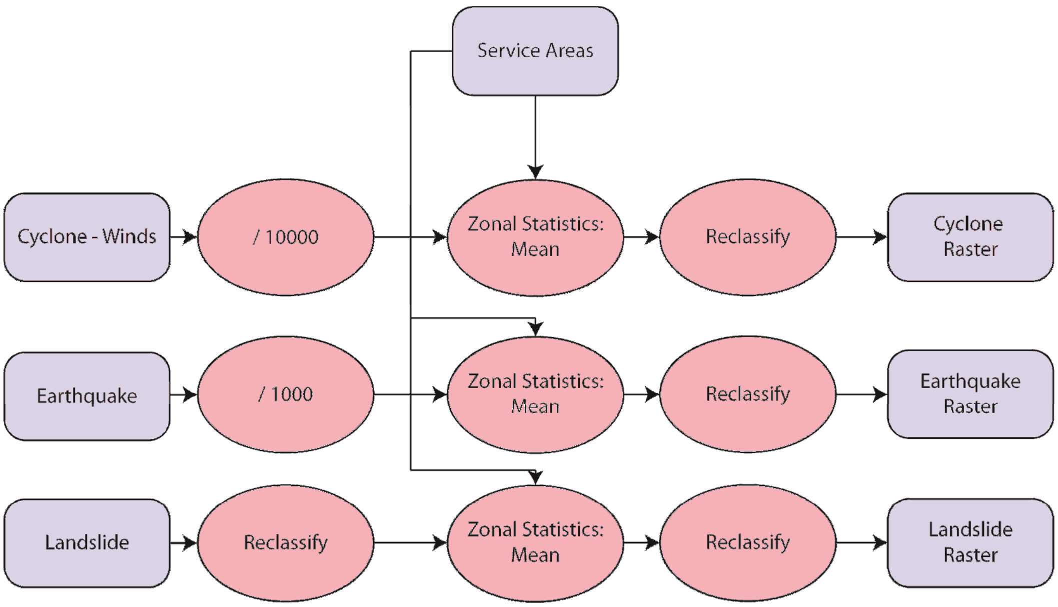


Figure 3.17: Data flow diagram for calculating the natural hazards raster utilized in the analysis.

Table 3.11: Algebraic functions for determining like values for natural hazard datasets.

Data Set	Units	Map Algebra Function
Cyclone – Winds	Average Number per 100 years Times 100	Value/10000
Earthquakes	Average Number per 1000 years	Value/1000
Landslides	Index of Annual Probability and Size	Reclassify by Potential Landslides/Year

Even though the hazard frequencies were in comparable units, the data were still not in the same unit as the other criteria utilized in the analysis. Each individual hazard frequency per year was averaged at the service area level, then reclassified into individual raster datasets indicating the relative frequency of each hazard (1 indicating least frequent and 5 indicating most frequent). Zonal means were used instead of sums, because a service area with a larger area would have been over represented as having a high frequency of events, when in reality the frequency was relatively low in comparison to other service areas. Additionally, individual hazards were used instead of an aggregated sum of all hazard frequencies to provide additional flexibility in the model structure. For example, there may be seasonal variances in occurrence that may exclude certain hazards from analysis at a given time (e.g. there is a season for cyclones, and when the model is run for times when cyclones are not prevalent, that particular dataset can be excluded from the analysis or weighted less). Tables 3.11 through 3.13 indicate the frequency ranges and their subsequent reclassified values for use in the final AHP analysis (Table 3.12, Table 3.13, and Table 3.14).

Table 3.12: Cyclone reclassified values.

Cyclone	Reclassified Value
0	1
0.0001 – 0.00033	2
0.00034 – 0.00061	3
0. 00062 – 0.0010	4
0.0011 – 0.002	5

Table 3.13: Landslide reclassified values.

Landslides	Reclassified Value
0	1
0.0001 – 0.00077	2
0. 00078 – 0.0021	3
0. 0022 – 0.0039	4
0.0040 – 0.0085	5

Table 3.14: Earthquake reclassified values.

Earthquakes	Reclassified Value
0	1
0.0001 – 0.022479	2
0.02248 - 0.03693	3
0.03694 – 0.06529	4
0.0653 – 0.13648	5

3.3.1.6. Climate Extremes

Reports explaining recent blackouts have often cited climatic extremes as being part of the cause of the disturbance (Ministry of Power 2012). Unfortunately, few attempts have been made to incorporate climate extremes with critical infrastructure models. Those models that have incorporated temperature extremes have done so by estimating the increase in demand per temperature degree of increase; however, this has only been tested in the mid-latitudes of the United States, and may not be advisable for this study area (Young 2009). However, this temperature increase model does indicate that with increased temperature, there is increased stress on the grid. This dissertation research used a basic approach to including temperature extremes. The approach to including climate extremes was to compare temperatures on any given day to the temperature normal for that month. The deviation from the average monthly temperature normal was then ranked from smallest deviation to largest deviation, thus showing increasing stress on the grid.

Figure 3.18 details the data preparation for temperature extremes. Monthly average temperature maximums (1950-2000) were available from WorldClim with a spatial resolution of 1 km. Since this dissertation research compared the model to the July 2012 Indian blackout, it was best to replicate those conditions. As such, the temperature normal for July was used and resampled to a resolution of 1 km to match the rest of the raster data.

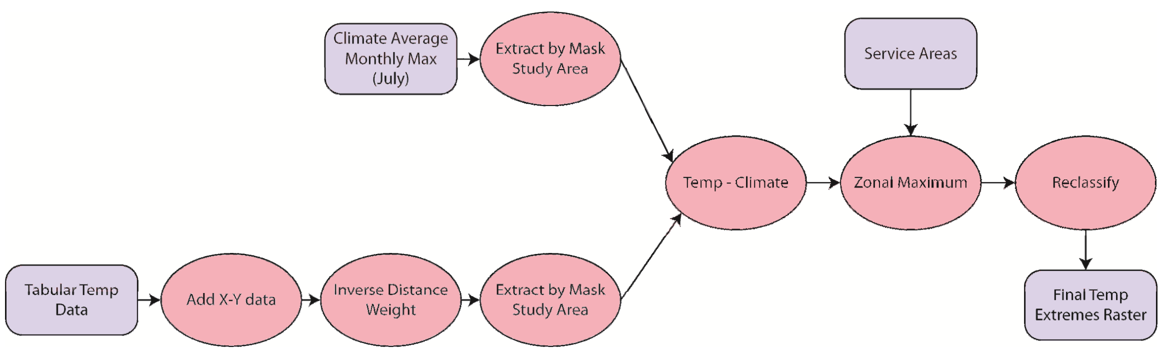


Figure 3.18: Data flow diagram for calculating the climate extremes raster utilized in the analysis.

The National Climatic Data Center (NCDC) maintains a database of historical global temperature data. This data is maintained in a tabular format for 70 stations in India. The dates of the Indian blackout were July 30 and 31, 2012, so temperature data for those days were obtained for the available stations, and transposed into a format readable in a GIS. This data was imported as points, and a 1-km raster surface was generated using the Inverse Distance Weight (IDW) algorithm.

The individual IDWs were subtracted from the July temperature normal data to determine which areas were above or below average for that particular day. The zonal maximum for each individual service area was then calculated to determine the maximum deviation from the normal temperature for a given service area on that day. The service areas were then converted into a 1-km raster dataset based on the maximum deviation from the normal temperature, and classified into five classes using natural breaks. Natural breaks were utilized to remain consistent with the other factors, and to also convert the climate data into the same vulnerability units for comparison to the other factors in this dissertation analysis. The data was then reclassified into values ranging between 1 and 5 (1 indicating the least deviation from normal (least critical to the grid) and 5 indicating the greatest deviation from normal (most critical to the grid)). Table 3.15 identifies the reclassified values for both days.

Table 3.15: Climate extremes reclassified values.

Temperate Extreme Range	Reclassified Value
-5.21048 – 0.000	1
0.001 – 6.449	2
6.450 – 14.496	3
14.497 – 23.8568	4
23.8569 – 36.336	5

3.3.2. Weighted Linear Combination (WLC)

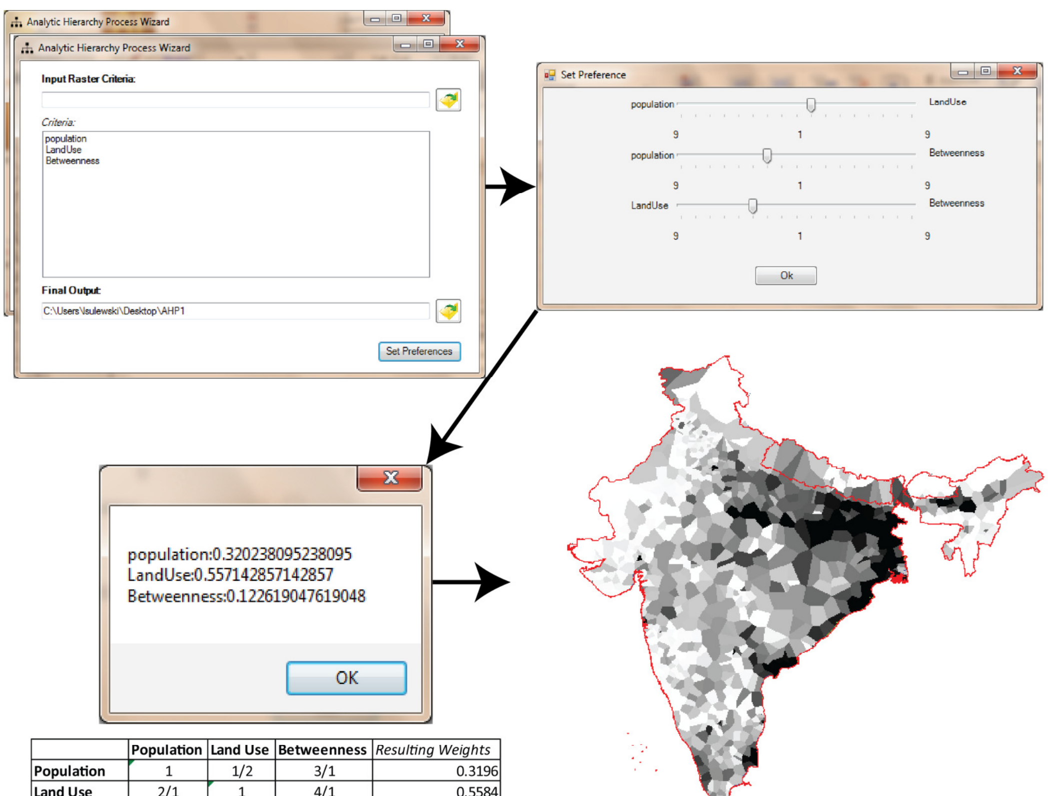
The base of the model was an AHP approached combined with a WLC. The AHP tool was packaged as an ArcGIS add-in for the ArcGIS version 10, as all legacy AHP scripts for ArcGIS were not compatible with the ArcGIS version 10 and creating a new code to the specifications of this research was more time efficient than upgrading existing code to the current version of ArcGIS. It is important to note that the implementation of AHP in this dissertation research did not use a hierarchical pairing of factors, but a flat implementation with pairwise comparisons.

The AHP tool, as stated, was an ArcGIS add-in developed in a VB.net framework that calculates weights based on AHP. The reclassified raster datasets from the data preparation stage as input into the “Analytic Hierarchy Process Wizard” were selected. These raster datasets then populated the “Criteria” box (Figure 3.20). The “Final Output” was where the output of the resulting WLC was stored. After the criteria were input, a dialogue box was created that dynamically generated track bars for the appropriate number of pairwise comparisons given the number of factors provided. These track bars allowed the user to perform the pairwise comparisons of the various factors, without filling in the comparison matrix. The sliders were used on the track bar to indicate the relative importance of one factor versus another.

After the pairwise comparisons for all combinations of the factors were performed, a dialogue box was generated providing the resulting weights for each factor and processed the rasters by multiplying the cell values by the factor weights calculated by the AHP algorithm. Figure 3.19 provides a small example of how the AHP tool works.

The minor differences in the resulting weights were due to rounding differences. For this dissertation research, experts in the region performed the pairwise comparison.

The primary purpose of the program was to calculate the weights; however, it had the added advantage of preforming the WLC. Figure 3.20 identifies the processes that occurred after the weights were calculated. After the weights were calculated, they were multiplied by their corresponding factor raster. These weighted surfaces were then summed together to provide a final vulnerability surface. After the final vulnerability surface was calculated, the raster was reclassified into five classes to indicate relative vulnerability classes: very low, low, medium, high, and very high vulnerabilities.



	Population	Land Use	Betweenness	*Resulting Weights*
Population	1	1/2	3/1	0.3196
Land Use	2/1	1	4/1	0.5584
Betweenness	1/3	1/4	1	0.122

Figure 3.19: AHP tool graphical user interface and outputs.

There were several advantages to producing an AHP tool for ArcGIS in this fashion. First, the track bars made the AHP tool much easier to use than simply having the user fill in the pairwise comparison matrix. The matrix can be difficult to understand, and the user might be unsure how to conduct the pairwise comparison. Providing track bars for the user made placing higher rankings on more important factors easier. Additionally, ArcGIS add-ins are easier to transport from one user to another. It is the first add-in of its kind for ArcGIS 10, as most other AHP tools are for legacy versions of ArcGIS that have either not been updated or have ArcGIS 10 versions in production.

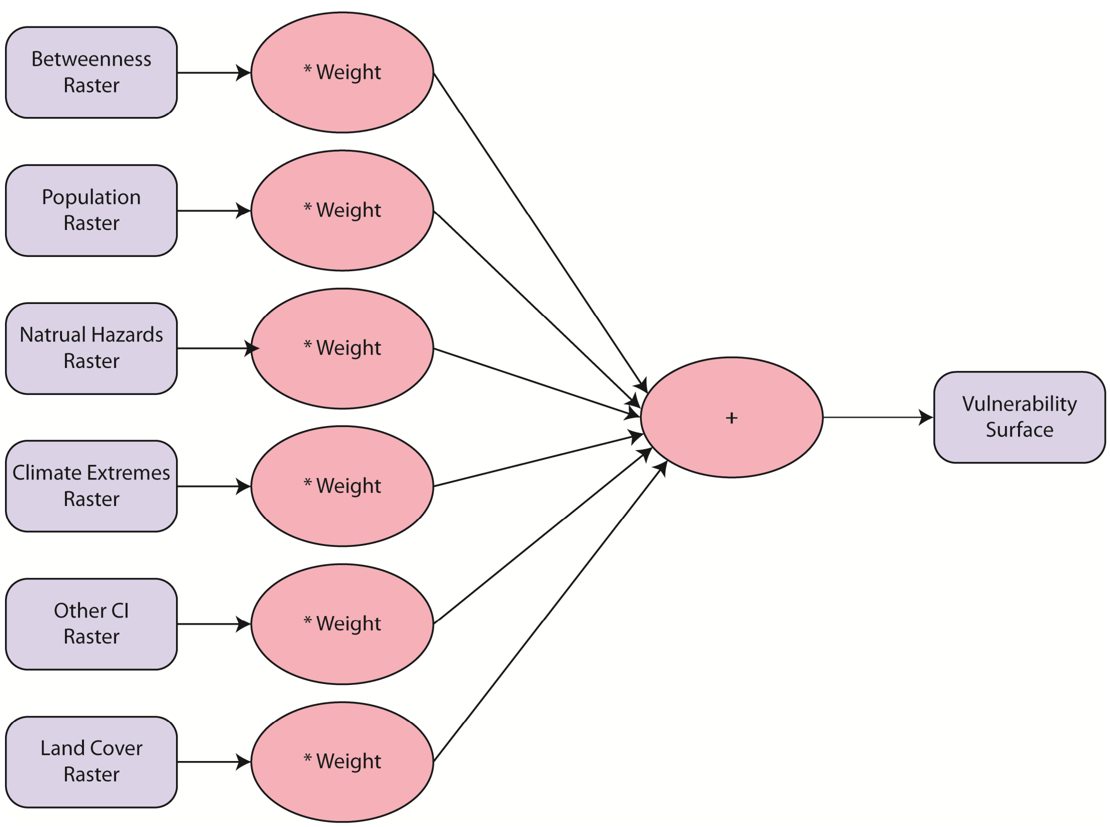


Figure 3.20: Weighted Linear Combination process performed in the AHP tool.

3.4. Model Evaluation

3.4.1. Accuracy Assessment/Validation

The model was performed using three different approaches, each building on the other. The first approach was based solely on the graph metrics, where the results were classified into five vulnerability classes using natural breaks. The graph metrics of degree, betweenness, and closeness were evaluated to determine the best performing graph metric. The second approach only used graph metrics and non-climatic data (betweenness, population, land cover, and other critical infrastructure data), as opposed to the first that only used graph metrics. The resulting vulnerability surfaces were also classified into five vulnerability classes using natural breaks. Natural breaks were used to classify the data based on its ability to derive natural classes, which can be reclassified into levels of vulnerability. This was also the same method used to classify the factors in this dissertation research. The final approach included the best performing combinations of the variables utilized in the previous approaches (betweenness, population, land cover, and other critical infrastructure) and combined them with natural disaster frequency and climatic extremes. As with the other two approaches, the vulnerability surface was classified into five vulnerability classes.

These three approaches were first compared and validated based on reference data from India's 2012 blackout by performing an accuracy assessment of the results. This accuracy assessment consisted of comparing the substation rankings for each of the variable combinations for the substations that were involved in the July 2012 Indian Blackout. The accuracy assessment metric was defined as the number of substations identified as having a high or very high vulnerability that were involved in the blackout compared to the total number of substations involved in the blackout (For example, if 10 of 30 substations involved in the blackout were identified by a given model, the accuracy assessment metric would yield a value of 33.3%.) Of particular interest were the following metrics and questions:

1) What service areas were consistently ranked as vulnerable by all three scenarios? What were the possible explanations?
2) What service areas were not consistently ranked as vulnerable by all three approaches?
3) Where did the differences lie in the most highly vulnerable substations?
4) What did the resulting vulnerability rankings indicate about the addition of geographic data to graph metrics for analyzing substation and service area vulnerability?
5) How would the policy recommendations from each approach differ, if any?

The results were given within the context of the 2012 Indian blackout. Would any of these models be able to predict the vulnerability of the key substations in the 2012 blackout?

3.4.2. Statistical Analyses

Two statistical analyses were performed to further determine the validity of the results. The first statistical analysis that was performed was a principal component analysis. One of the purposes of the second and third approaches was to determine which factors were most pertinent to node vulnerability and those factors that were not. A factor analysis helped to determine if there were redundancies in the data and where those redundancies were. Additionally, factor analysis determined whether the best performing model had or did not have data redundancies. A second statistical analysis that was performed was an Analysis of Variance (ANOVA) test to assess the statistical differences between the best performing scenarios and determine if there were any statistically significant differences between the scenarios.

3.4.3. Sensitivity Analyses

Sensitivity analyses on MCDM analyses typically represent variations in the weights, geography, or scale. For this dissertation research, the three major sensitivity analyses were conducted: based on the combination of factors, based on the factor weights, and based on the spatial scope of analysis.

The first sensitivity analysis was a sensitivity analysis based on the combination of factors. For example, how did the combination of land use and betweenness compare to the combination of land use, betweenness, and population? Did one of these combinations correctly identify vulnerable substations from the July 2012 black out versus the other? Some combinations of factors out-performed others, and thus it was important to find the combinations of factors that best identified the vulnerable substations.

A sensitivity analysis was also conducted on the factor weights. In this sensitivity analysis, the weights of the factors in the pairwise combinations were varied to see the effect the combination's performance.

A sensitivity analysis was also conducted for the spatial scope of analysis. This dissertation research was conducted initially for the entire study area (Bhutan, India, and Nepal). The spatial scope sensitivity analysis took the analysis that was first conducted for the entire study area and conducted it only for the Northern Indian Grid (Figure 3.3). The factors were reclassified using natural breaks based on the Northern Grid region only. These results were compared to the outcome of the model when analyzed for the entire study area to see if there were sensitivities in the analysis results based on the scale of analysis.

CHAPTER IV

RESULTS

4.1. Introduction

The methods for determining substation vulnerability described in Chapter 3 were evaluated utilizing the Indian electrical grid context, specifically the Indian blackout that occurred on July 30 and 31, 2012. The framework was evaluated for its value in determining substation vulnerability and applicability to other research areas. Vulnerability, for the purposes of this research, is defined as "…physical feature or operational attribute that renders an entity open to exploitation or susceptible to a given hazard" (DHS 2008, pg. 34). Essentially, this definition refers to a critical infrastructure network's ability to maintain a given critical function. In the case of this dissertation research, the critical function was maintaining power to the largest number of customers. This dissertation method did not indicate risk, as consequence was not evaluated. The results were used not only to understand the vulnerabilities of individual substations but to understand the system as a whole.

4.2. Indian 2012 Blackout – Context

On July 30 and July 31, 2012, two major blackouts impacted the Indian electrical grid leaving 10% of the world's population without power. Both blackouts had major consequences for the northern grid of India, resulting from a series of transmission line losses throughout the regional grids (Figure 3.3). A report of the Enquiry Committee on the Grid Disturbance cited weak inter-regional transmission corridors and lack of situational awareness as the major reasons for the blackouts (Ministry of Power 2012).

The July 30, 2012, event consisted of the loss of substations and transmission lines over the course of 20 different events, taking approximately 20 seconds, and affecting 29 substations throughout northern India. The first event was the loss of the Bina-Gwalior 400 kV Transmission line, which is the main connection between the western region and the northern region. Figure 4.1 illustrates the sequence of events and lost assets during the July 30, 2012 event. Figure 4.2 illustrates the lost substations. Eighty of power was restored after 15 hours of work on July 30, 2012; however, another, larger event occurred the next day: July 31, 2012.

Sixteen substations were impacted on both days, with 13 impacted only on July 30, 2012, and 29 impacted only on July 31, 2012. The second day's events also extended further south and north. The July 31st event greatly impacted the New Delhi area, with five additional substations being impacted in that area (Figure 4.3). The second event occurred over the course of 3 minutes and 5 seconds. A little less than double the number of substations was affected on this second day, with a total of 45 substations impacted over the course of 30 distinct outages (Figure 4.2). The initiating event was the loss of Bina-Gwalior 400 kV Transmission line, which is the main connection between the western region and the northern region.

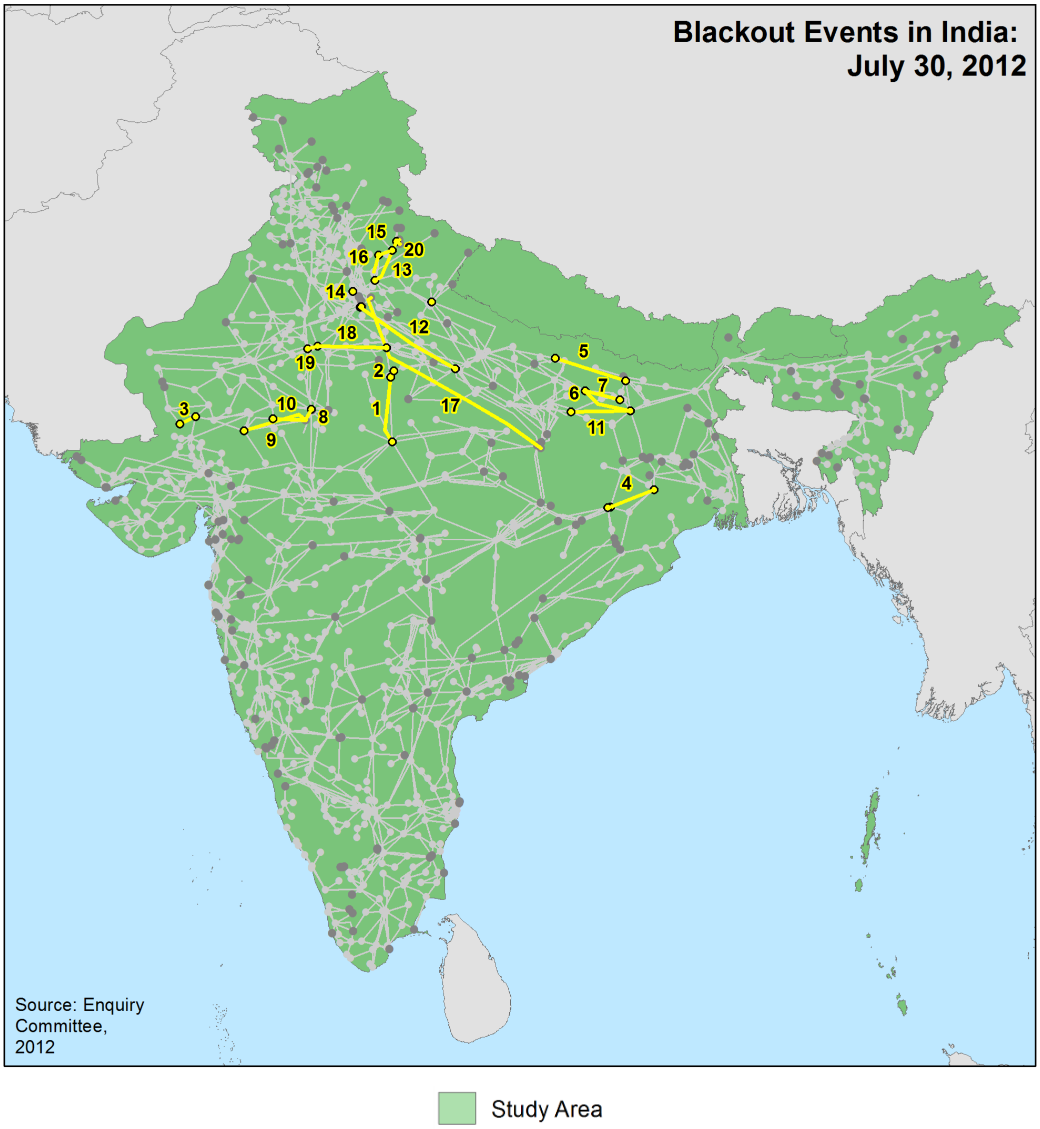


Figure 4.1: July 30, 2012, sequence of events shown by the transmission line that was lost and the order (by number).

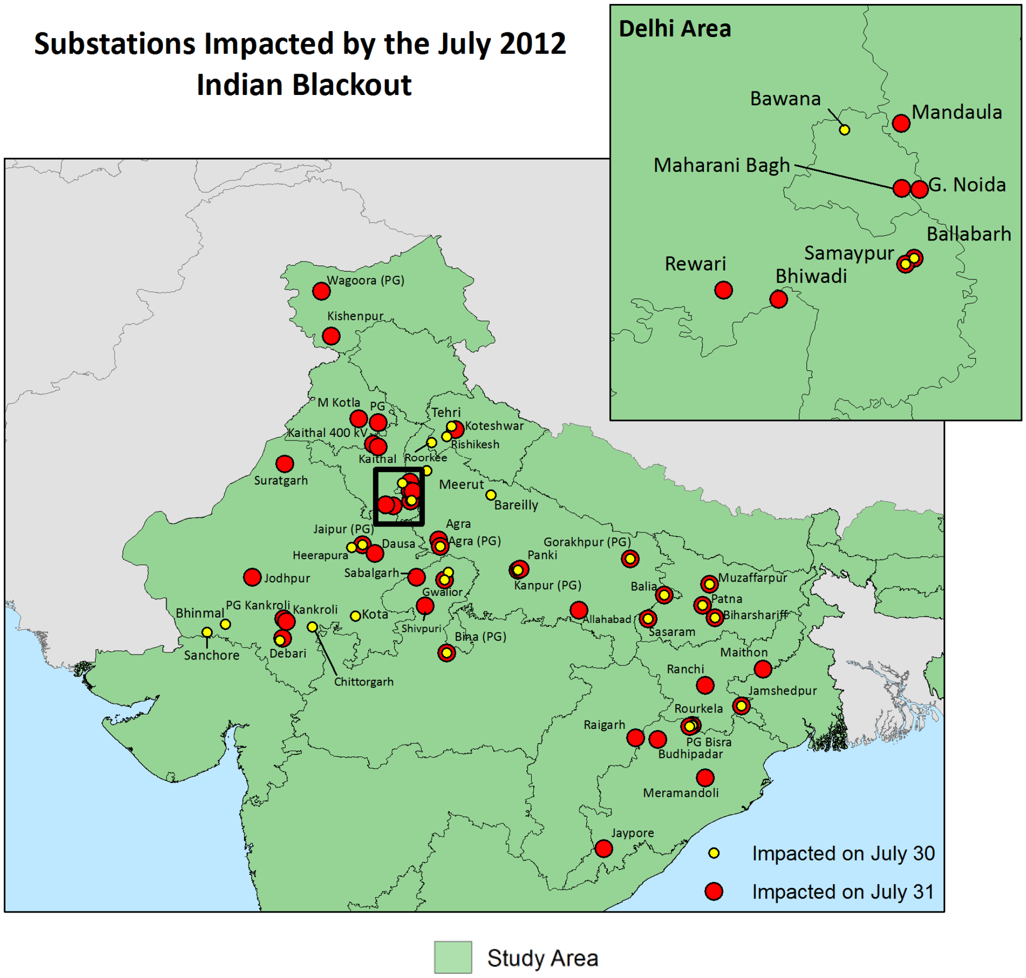


Figure 4.2: Substations impacted by the blackouts.

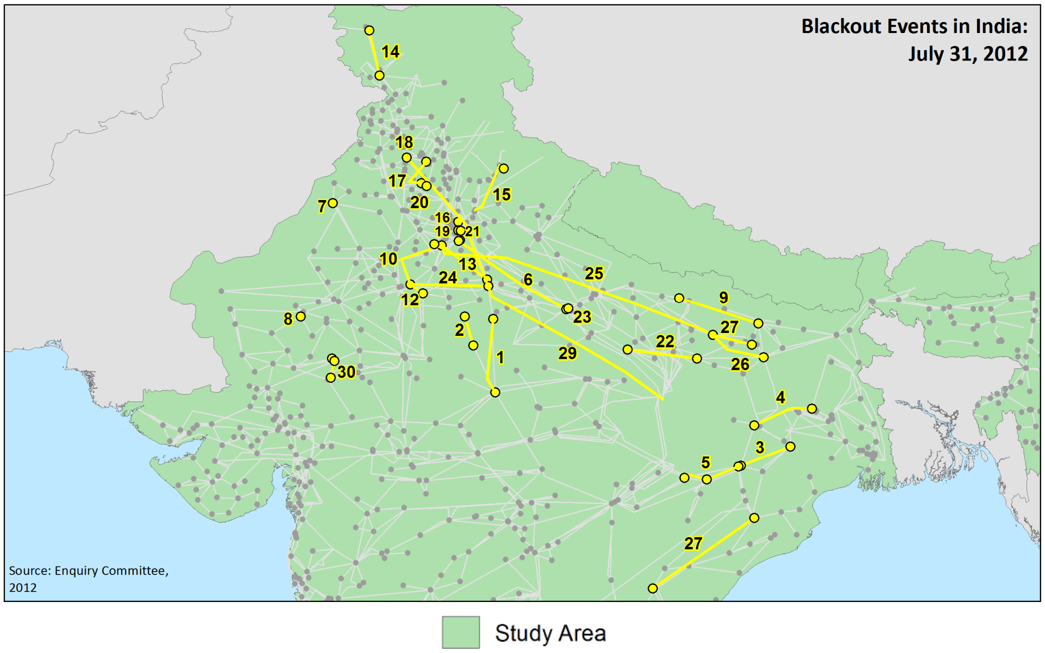


Figure 4.3: Sequence of events for the July 31, 2012 blackout.

Could this blackout have been predicted and perhaps avoided with knowledge of the vulnerabilities existing in the grid? The detailed report of the events of both days helped determine if this blackout could have been predicted with graph metrics alone, a more or detailed, place-based framework. The two days were analyzed as two separate events.

4.3. Graph Metric Results

Literature describing the use of betweenness to determine vulnerability in electrical grids have cited that betweenness is the most appropriate measurement, as it is the only measure that measures flow in the network (See Chapter 2 for a more detailed explanation of betweenness) (Rocco et al. 2011, Desmar et al. 2008). To further demonstrate the importance of betweenness using real-world data, three centrality metrics are tested: betweenness, degree, and closeness. Table 4.1 details the usefulness of each graph metric in identifying substations involved in the 2012 Indian blackout as a high or very high vulnerability. Numbers 1 through 5 indicated relative vulnerability, 1 indicating very low vulnerability and 5 indicating very high vulnerability. Closeness was not included in the table, because it did not identify any of the substations as having a high or very high vulnerability. Betweenness and degree had similar results; however, betweenness identified more substations as having a high or very high vulnerability (27.6% versus 20.6% for degree for July 30th substations). There was between a 36% (for the July 31st substations) and 57% (for the July 30th substations) agreement between betweenness and degree, meaning 36% of the substations identified as having a high or very high vulnerability were identified by both the betweenness metric and the degree metric.

In this section and the sections to follow, a variety tables were generated to help visualize the vulnerability of the nodes involved in the Indian Blackout. These charts are color coded based on the color scheme in Figure 4.4.

Figure 4.4: Color Scheme used.

Table 4.1: Comparison of centrality metrics for determining spatial graph Vulnerabilities, based on betweenness.

July 30, 2012		
	Vulnerability Rankings	
Substation	Betweenness	Degree
Agra (PG)	2	3
Balia	1	3
Ballabarh	1	3
Bareilly	1	3
Bawana	2	3
Bhinmal	3	3
Biharshariff	4	5
Bina (PG)	3	3
Chittorgarh	1	3
Gorakhpur (PG)	4	3
Gwalior (PG)	5	3
Heerapura	4	5
Jaipur (PG)	3	5
Jamshedpur	4	3
Kanpur (PG)	4	5
Kota	5	5
Malanpur	2	2
Meerut	3	5
Muzaggarpur	3	3
Patna	1	3
PG Bisra	3	3
Rishikesh	1	3
Roorkee	2	3
Rourkela	4	3
Samaypur	2	5
Sanchore	1	2
Sasaram	3	3
Tehri	1	3
Udaipur	1	3
%(Very)High	27.60%	20.60%

July 31, 2012		
	Vulneraiblity Rankings	
Substation	Betweenness	Degree
Agra	5	5
Agra (PG)	2	3
Allahabad	4	5
Balia	1	3
Ballabarh (BBMB)	1	3
Bhiwadi	5	5
Biharshariff	4	5
Bina (PG)	3	3
Budhipadar	2	3
Dausa	3	3
Debari	1	1
Greater Noida	1	2
Gorakhpur (PG)	4	3
Gwalior (PG)	5	3
Jaipur (PG)	3	5
Jamshedpur	4	3
Jaypore	4	3
Jodhpur	3	3
Kaithal	2	4
Kaithal (400 kV)	1	3
Kankroli	1	2
Kanpur (PG)	4	5
Kishenpur	3	5
Koteshwar	1	3
Maler Kotla	2	3
Maharani Bagh	2	2
Maithon	2	3
Mandaula	2	3
Meramandoli	4	3
Muzaffarpur	3	3
Panki	3	4
Patna	1	3
PG	2	5
PG Bisra	3	3
PG Kankroli	2	3
Raigarh	2	3
Ranchi	3	3
Rewari	3	2
Rourkela	4	3
Sabalgarh	1	1
Samaypur	2	5
Sasaram	3	3
Shivpuri	1	2
Suratgarh	1	3
Wagoora (PG)	1	3
Percent High/Very High	24.40%	24.40%

In keeping with previous research for determining vulnerabilities in graph networks, betweenness was calculated for this research and compared to the events for the 2012 blackout (Rocco et al. 2011, Desmar et al. 2007). The betweenness results for the entire study area are summarized in Table 4.2. Ideally, with critical infrastructure, a country does not want a larger number of service areas (substations) to have a very high vulnerability. For this blackout, there were only about 2% of the substations were considered to be very high vulnerability nodes. The high and very high vulnerability rankings accounted for less than 10% of the substations in the study area. Hines et al. (2010), however, caution that only assessing a network's vulnerability based on graph metrics can be misleading. Was betweenness able to identify the majority of the substation service areas that were affected by the 2012 Indian Blackout?

Table 4.2: Summary of Betweenness Metric Results.

Vulnerability Ranking	Number of Service Areas Identified	Percent of Total Service Areas
Very Low	652	68.8%
Low	149	15.7%
Medium	79	8.3%
High	50	5.3%
Very High	18	1.9%

4.3.1. July 30th, 2012

Overall, betweenness was a poor indicator of substations vulnerable during the July 30th event. Overall, betweenness was only able to identify two out of 29 of the impacted substations as high vulnerability substations, and only 8 (27.7%) of the 29 as high or very high vulnerability. The majority of the impacted substations were considered medium, low, or very low vulnerability (76.1%) with regards to betweenness (Figure 4.5). However, it is important to note that one of the initiating events of the blackout, the loss of the transmission line between Bina and Gwalior, would have been identified using the betweenness metric, as the Gwalior substation has a very high vulnerability ranking.

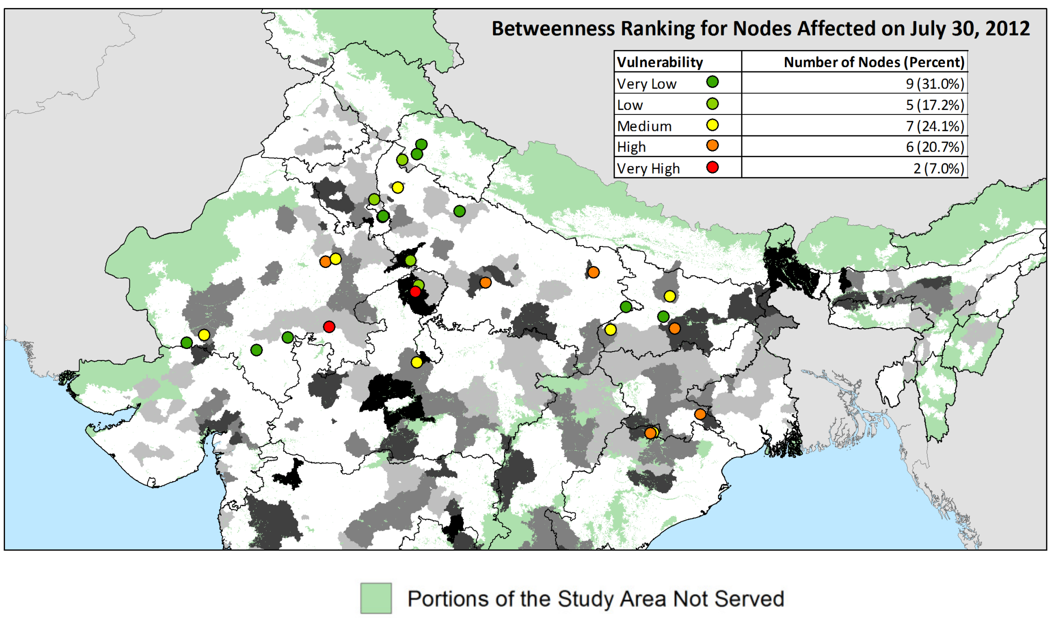


Figure 4.5: Vulnerability rankings for substations affected by the July 30th, 2012 blackout.

4.3.2. July 31st, 2012

Similar to the events on July 30th, overall, the betweenness measure was a poor indicator of substations vulnerable during the July 31st event. Only three (6.5%) of the nodes were ranked as having a very high vulnerability, and only 12 (26.1%) nodes were captured in the high and very high vulnerability rankings (Figure 4.6). If policy were enacted solely on the basis of the betweenness measure, only 12 of the nodes affected on this day would have been protected. However, it is important to note that one of the initiating events of the blackout, the loss of the transmission line between Bina and Gwalior, would have been identified using the betweenness metric, as the Gwalior substation has a very high ranking.

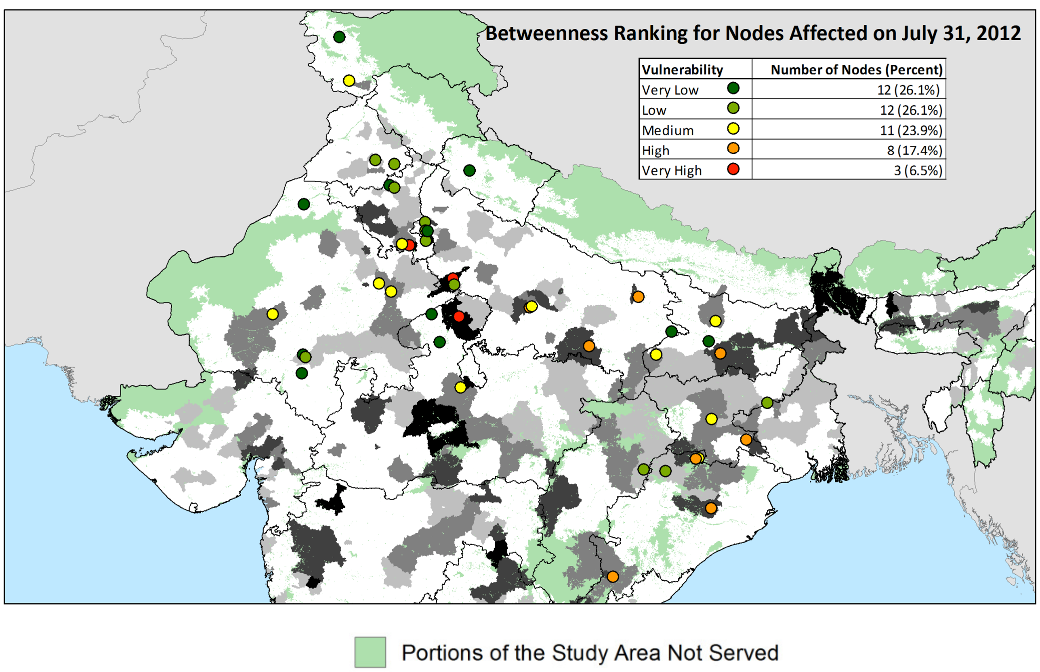


Figure 4.6: Vulnerability rankings for substations affected by the July 31st, 2012 blackout.

For both days of the blackout event, betweenness was not a sufficient metric for identifying vulnerable substations. For each day, it accounted for less than half of the total affected substations.

4.4. Inclusion of Non-Climate Related Variable Results

The first approach indicated the ineffectiveness of only using betweenness to measure vulnerabilities of substations within the Indian grid. While there are benefits of using betweenness, such as its ability to identify the vulnerabilities in the initiating substations, it missed around 80% of the other substations affected. Will the inclusion of characteristics of the substations help identify additional vulnerabilities?

4.4.1. Descriptions of the Datasets

4.4.1.1. Land Use

The land use dataset, similar to the betweenness dataset, was ranked from 1 to 5 indicating varying ranks of vulnerability. The land use data included such categories as urban, water, and partly developed. For a more robust definition of land use, please refer to Chapter 3. The results for the vulnerability rankings for land use are summarized in Table 4.3. The percentage of total service areas in each category was fairly comparable to that of the betweenness metric, except there was approximately 3% more service areas found to have a very high vulnerability ranking; however, that increase was coupled with an approximate 5% increase in the percentage of service areas ranked as having a very low vulnerability ranking.

Table 4.3: Summary of land use results.

Vulnerability Ranking	Number of Service Areas Identified	Percent of Total Service Areas
Very Low	701	73.9%
Low	124	13.1%
Medium	49	5.2%
High	34	3.6%
Very High	40	4.2%

In addition to the differences in the total number of service areas represented by each category, there were also differences in the geographic distribution of the vulnerability (Figure 4.7). While over half (55.8%) of the substations had no change in their ranking, it was interesting to note where there were differences in the vulnerability ranking (44.2%).

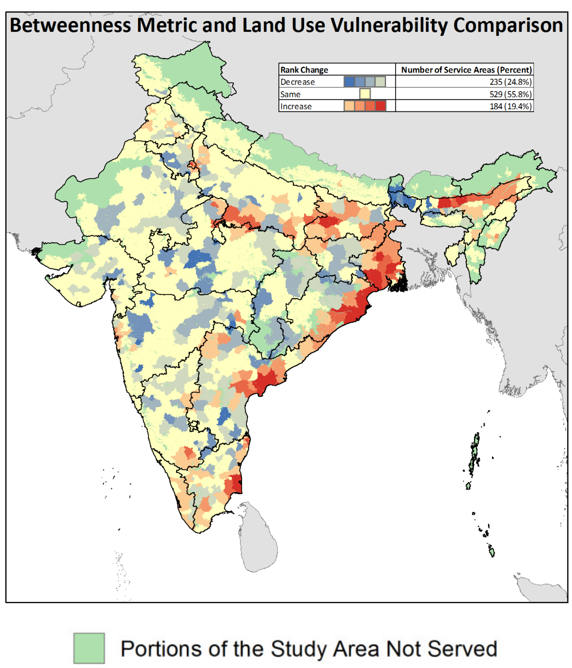


Figure 4.7: Geographic difference in vulnerability distribution for betweenness versus land use.

4.4.1.2. Population

Population totals, as acquired from LandScan Global, were also ranked from 1 to 5 indicating ranks of vulnerability with 1 being very low vulnerability and 5 being very high (see Chapter 3 for more information). The results for the vulnerability rankings for population are summarized in Table 4.4. The population rankings had a different spatial distribution, with the majority of the service areas with a ranking of very low through medium vulnerability (97.2%), and only 2.8% of the service areas with a ranking of high or very high vulnerability. This could mean that the service areas were well divided amongst the served population, with few service areas serving large populations.

Table 4.4: Summary of population results.

Vulnerability Ranking	Number of Service Areas Identified	Percent of Total Service Areas
Very Low	565	59.6%
Low	250	26.4%
Medium	106	11.2%
High	25	2.6%
Very High	2	0.2%

The geographic distribution of the vulnerability rankings using population totals definitely revealed new information not contained in the betweenness metric or the land use vulnerability rankings (Figure 4.8). Between population and land use rankings, 56% of the service areas remained the same, and 44% changed. When comparing the betweenness metric and the population rankings, 51% of the service areas experienced a change in rankings. This indicates disparities in what each variable identifies as being vulnerable.

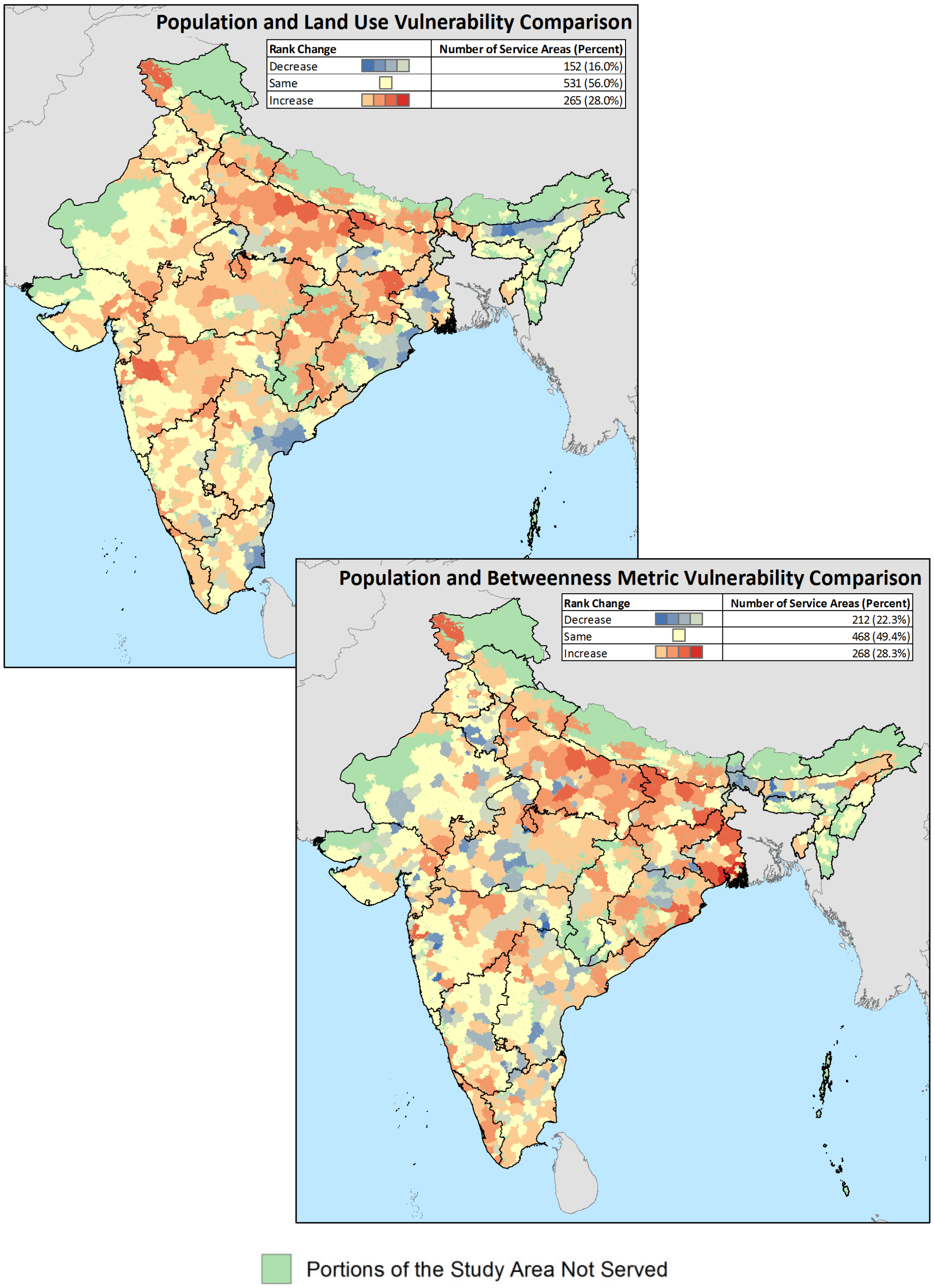


Figure 4.8: Geographic difference in vulnerability distribution for population versus the betweenness metric and land use.

4.4.1.3. Additional Critical Assets

The additional critical assets by service area dataset were ranked in the same fashion as the betweenness metric, land use, and population, with a service area containing few additional critical assets having a vulnerability ranking of 1. The results for the vulnerability rankings for population are summarized in Table 4.5. Similar to population, very few substations were ranked as high or very highly vulnerable (less than 1%) in terms of the occurrence of other critical assets beyond the electrical infrastructure.

Table 4.5: Summary of additional critical assets results.

Vulnerability Ranking	Number of Service Areas Identified	Percent of Total Service Areas
Very Low	843	89.0%
Low	71	7.5%
Medium	26	2.7%
High	7	0.7%
Very High	1	0.1%

The geographic distribution of the ranking of additional critical infrastructure asset vulnerability differed from the vulnerability distributions of the other three variables (Figure 4.9). When compared to all three of the other variables, between 61 and 71% of the service area rankings remained the same.

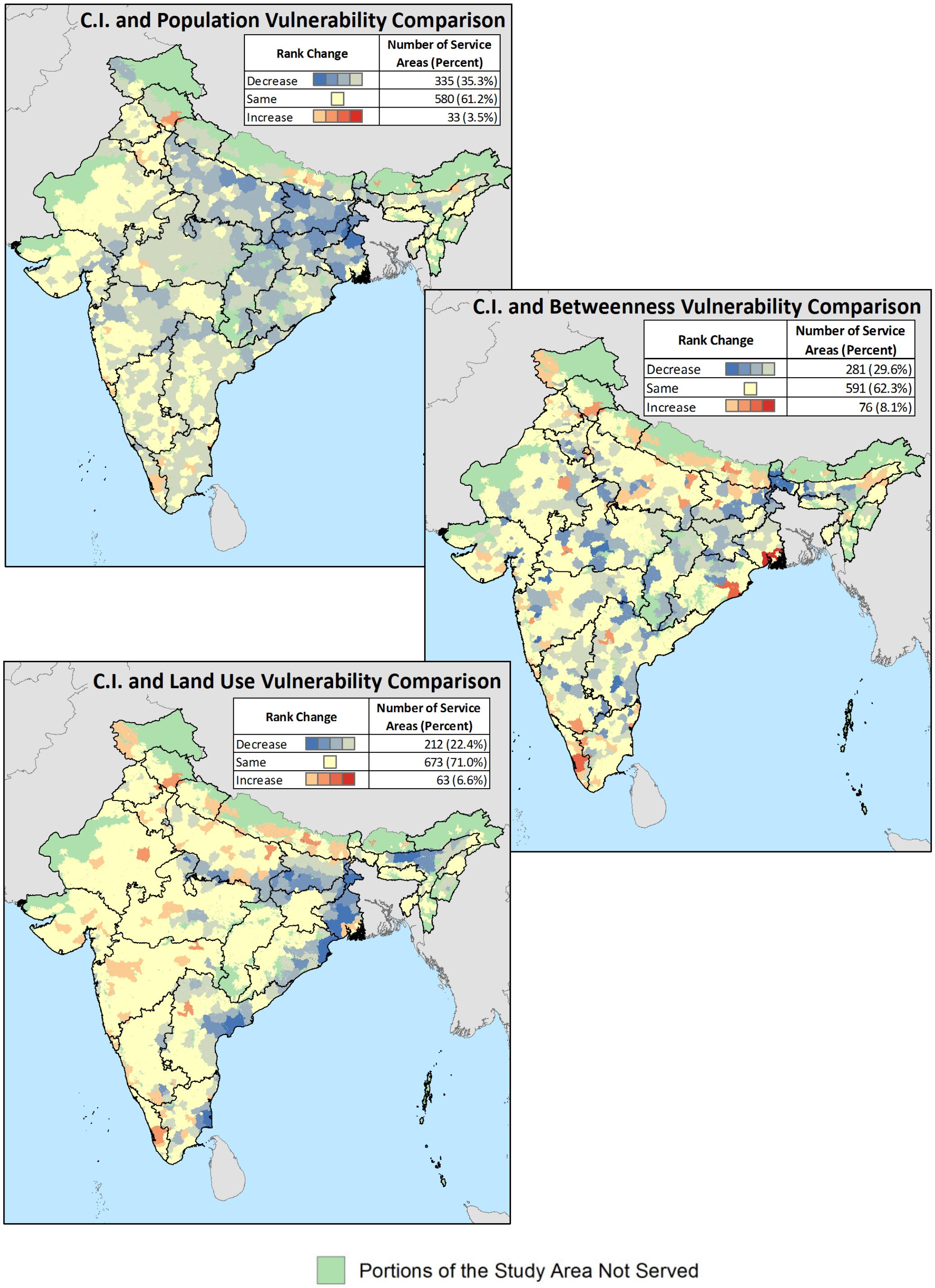


Figure 4.9: Geographic difference in vulnerability distribution for other critical infrastructure assets versus the betweenness metric, land use, and population.

4.4.1.4. Comparison of Factors for Blackout

The factors discussed all have different geographic distributions. To understand how to combine them, it was important to understand how the variables interact with each other and how substations involved ranked on each factor.

To assess how the factors interact with each other, the Spearman correlation coefficient is calculated and analyzed (Table 4.6). The correlations between any pair of variables were low to moderate; however, the highest correlations occurred between population and land use, and population and other critical assets. The moderate relationship between land use and population was explainable by the fact that land use and population are directly tied. The more people in a particular area, the more likely it is to be an "urban" area. Also, the moderate relationship between population and other critical assets can be explained by the fact that the more people in a given area, the larger the need for more critical assets, such as hospitals.

Table 4.6: Variable correlations using the Spearman method, Values significant at the 0.05 level are marked in bold.

	Betweenness	Land Use	Population	Other Critical Assets
Betweenness	1	**0.156**	**0.164**	**0.132**
Land Use	**0.156**	1	**0.369**	**0.339**
Pouplation	**0.164**	**0.369**	1	**0.422**
Other Critical Assets	**0.132**	**0.339**	**0.422**	1

For July 30th, betweenness alone could have identified 27.6% of the substations that were impacted as being vulnerable. However, by including other landscape metrics four additional substations were identified as vulnerable and reiterated other station vulnerabilities (for example, Biharshariff having vulnerabilities both in betweenness and land use). Interestingly, other critical assets would not have added any additional information, providing no indication of vulnerabilities in any of these substations. Additional critical assets may be impacted as a consequence of the vulnerabilities in the electrical grid; however, additional critical assets themselves may not be important for identifying vulnerabilities. Utilizing betweenness, population, and land use data, 41.4% of the substations affected on July 30th were identified as vulnerable, which is a 13.8% increase over betweenness alone (Table 4.7). While this percentage metric helped to explain omission errors, commission errors in this type of predictive model are difficult to assess. Since validation data was based on one instance (in this case, the July 2012 Indian blackout), it is difficult to assess if the attribution of a substations as vulnerable is an error of commission. Just because the substation was not impacted by this particular event, does not mean that it is not vulnerable.

The rankings for the substations for July 31st reflect similar results to those of July 30th. Once again, betweenness identified the most substations impacted on this day, with 24.4%, followed by land use (8.9%), and finally population (6.7%). Other critical assets continued to provide little insights into the vulnerability of these particular substations. Land use and population, however, did rank four substations as having a high or very high vulnerability that were not ranked by betweenness alone. Combined, all four variables rank 33.3% of the substations impacted on July 31st as having high or very high vulnerability, which was 8.9% more than graph metrics alone (Table 4.8).

The major benefit of comparing the variables in this fashion was to see how important each variable is to the overall goal. For example, the "other" critical assets variable did not greatly contribute to the identification of any vulnerable substations that could have been useful for preventing the 2012 Indian blackout. Betweenness ranked around a quarter of the substations as having high or very high vulnerability; land use ranked about a tenth as having high or very high vulnerability; and population ranked about six-hundredths (6%) as having high or very high vulnerability. This may be useful information performing the pairwise comparisons in the AHP portion of the research.

Table 4.7: Vulnerability rankings by variable for substations affected by July 30, 2012 blackout.

	Vulnerability Rankings				
Substation	**Betweenness**	**Land Use**	**Population**	**Other Critical Assets**	**Number of Factors Account for High Vulnerability**
Agra (PG)	2	1	1	1	0
Balia	1	3	3	1	0
Ballabarh	1	2	1	1	0
Bareilly	1	1	1	1	0
Bawana	2	1	1	1	0
Bhinmal	3	1	1	1	0
Biharshariff	4	4	4	1	2
Bina (PG)	3	1	2	1	0
Chittorgarh	1	1	1	1	0
Gorakhpur (PG)	4	1	1	1	1
Gwalior (PG)	5	1	3	2	1
Heerapura	4	1	1	1	1
Jaipur (PG)	3	1	1	1	0
Jamshedpur	4	1	1	1	1
Kanpur (PG)	4	2	1	1	1
Kota	5	1	2	1	1
Malanpur	2	5	1	1	1
Meerut	3	2	1	1	0
Muzaggarpur	3	2	4	1	1
Patna	1	5	1	1	1
PG Bisra	3	1	1	1	0
Rishikesh	1	1	2	1	0
Roorkee	2	1	2	1	0
Rourkela	4	1	2	1	1
Samaypur	2	1	1	1	0
Sanchore	1	1	1	1	0
Sasaram	3	4	2	1	1
Tehri	1	1	2	1	0
Udaipur	1	1	2	1	0
%(Very)High	27.60%	13.80%	6.90%	0.0.%	41.40%

Table 4.8: Vulnerability rankings by variable for substations affected by July 31, 2012 blackout.

	Vulnerability Rankings				
Substation	Betweenness	Land Use	Population	Other Critical Assets	Number of Factors Account for High Vulnerability
Agra	5	2	3	3	1
Agra (PG)	2	1	1	1	0
Allahabad	4	2	4	3	2
Balia	1	3	3	1	0
Ballabarh (BBMB)	1	2	1	1	0
Bhiwadi	5	1	1	1	1
Biharshariff	4	4	4	1	3
Bina (PG)	3	1	2	1	0
Budhipadar	2	1	3	1	0
Dausa	3	1	1	1	0
Debari	1	1	2	1	0
Greater Noida	1	3	3	2	0
Gorakhpur (PG)	4	1	1	1	1
Gwalior (PG)	5	1	3	2	1
Jaipur (PG)	3	1	1	1	0
Jamshedpur	4	1	1	1	1
Jaypore	4	1	1	1	1
Jodhpur	3	1	2	2	0
Kaithal	2	1	1	1	0
Kaithal (400 kV)	1	1	1	1	0
Kankroli	1	1	1	1	0
Kanpur (PG)	4	2	1	1	1
Kishenpur	3	1	1	1	0
Koteshwar	1	1	2	1	0
Maler Kotla	2	1	1	1	0
Maharani Bagh	2	5	3	3	1
Maithon	2	2	1	1	0
Mandaula	2	2	1	1	0
Meramandoli	4	3	2	1	1
Muzaffarpur	3	2	4	1	1
Panki	3	2	2	1	0
Patna	1	5	1	1	0
PG	2	1	1	1	1
PG Bisra	3	1	1	1	0
PG Kankroli	2	1	1	1	0
Raigarh	2	1	2	1	0
Ranchi	3	1	3	1	0
Rewari	3	1	1	1	0
Rourkela	4	1	2	1	1
Sabalgarh	1	1	1	1	0
Samaypur	2	1	1	1	0
Sasaram	3	4	2	1	1
Shivpuri	1	1	2	1	0
Suratgarh	1	1	2	1	0
Wagoora (PG)	1	1	1	1	0
Percent High/Very High	24.40%	8.90%	6.70%	0.00%	33.30%

4.4.2. Assessing Combinations of Variables with Equal Weights

Describing the vulnerability distributions of the factors indicated that each variable may add an extra dimension to assessing the vulnerability of a service area, and thus the vulnerability of the substation. This section details the different combinations of variables and how well the combinations assess the vulnerability of the substations affected by the Indian blackout. The various combinations included evenly weighted and a weighted analysis based on knowledge of the region to assess the impact of the weights on the results.

4.4.2.1. Two Variable Combinations

There were six two-variable combinations assessed:

- land use and population,
- land use and other critical assets,
- land use and betweenness,
- population and other critical assets,
- population and betweenness, and
- other critical assets and betweenness.

The resulting ranks for the substations affected on July 30th (Table 4.9) and July 31st (Table 4.10) were compared to the rankings for the same nodes for betweenness.

For the 29 substations affected on July 30th, the two-variable combination that ranked the greatest percentage of affected substations as having high or very high vulnerability was land use and betweenness (44.8%), followed closely by population and betweenness (37.9%). Using betweenness alone, 27.6% of the affected substations were identified as having a high or very high vulnerability. Additionally, the two-variable pairs with the highest percentages identified the same substations as having high and very high vulnerabilities as betweenness alone; however, the two variable pairs, such as land use and betweenness, added additional substations to the high and very highly vulnerable categories. Combinations that included other critical assets did not have high percentages of substations identified. Additionally, the combination of land use and population only identified 20.7% of the affected substations as having high or very high vulnerability (Table 4.9).

The substations impacted by the blackout on July 31st exhibited similar results. The two-variable combinations with the highest percentage of affected substations being ranked as having high or very high vulnerability are the land use and betweenness, along with the population and betweenness combinations. Like for the July 30th event, these two combinations identified the same substations as betweenness alone, but also identified other vulnerable substations. As with July 30th, the combinations containing other critical assets had the lowest identification percentages (Table 4.10).

Table 4.9: Vulnerability rankings by two-variable pair with even weights and by betweenness for substations affected by July 30, 2012 blackout.

	Vulneraiblity Rankings - Even Weights						
Substation	**Betweenness**	**Land Use-Population**	**Land Use - Other Critical Assets**	**Land Use - Betweennness**	**Population - Other Critical Assets**	**Population - Betweenness**	**Other Critical Assets - Betweenness**
Agra (PG)	2	1	1	2	1	2	2
Balia	1	4	3	3	3	3	1
Ballabarh (BBMB)	1	2	2	2	1	1	1
Bareilly	1	1	1	1	1	1	1
Bawana	2	1	1	2	1	2	2
Bhinmal	3	1	1	3	1	3	3
Biharshariff	4	5	4	5	4	5	4
Bina (PG)	3	2	1	3	2	4	3
Chittorgarh	1	1	1	1	1	1	1
Gorakhpur (PG)	4	1	1	4	1	4	4
Gwalior (PG)	5	3	2	4	4	5	5
Heerapura	4	1	1	4	1	4	4
Jaipur (PG)	3	1	1	3	1	3	3
Jamshedpur	4	1	1	4	1	4	4
Kanpur (PG)	4	2	2	4	1	4	4
Kota	5	2	1	4	2	5	5
Malanpur	2	4	4	5	1	2	2
Meerut	3	2	2	4	1	3	3
Muzaggarpur	3	4	2	4	4	5	3
Patna	1	4	4	4	1	1	1
PG Bisra	3	1	1	3	1	3	3
Rishikesh	1	2	1	1	2	2	1
Roorkee	2	2	1	2	2	3	2
Rourkela	4	2	1	4	2	4	4
Samaypur	2	1	1	2	1	2	2
Sanchore	1	1	1	1	1	1	1
Sasaram	3	4	4	5	2	4	3
Tehri	1	2	1	1	2	2	1
Udaipur	1	2	1	1	2	2	1
Percent High/Very High	27.60%	20.70%	13.80%	**44.80%**	10.30%	**37.90%**	27.60%

Table 4.10: Vulnerability rankings by two-variable pair with even weights and by betweenness for substations affected by July 31, 2012 blackout.

	Vulneraiblity Rankings						
Substation	Betweenness	Land Use-Population	Land Use - Other Critical Assets	Land Use - Betweennness	Population - Other Critical Assets	Population - Betweenness	Other Critical Assets - Betweenness
Agra	5	3	4	5	4	5	5
Agra (PG)	2	1	1	2	1	2	2
Allahabad	4	4	4	4	4	5	5
Balia	1	4	3	4	3	3	1
Ballabarh (BBMB)	1	2	2	2	1	1	1
Bhiwadi	5	1	1	4	1	4	5
Biharshariff	4	5	4	5	4	5	4
Bina (PG)	3	2	1	4	2	4	3
Budhipadar	2	3	1	2	3	4	2
Dausa	3	1	1	3	1	3	3
Debari	1	2	1	1	2	2	1
Greater Noida	1	4	4	4	4	3	2
Gorakhpur (PG)	4	1	1	4	1	4	4
Gwalior (PG)	5	3	2	4	4	5	5
Jaipur (PG)	3	1	1	3	1	3	3
Jamshedpur	4	1	1	4	1	4	3
Jaypore	4	1	1	4	1	4	4
Jodhpur	3	2	2	4	3	4	4
Kaithal	2	1	1	2	1	2	2
Kaithal (400 kV)	1	1	1	1	1	1	1
Kankroli	1	1	1	1	1	1	1
Kanpur (PG)	4	2	2	4	1	4	4
Kishenpur	3	1	1	3	1	3	3
Koteshwar	1	2	1	1	2	2	1
Maler Kotla	2	1	1	2	1	2	1
Maharani Bagh	2	5	5	5	4	4	4
Maithon	2	2	2	3	1	2	2
Mandaula	2	2	2	3	1	2	2
Meramandoli	4	3	3	5	2	4	4
Muzaffarpur	3	4	2	4	4	5	3
Panki	3	3	2	4	2	4	3
Patna	1	4	4	4	1	1	1
PG	2	1	1	2	1	2	2
PG Bisra	3	1	1	3	1	3	3
PG Kankroli	2	1	1	2	1	2	2
Raigarh	2	2	1	2	2	3	2
Ranchi	3	3	1	4	3	4	3
Rewari	3	1	1	3	1	3	3
Rourkela	4	2	1	4	2	4	4
Sabalgarh	1	1	1	1	1	1	1
Samaypur	2	1	1	2	1	2	2
Sasaram	3	4	4	5	2	4	3
Shivpuri	1	2	1	1	2	2	1
Suratgarh	1	2	1	1	2	2	1
Wagoora (PG)	1	1	1	1	1	1	1
Percent High/Very High	24.40%	17.80%	15.60%	**46.60%**	15.60%	**42.20%**	28.90%

4.4.2.2. Three Variable Combinations

There were four three-variable combinations analyzed with equal weights:

- land use – population – other critical assets,
- land use – population – betweenness,
- land use – other critical assets – betweenness, and
- population – other critical assets – betweenness.

These combinations were compared to events on both July 30 and 31 to determine each combination's effectiveness.

With reference to the substations that were impacted during the July 30th event, the Land Use – Population - Betweenness (21%) combination performed the best of the three variable combinations, only second to assessing betweenness alone (27.6%). While most of the substations maintained the same ranking when comparing the betweenness-only vulnerabilities with the Land Use – Population – Betweenness combination, five of the substations that were ranked with high or very high vulnerabilities in the betweenness-only analysis had a moderate vulnerability in the Land Use - Population – Betweenness combination. In addition to these changes, three of the substations impacted on July 30th changed rank from either very low, low, or moderate vulnerabilities to high or very high vulnerabilities. By combining the results of betweenness-only analysis and the land use – population – betweenness combination, 37.9% of the substations were ranked as having high or very high vulnerabilities (Table 4.11).

Table 4.11: Vulnerability rankings by three-variable pairs with even weights and by betweenness for substations affected by July 30, 2012 blackout.

	Vulnerability Rankings				
Substation	**Betweenness**	**Land Use - Population - Other Critical Assets**	**Land Use - Population - Betweenness**	**Land Use - Other Critical Assets - Betweenness**	**Population - Other Critical Assets - Betwenness**
Agra (PG)	2	1	2	3	2
Balia	1	3	3	3	3
Ballabarh	1	2	2	3	1
Bareilly	1	1	1	1	1
Bawana	2	1	2	3	2
Bhinmal	3	1	2	3	3
Biharshariff	4	4	5	5	5
Bina (PG)	3	2	2	3	3
Chittorgarh	1	1	1	1	1
Gorakhpur (PG)	4	1	2	3	3
Gwalior (PG)	5	3	4	5	5
Heerapura	4	1	2	3	3
Jaipur (PG)	3	1	2	3	3
Jamshedpur	4	1	2	3	3
Kanpur (PG)	4	2	3	3	3
Kota	5	2	4	4	4
Malanpur	2	3	4	2	2
Meerut	3	2	2	3	3
Muzaggarpur	3	3	4	4	4
Patna	1	3	3	1	1
PG Bisra	3	1	2	3	3
Rishikesh	1	2	2	2	2
Roorkee	2	2	2	3	3
Rourkela	4	2	3	4	4
Samaypur	2	1	2	2	2
Sanchore	1	1	1	1	1
Sasaram	3	3	4	3	3
Tehri	1	2	2	2	2
Udaipur	1	2	2	2	2
%(Very)High	**27.60%**	3.40%	**20.70%**	17.20%	17.20%

The substations impacted by the July 31st event exhibited similar results; however, the highest scoring combinations in addition to the betweenness-only analysis, were the land use – other critical assets – betweenness and the population – other critical assets – betweenness combinations. These high percentages were likely due to critical infrastructure having little impact on the vulnerability rankings of any of these substations, so the results are similar to that of using population – betweenness and land use –betweenness pairings, dampened by the inclusion of critical infrastructure. With this knowledge, the land use – population – betweenness combination was the best performer. Three substations that were ranked as having high or very high vulnerabilities in the betweenness-only analysis were ranked as having a moderate vulnerability in the land Use – population – betweenness combination, but that was in addition to the three impacted substations in the high or very high vulnerability categories. Another important distinction was that none of the New Delhi substations impacted by the July 31st blackout were ranked as having a high or very high vulnerability in the betweenness-only analysis, but one, Maharani Bagh substation, was identified as having a very high vulnerability in the land use – population – betweenness combination. Using only the betweenness metric would have missed entirely the New Delhi area's vulnerability to blackouts (Table 4.12).

Table 4.12: Vulnerability rankings by three-variable pair with even weights and by betweenness for substations affected by July 31, 2012 blackout.

	Vulnerability Rankings				
Substation	Betweenness	Land Use - Population - Other Critical Assets	Land Use - Population - Betweenness	Land Use - Other Critical Assets - Betweenness	Population - Other Critical Assets - Betweenness
Agra	5	4	5	5	5
Agra (PG)	2	1	2	3	2
Allahabad	4	4	5	5	5
Balia	1	3	3	3	3
Ballabarh (BBMB)	1	2	2	3	1
Bhiwadi	5	1	3	4	4
Biharshariff	4	4	5	5	5
Bina (PG)	3	2	2	3	3
Budhipadar	2	2	2	3	3
Dausa	3	1	2	3	3
Debari	1	2	2	1	2
Greater Noida	1	4	3	3	3
Gorakhpur (PG)	4	1	2	3	3
Gwalior (PG)	5	3	4	5	5
Jaipur (PG)	3	1	2	3	3
Jamshedpur	4	1	2	3	3
Jaypore	4	1	2	3	3
Jodhpur	3	2	2	3	4
Kaithal	2	1	2	3	2
Kaithal (400 kV)	1	1	1	1	1
Kankroli	1	1	1	1	1
Kanpur (PG)	4	2	3	4	3
Kishenpur	3	1	2	3	3
Koteshwar	1	2	2	1	2
Maler Kotla	2	1	2	1	1
Maharani Bagh	2	5	5	5	4
Maithon	2	2	2	3	2
Mandaula	2	2	2	3	2
Meramandoli	4	3	4	5	4
Muzaffarpur	3	3	4	3	4
Panki	3	2	3	3	3
Patna	1	3	3	4	1
PG	2	1	2	3	2
PG Bisra	3	1	2	3	3
PG Kankroli	2	1	2	3	2
Raigarh	2	2	2	3	3
Ranchi	3	2	3	3	4
Rewari	3	1	2	3	3
Rourkela	4	2	3	3	4
Sabalgarh	1	1	1	1	1
Samaypur	2	1	2	3	2
Sasaram	3	3	4	5	3
Shivpuri	1	2	2	1	2
Suratgarh	1	2	2	1	2
Wagoora (PG)	1	1	1	1	1
Percent High/Very High	24.40%	11.10%	17.80%	22.20%	24.40%

4.4.2.3. Four-Variable Combination

Based on the previous results of the combinations of the variables, it was clear that the four-variable combination would likely not be the best choice, as the other critical assets information added no additional information when combined with any of the other factors. 20.7% of the substations impacted on July 30th were ranked as having high or very high vulnerability, compared to 27.6% identified with the same rankings using betweenness alone. Three substations not identified only using betweenness were identified, while losing five impacted substations to the moderate vulnerability class (Table 4.12). The substations impacted on July 31st showed a similar pattern to the two-variable and three-variable combinations where all of the combinations were less effective in identifying impacted substations as having high or very high vulnerabilities. Six of the substations impacted on July 31st that were identified as having high or very high vulnerability in the betweenness-only analysis were considered as having moderate vulnerability in the all-variable combination. That being said, four substations not identified by the betweenness-only analysis were identified as having high or very high vulnerability in the all-variable combination. Like in the three-variable combinations, this all-variable combination identified substations in New Delhi as having very high vulnerability, and thus, would have been potentially protected beforehand for having known vulnerabilities (Table 4.13).

Table 4.13: Vulnerability rankings for an all-variable combination with even weights and for betweenness for substations affected by both days of the Indian blackout.

July 30, 2012

	Vulnerability Rankings	
Substation	**Betweenness**	**Even All Four**
Agra (PG)	2	2
Balia	1	3
Ballabarh	1	2
Bareilly	1	1
Bawana	2	2
Bhinmal	3	2
Biharshariff	4	5
Bina (PG)	3	3
Chittorgarh	1	1
Gorakhpur (PG)	4	3
Gwalior (PG)	5	4
Heerapura	4	3
Jaipur (PG)	3	2
Jamshedpur	4	3
Kanpur (PG)	4	3
Kota	5	4
Malanpur	2	4
Meerut	3	3
Muzaggarpur	3	4
Patna	1	3
PG Bisra	3	2
Rishikesh	1	2
Roorkee	2	2
Rourkela	4	3
Samaypur	2	2
Sanchore	1	1
Sasaram	3	4
Tehri	1	2
Udaipur	1	2
Percent High/Very High	27.60%	20.70%

July 31, 2012

	Vulnerability Rankings	
Substation	**Betweenness**	**Even All Four**
Agra	5	5
Agra (PG)	2	2
Allahabad	4	5
Balia	1	3
Ballabarh (BBMB	1	2
Bhiwadi	5	3
Biharshariff	4	5
Bina (PG)	3	3
Budhipadar	2	3
Dausa	3	2
Debari	1	2
Greater Noida	1	4
Gorakhpur (PG)	4	3
Gwalior (PG)	5	4
Jaipur (PG)	3	2
Jamshedpur	4	3
Jaypore	4	3
Jodhpur	3	3
Kaithal	2	2
Kaithal (400 kV)	1	1
Kankroli	1	1
Kanpur (PG)	4	3
Kishenpur	3	2
Koteshwar	1	2
Maler Kotla	2	1
Maharani Bagh	2	5
Maithon	2	2
Mandaula	2	2
Meramandoli	4	4
Muzaffarpur	3	4
Panki	3	3
Patna	1	3
PG	2	2
PG Bisra	3	2
PG Kankroli	2	2
Raigarh	2	2
Ranchi	3	3
Rewari	3	2
Rourkela	4	3
Sabalgarh	1	1
Samaypur	2	2
Sasaram	3	4
Shivpuri	1	2
Suratgarh	1	2
Wagoora (PG)	1	1
Percent High/Very High	24.40%	20.00%

4.4.3. Weighted Versions of Best Performer of Variable Combinations

Some variable combinations performed better than others. For example, those combinations including other critical assets tended to perform more poorly than those containing land use or population. This section used AHP to weight the best variable combinations based on the analysis of how much each variable contributed to the identifying vulnerabilities in those substations impacted by the Indian blackout. Pairwise comparisons were performed based on the pertinent information and averaged based on response of several experts.

4.4.3.1. Weighted Two-Variable Combinations

The best performers of the two-variable combinations were the land use and betweenness combination and the land use and population combination. When analyzing each variable individually, betweenness was able to account for more substations during both days, so it is considered a more important variable, followed by land use, and then population. The pairwise comparisons reflect these findings. In the land use and betweenness combination, the resulting weights from the AHP analysis had betweenness weighted 0.75 and land use weighted 0.25. In the land use and population combination, the resulting weights from the AHP analysis had betweenness weighted as 0.89 and population weighted as 0.12, as population was the least important variable in determining the vulnerabilities of substations. For both July 30th and 31st, the land use and betweenness combination performed the best (41.38% and 33.3%, respectively). Comparing the results of the land use and betweenness and the population and betweenness combinations indicate that the population and betweenness combination added no new information to the analysis, identifying fewer substations as high or very highly vulnerable. Those substations that were ranked as having high or very high vulnerability were also identified as such by the land use and betweenness combination (Table 4.14).

Table 4.14: Weighted two-variable combination results.

July 30, 2012			
	Vulneraiblity Rankings		
Substation	Betweenness	Land Use - Betweennness	Population - Betweenness
Agra (PG)	2	2	2
Balia	1	2	1
Ballabarh (BBMB)	1	2	1
Bareilly	1	1	1
Bawana	2	2	2
Bhinmal	3	3	3
Biharshariff	4	5	5
Bina (PG)	3	3	3
Chittorgarh	1	1	1
Gorakhpur (PG)	4	4	4
Gwalior (PG)	5	5	5
Heerapura	4	4	4
Jaipur (PG)	3	3	3
Jamshedpur	4	4	4
Kanpur (PG)	4	5	4
Kota	5	5	5
Malanpur	2	4	2
Meerut	3	4	3
Muzaggarpur	3	4	4
Patna	1	3	1
PG Bisra	3	3	3
Rishikesh	1	1	1
Roorkee	2	2	2
Rourkela	4	4	4
Samaypur	2	2	2
Sanchore	1	1	1
Sasaram	3	4	3
Tehri	1	1	1
Udaipur	1	1	1
Percent High/Very High	27.60%	**41.38%**	31.03%

July 31, 2012			
	Vulneraiblity Rankings		
Substation	Betweenness	Land Use - Betweennness	Population - Betweenness
Agra	5	5	5
Agra (PG)	2	2	2
Allahabad	4	5	5
Balia	1	2	1
Ballabarh (BBMB)	1	2	1
Bhiwadi	5	5	5
Biharshariff	4	5	5
Bina (PG)	3	3	3
Budhipadar	2	2	2
Dausa	3	3	3
Debari	1	1	1
Greater Noida	1	2	1
Gorakhpur (PG)	4	4	4
Gwalior (PG)	5	5	5
Jaipur (PG)	3	3	3
Jamshedpur	4	4	4
Jaypore	4	4	4
Jodhpur	3	3	3
Kaithal	2	2	2
Kaithal (400 kV)	1	1	1
Kankroli	1	1	1
Kanpur (PG)	4	5	4
Kishenpur	3	3	3
Koteshwar	1	1	1
Maler Kotla	2	2	2
Maharani Bagh	2	4	2
Maithon	2	3	2
Mandaula	2	3	2
Meramandoli	4	5	4
Muzaffarpur	3	4	4
Panki	3	4	3
Patna	1	3	1
PG	2	2	2
PG Bisra	3	3	3
PG Kankroli	2	2	2
Raigarh	2	2	2
Ranchi	3	3	3
Rewari	3	3	3
Rourkela	4	4	4
Sabalgarh	1	1	1
Samaypur	2	2	2
Sasaram	3	4	3
Shivpuri	1	1	1
Suratgarh	1	1	1
Wagoora (PG)	1	1	1
Percent High/Very High	24.40%	**33.30%**	26.70%

4.4.3.2. Weighted Three-Variable Combinations

The three variable combinations selected for weighted analysis were the land use – betweenness – population, the land use – other critical assets – betweenness, and the population – other critical assets – betweenness combinations. The weights were once again determined based on pairwise comparisons of the relationship of each variable to explaining the substations impacted by the black out. In the land use – betweenness – population combination, land use was weighted as 0.24, population as 0.06, and betweenness at 0.70. In the land use – other critical assets – betweenness combination, land use was weighted as 0.25, betweenness as 0.70, and other critical assets as 0.05. The final combination, population – other critical assets – betweenness, had population weighted as 0.21, other critical assets weighted as 0.05, and betweenness weighted as 0.74. For both blackouts, each combination had relatively the same results, only slightly better than the betweenness-only analysis (between 2 and 4% better) (Table 4.15). Additionally, any substations identified with high or very high vulnerabilities were also identified in the land use and betweenness combination from the two-variable combinations as having high or very high vulnerabilities.

Table 4.15: Weighted three-variable combination results.

July 30, 2012

Substation	Vulnerability Rankings: Betweenness	Land Use - Population - Betweenness	Land Use - Other Critical Assets - Betweenness	Population - Other Critical Assets - Betwenness
Agra (PG)	2	2	2	2
Balia	1	2	2	1
Ballabarh	1	1	1	1
Bareilly	1	1	1	1
Bawana	2	2	2	2
Bhinmal	3	3	3	3
Biharshariff	4	5	5	5
Bina (PG)	3	3	3	3
Chittorgarh	1	1	1	1
Gorakhpur (PG)	4	4	4	4
Gwalior (PG)	5	5	5	5
Heerapura	4	4	4	4
Jaipur (PG)	3	3	3	3
Jamshedpur	4	4	4	4
Kanpur (PG)	4	4	4	4
Kota	5	5	5	5
Malanpur	2	3	3	2
Meerut	3	3	3	3
Muzaggarpur	3	3	3	4
Patna	1	2	2	1
PG Bisra	3	3	3	3
Rishikesh	1	1	1	1
Roorkee	2	2	2	2
Rourkela	4	4	4	4
Samaypur	2	2	2	2
Sanchore	1	1	1	1
Sasaram	3	4	4	3
Tehri	1	1	1	1
Udaipur	1	1	1	1
Percent High/ Very High	27.60%	31.00%	31.00%	31.00%

July 31, 2012

Substation	Vulnerability Rankings: Betweenness	Land Use - Population - Betweenness	Land Use - Other Critical Assets - Betweenness	Population - Other Critical Assets - Betwenness
Agra	5	5	5	5
Agra (PG)	2	2	2	2
Allahabad	4	4	5	5
Balia	1	2	2	1
Ballabarh (BBMB)	1	1	1	1
Bhiwadi	5	5	5	5
Biharshariff	4	5	5	5
Bina (PG)	3	3	3	3
Budhipadar	2	2	2	2
Dausa	3	3	3	3
Debari	1	1	1	1
Greater Noida	1	2	2	1
Gorakhpur (PG)	4	4	4	4
Gwalior (PG)	5	5	5	5
Jaipur (PG)	3	3	3	3
Jamshedpur	4	4	4	4
Jaypore	4	4	4	4
Jodhpur	3	3	3	3
Kaithal	2	2	2	2
Kaithal (400 kV)	1	1	1	1
Kankroli	1	1	1	1
Kanpur (PG)	4	4	4	4
Kishenpur	3	3	3	3
Koteshwar	1	1	1	1
Maler Kotla	2	2	1	1
Maharani Bagh	2	3	4	3
Maithon	2	2	2	2
Mandaula	2	2	2	2
Meramandoli	4	5	5	4
Muzaffarpur	3	3	3	4
Panki	3	3	3	3
Patna	1	2	2	1
PG	2	2	2	2
PG Bisra	3	3	3	3
PG Kankroli	2	2	2	2
Raigarh	2	2	2	2
Ranchi	3	3	3	3
Rewari	3	3	3	3
Rourkela	4	4	4	4
Sabalgarh	1	1	1	1
Samaypur	2	2	2	2
Sasaram	3	4	4	3
Shivpuri	1	1	1	1
Suratgarh	1	1	1	1
Wagoora (PG)	1	1	1	1
Percent High/Very High	24.40%	26.70%	28.90%	26.70%

4.4.3.3. All Variable Weighted Combination

The final combination includes all of the variables: betweenness, population, land use, and other critical assets. As in the two and three-variable combinations, the weights were determined based on AHP with pairwise comparisons of each variable.

For this analysis, betweenness was weighted as 0.62, population as 0.10, land use as 0.24, and additional critical assets as 0.04. For both blackout days this four variable combination performed better than betweenness alone; however, this performance was still not better than the betweenness and land use combination. Additionally, the four-variable combination provided no new information when compared to the betweenness and land use combination; all of the substations identified as having a high or very high vulnerability were identified by both combinations, with more information being provided by the betweenness – land use combination (Table 4.16).

Table 4.16: Weighted all-variable combination results.

July 30, 2012		
	Vulnerability Rankings	
Substation	Betweenness	Weighted All Four
Agra (PG)	2	2
Balia	1	2
Ballabarh	1	1
Bareilly	1	1
Bawana	2	2
Bhinmal	3	3
Biharshariff	4	5
Bina (PG)	3	3
Chittorgarh	1	1
Gorakhpur (PG)	4	4
Gwalior (PG)	5	5
Heerapura	4	4
Jaipur (PG)	3	3
Jamshedpur	4	4
Kanpur (PG)	4	4
Kota	5	5
Malanpur	2	3
Meerut	3	3
Muzaggarpur	3	4
Patna	1	2
PG Bisra	3	3
Rishikesh	1	1
Roorkee	2	2
Rourkela	4	4
Samaypur	2	2
Sanchore	1	1
Sasaram	3	4
Tehri	1	1
Udaipur	1	1
Percent High/Very High	27.60%	38.50%

July 31, 2012		
	Vulnerability Rankings	
Substation	Betweenness	Weighted All Four
Agra	5	5
Agra (PG)	2	2
Allahabad	4	5
Balia	1	2
Ballabarh (BBMB)	1	1
Bhiwadi	5	5
Biharshariff	4	5
Bina (PG)	3	3
Budhipadar	2	2
Dausa	3	3
Debari	1	1
Greater Noida	1	2
Gorakhpur (PG)	4	4
Gwalior (PG)	5	5
Jaipur (PG)	3	3
Jamshedpur	4	4
Jaypore	4	4
Jodhpur	3	3
Kaithal	2	2
Kaithal (400 kV)	1	1
Kankroli	1	1
Kanpur (PG)	4	4
Kishenpur	3	3
Koteshwar	1	1
Maler Kotla	2	1
Maharani Bagh	2	4
Maithon	2	2
Mandaula	2	2
Meramandoli	4	5
Muzaffarpur	3	4
Panki	3	3
Patna	1	2
PG	2	2
PG Bisra	3	3
PG Kankroli	2	2
Raigarh	2	2
Ranchi	3	3
Rewari	3	3
Rourkela	4	4
Sabalgarh	1	1
Samaypur	2	2
Sasaram	3	4
Shivpuri	1	1
Suratgarh	1	1
Wagoora (PG)	1	1
Percent High/Very High	24.40%	31.10%

4.4.4. Graph Metrics and Non-Climatic Factor Findings

The analyses performed for the second approach indicate the importance of landscape factors in helping to assess the vulnerability of the grid. Of the greatest importance (in addition to the betweenness metric) was land use. The combination of land use and betweenness provided the most information about substations that were and are vulnerable. Of particular interest was the addition of Maharani Bagh substation as being vulnerable with the addition of land use. Using only betweenness Maharani Bagh had a low vulnerability, as did all of the substations surrounded New Delhi; however, with the addition of land use, the New Delhi Area substations, such as Maharani Bagh, had elevated vulnerability, as Delhi is a highly urban area of India. Interestingly, the evenly weighted combinations had a greater accuracy the weighted combinations by around 3%. Figure 4.10 illustrates the even weights results for the impacted areas. When compared to the betweenness-only analysis (Figures 4.4 and 4.5), a much larger portion of the impacted area was identified as having a high or very high vulnerability.

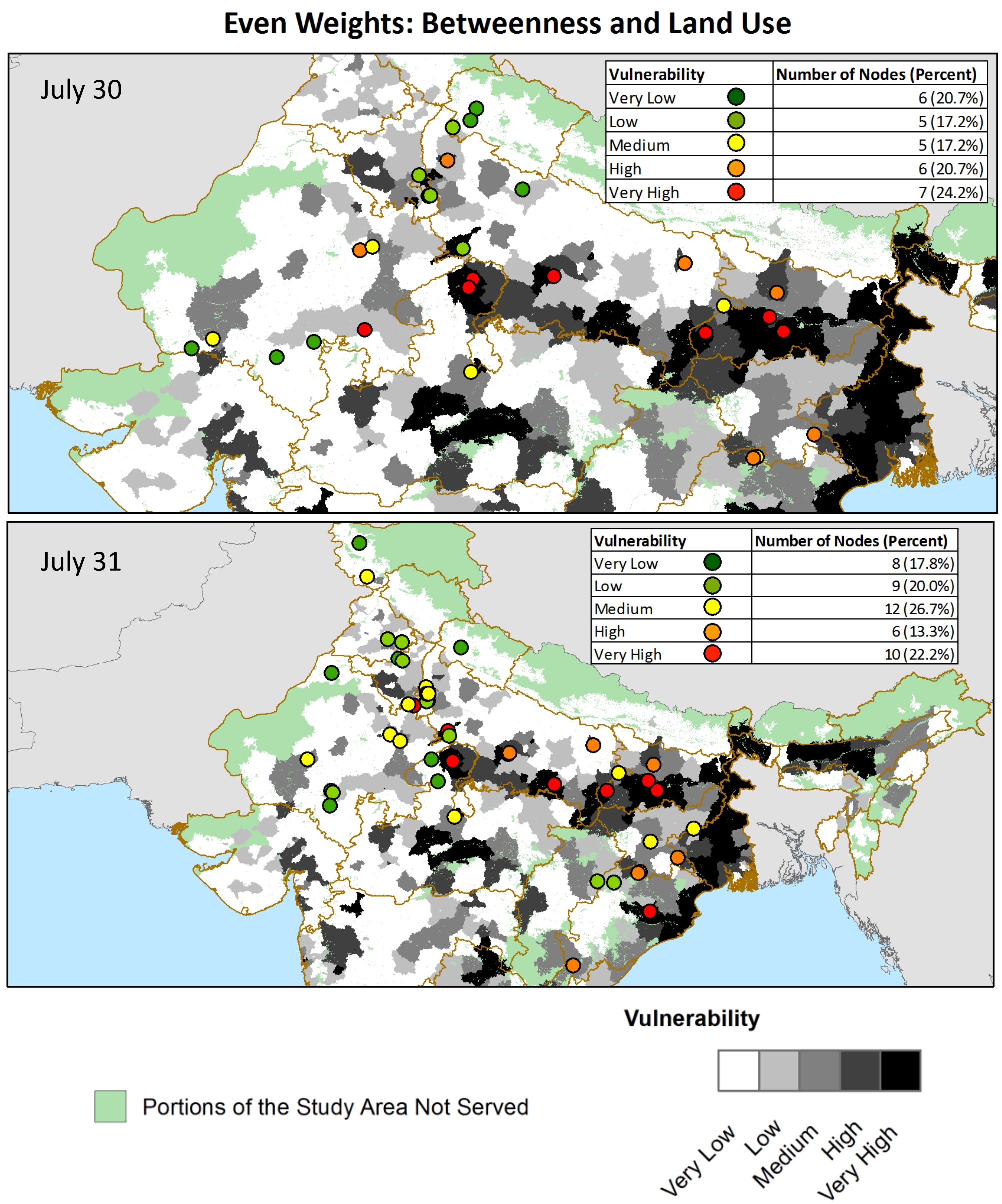


Figure 4.10: Even weights results for impacted substations for the betweenness and land use combination.

There is no difference in the rankings of the substations between the weighted and even weights for the betweenness – land use combination for July 30th; however, there are slight differences in the rankings for July 31st (Figure 4.10). It is also interesting to note the effects of the weights. In even weights, both variables are weighted 0.5, while in the weighted analysis, betweenness is weighted 0.75 and land use is weighted 0.25. That small change in the weights dampens the scores of some substations that were vulnerable when evenly weighted (Figure 4.10 and 4.11).

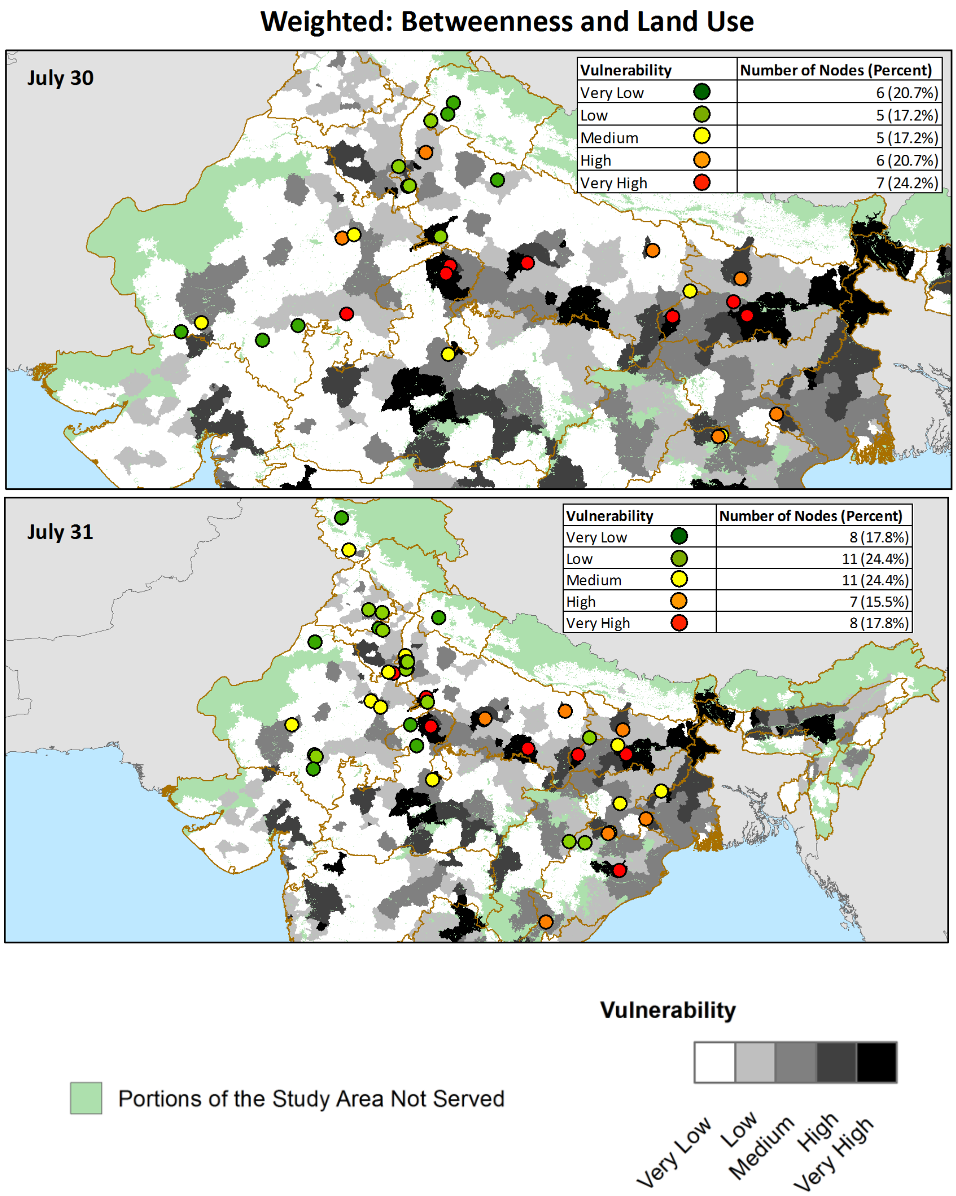


Figure 4.11: Weighted results for impacted substations for the betweenness and land use combination.

Overall this combination shows an interesting pattern of vulnerability (Figure 4.12). The majority of the substations ranked with high or very high vulnerabilities were either very large substations or were located in the eastern region of the study area. In both analyses there is a belt of vulnerable service areas running through north central India across and through east India. Additionally, there are substations ranked with high and very high vulnerabilities disbursed throughout the country, but once again, mainly in the eastern part of the study area. Interestingly, neither Nepal nor Bhutan have any service areas ranked with high or very high vulnerabilities.

There may be several reasons why the betweenness and land use combination was the best performing combination assessed. Firstly, land use can often be seen as a proxy for demand. For instance, urban areas obviously utilize more power than a forest, and thus can be more vulnerable to electrical outages. The Enquiry Committee report (Ministry of Power 2012) for the Indian blackout indicated Delhi's vulnerability because their metro system is dependent on electrical power. In addition to structural vulnerabilities, there are engineering vulnerabilities where other critical infrastructure systems depend on the electrical grid to provide them with the power they need to function for the citizenry. Additionally, media sources widely cite the drought and temperature extremes as reasons for the blackout, because farmers are not receiving the necessary rain water to keep their crops healthy (Philpott and Jones 2012). To circumvent this, farmers are using more power to run their water pumps to irrigate their crops and maintain their financial viability (Philpott and Jones 2012).

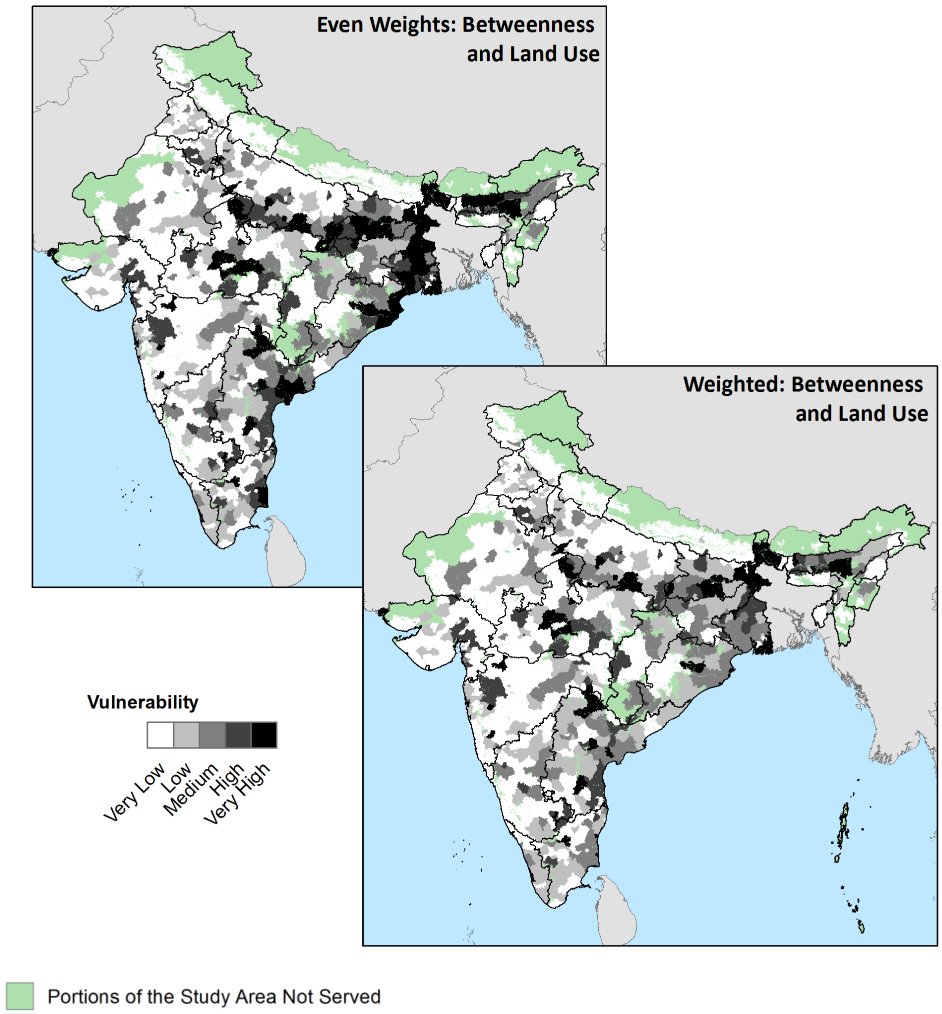


Figure 4.12: Geographic distribution of vulnerability rankings for the betweenness and land use combinations.

4.5. Inclusion of Climate and Natural Hazards Results

While there is little indication that climate or natural hazards played a role in the July 2012 black out, these variable may be important for other events, or assessing the overall vulnerability of the Indian – Nepali – Bhutanese grid system. Additionally, little attention has been given to the inclusion of climate variables and natural hazards in the critical infrastructure protection literature, so this dissertation approach seeks to provide a basic method for addressing the dearth of climate variable inclusion.

4.5.1. Assessment of the Variables

Natural hazards occur in any place where humans are, and knowing where past hazards have occurred can be useful information for knowing what substations are vulnerable to such hazards in the future. Of particular interest in this dissertation research were cyclones (called typhoons in this part of the world), landslides, and earthquakes. The geographic distribution of the vulnerability to these hazards was based on the frequency of past occurrences (see Chapter 3 for more information). For example, the eastern part of India tended to have more cyclones than the western part. Additionally, the northern part of the study area, including Bhutan and Nepal, was more prone to landslides, likely due to its mountainous terrain. The study area was also most prone to earthquakes in the northern region due to the plate boundary; additionally, Gujarat (western India) also experienced large earthquakes in its recent history. Note, that much of central India had a fairly low vulnerability to these hazards (Figure 4.13). It is important to note that these hazards have very little bearing on the events in July 2012. This information, however, should be used to assess general substation vulnerability.

Climate extremes were another variable that had not often been included in critical infrastructure vulnerability analyses; however, increases in temperature often lead to increases in demand. The temperature extremes dataset was derived from the daily maximum temperature and temperature maximum normal temperatures (see Chapter 3 for a more robust discussion). While temperature extremes appeared to have little to do with this particular blackout, temperature extremes explained some of the outages in the northern region that were not considered vulnerable either in terms of topology (betweenness) or land use (Figure 4.14). On July 30th, 6.9% of the impacted substations were located in regions that experienced a very high diversion from normal temperatures, and on July 31st 6.7% of the impacted substations experienced the same.

The Spearman correlations values between the variables (betweenness, land use, population, "other" critical assets, cyclones, earthquakes, landslides, and temperature extremes) show some interesting patterns (Table 4.17). For the majority of the variables, the correlations were only either slightly positive or slightly negative, with few moderate to strong correlations. One strong correlation was the correlation between the temperature extremes of July 30 and July 31. Of course the temperature extremes for July 30 and 31 should be highly correlated, because the temperatures at any given place should not be too different from day to day. Additionally, these two variables would never be in the same model, so their high correlation was of little concern to the research. One interesting moderate correlation was between landslides and the temperature extremes for July 30 and July 31. This may be indicative of the fact that the majority of the very high temperature extremes were in the northern part of the study area, located in the foothills of the Himalayas. This was where the majority of the landslides occur in the study region.

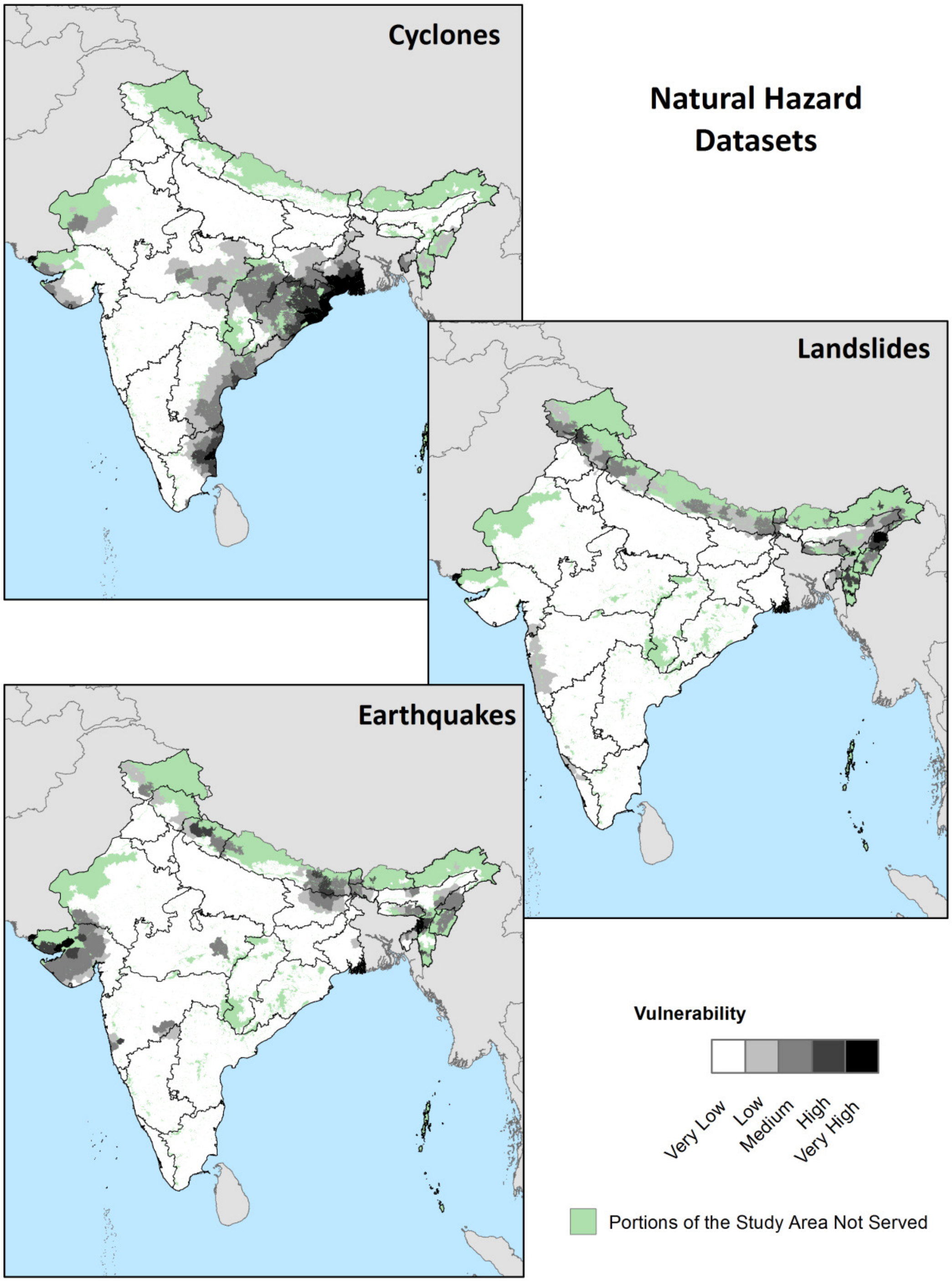


Figure 4.13: Geographic distribution of natural hazard vulnerability.

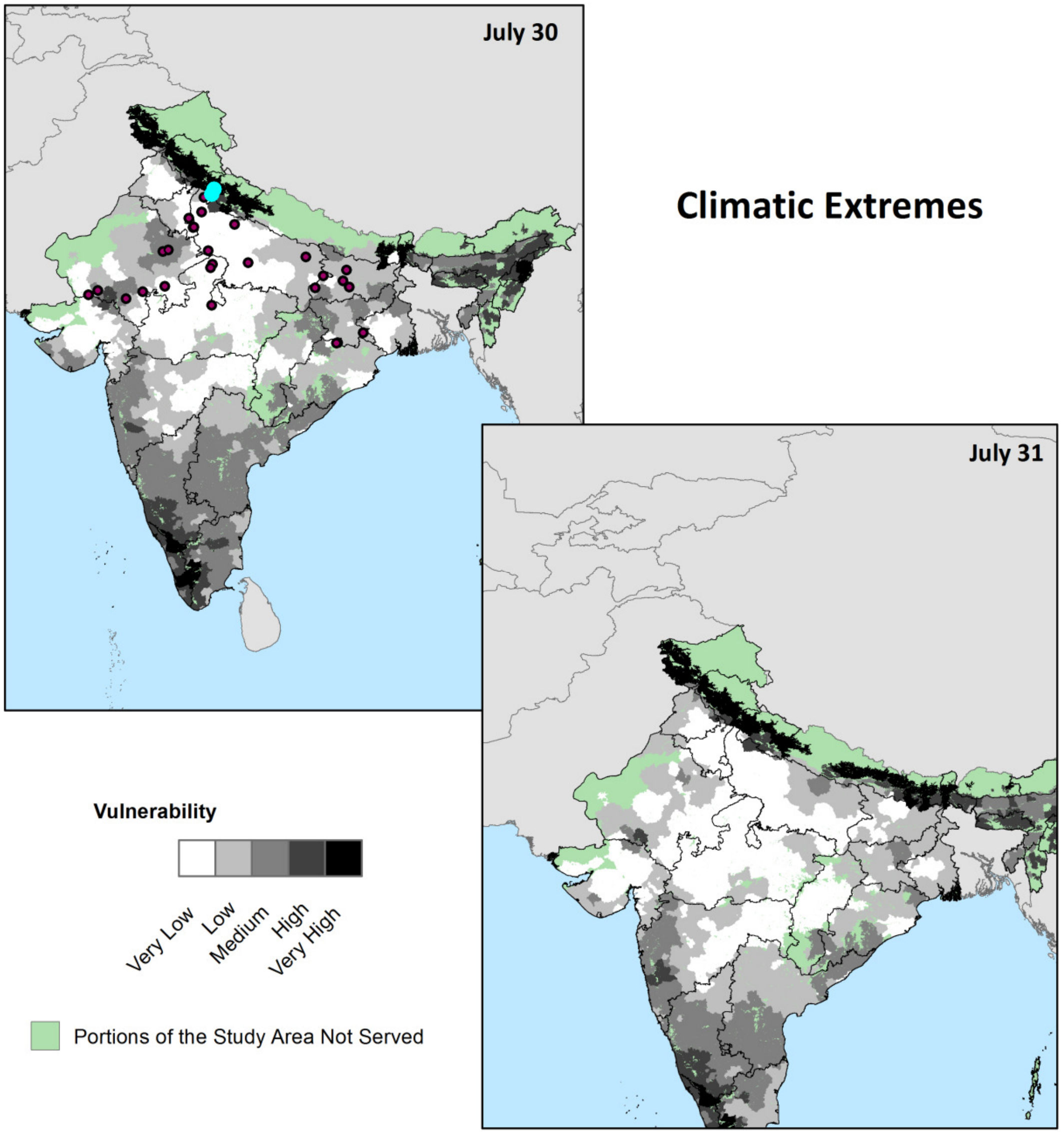


Figure 4.14: Geographic distribution of temperature extremes vulnerability.

Table 4.17: Variable correlations using the Spearman method; values significant at the 0.05 level are marked in bold.

	Betweenness	Land Use	Population	Other Critical Assets	Cyclones	Earthquakes	Landslides	July 30 Temperature Extremes
Betweenness	1	**0.156**	**0.164**	**0.132**	0.093	0.052	-0.033	0.028
Land Use	**0.156**	1	**0.369**	**0.339**	**0.295**	0.032	0.032	**0.116**
Pouplation	**0.164**	**0.369**	1	**0.422**	**0.152**	-0.005	-0.009	**0.095**
Other Critical Assets	**0.132**	**0.339**	**0.422**	1	**0.089**	**0.121**	**0.188**	**0.111**
Cyclones	**0.093**	**0.295**	**0.152**	**0.089**	1	**0.127**	**0.064**	**0.131**
Earthquakes	0.052	0.032	-0.005	**0.121**	**0.127**	1	**0.468**	**0.125**
Landslides	-0.033	0.032	-0.009	**0.188**	**0.064**	**0.468**	1	**0.369**
July 30 Temperature Extremes	0.028	**0.116**	**0.95**	**0.111**	**0.131**	**0.125**	**0.369**	1

Interesting correlations emerge when the variables of interest were correlated with the capacity of the substations (Table 4.18). The capacity of the substations was moderately correlated to its betweenness. This was probably due to larger substations being more important to the overall flow of the graph. Another interesting correlation was the moderate correlation between population and capacity. This may be due to the fact that when substations have a higher capacity, they were usually built to supply a larger population. With this knowledge, one might conclude that increasing the number of substations, but reducing the overall capacity of the substations may decrease the overall vulnerability of the electrical grid.

Table 4.18: Capacity Correlations; bold are significant to the 0.05 significance level.

	Capacity
Capacity	1
Betweenness	**0.339**
Land Use	**0.106**
Population	**0.221**
Other Critical Assets	**0.116**
Earthquakes	**-0.066**
Landslides	**-0.144**
Cyclones	**0.095**
July 30 Temperature Extremes	0.029

Another visualization mechanism for determining the relationship between and among variables used by this research was to identify the vulnerability rankings for the pairwise comparisons for each substation. Three substations were chosen: Biharshariff (Table 4.19), Gwalior (Table 4.20), and Rishikesh (Table 4.21). Biharshariff was chosen because the majority of pairwise comparisons found it to be vulnerable. Gwalior was chosen because the pairwise comparisons were a mix of vulnerable and not vulnerable outcomes. Rishikesh was chosen because most of the pairwise comparisons yielded low vulnerability ranks. The color scheme is the same color scheme described in Figure 4.6. These plots show interesting relationships between the variables. For instance, a large portion of Biharshariff's matrix was highly or very highly vulnerable; however, with the inclusion and combination of other critical assets, natural hazards, and temperature extremes, the matrix changed and the substations had a low or very low vulnerability ranking. Interestingly, the opposite is true for the Rishikesh substation.

Table 4.19: Pairwise comparison rankings for Biharshariff substation.

	Biharshariff					
	Betweenness	Land Use	Population	Other Critical Assets	Natural Hazards	Temperature Extremes
Betweenness	4	5	5	4	3	5
Land Use	5	4	5	4	3	5
Population	5	5	4	4	4	5
Other Critical Assets	1	4	4	1	1	2
Natural Hazards	3	3	4	1	1	2
Temperature Extremes	3	3	3	1	1	1

Table 4.20: Pairwise comparison rankings for Gwalior substation.

	Gwalior					
	Betweenness	Land Use	Population	Other Critical Assets	Natural Hazards	Temperature Extremes
Betweenness	5	4	5	5	4	5
Land Use	4	1	3	2	2	3
Population	5	3	3	4	3	3
Other Critical Assets	5	2	4	2	3	2
Natural Hazards	4	2	3	3	1	1
Temperature Extremes	5	3	3	2	1	1

Table 4.21: Pairwise comparison rankings for Rishikesh substation.

	Rishikesh					
	Betweenness	Land Use	Population	Other Critical Assets	Natural Hazards	Temperature Extremes
Betweenness	1	1	2	1	3	5
Land Use	1	1	2	1	3	5
Population	2	2	2	2	4	5
Other Critical Assets	1	1	2	1	4	5
Natural Hazards	3	3	4	4	2	4
Temperature Extremes	5	5	5	5	4	5

It is also interesting to note the similarities between the hazards data and the betweenness and land use combination results (Figure 4.15). Both natural hazards and betweenness identified the eastern coast as being vulnerable, as well as the eastern part of the country. There are, however, differences, where combining these data was useful. For example, the middle belt of India was where the betweenness and land use combination vulnerabilities were the highest (mainly irrigated cropland area) was not considered vulnerable when looking at only hazards information.

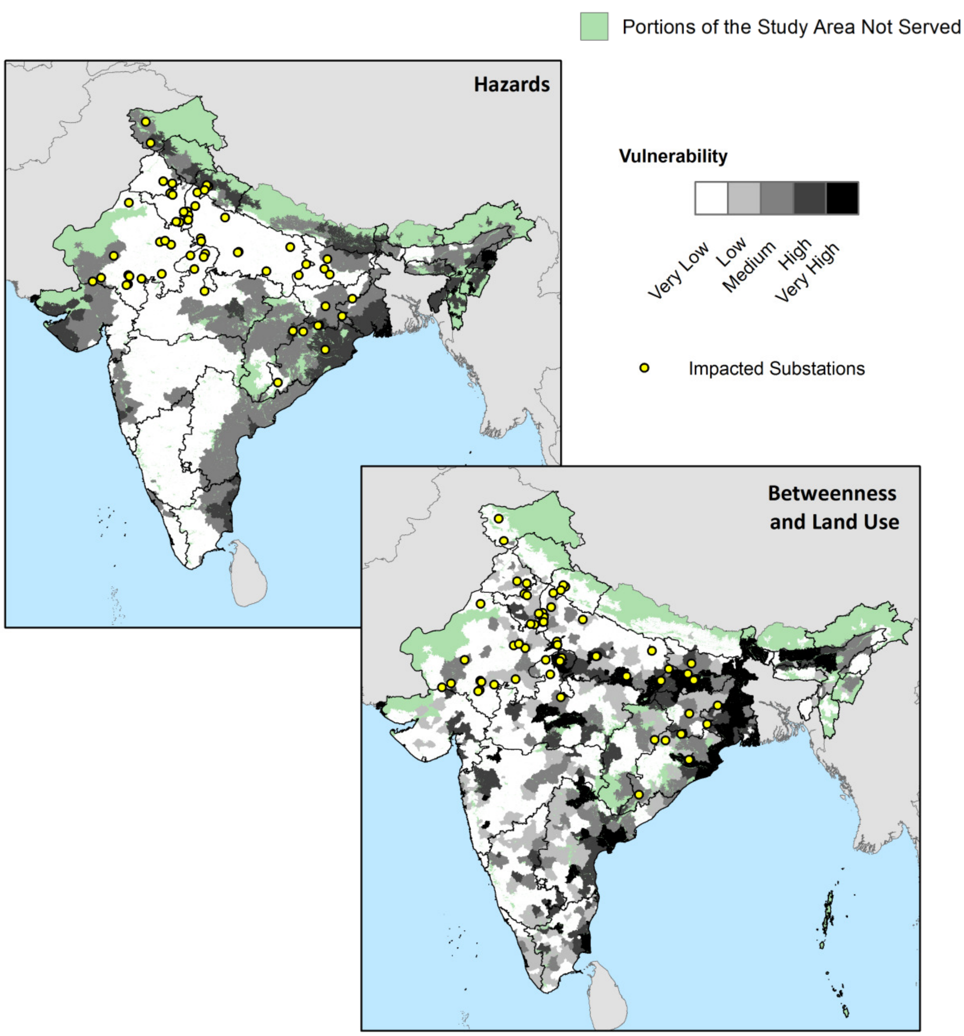


Figure 4.15: Betweenness and land use combination comparison with hazards.

4.5.2. Natural Hazards

In this approach natural hazards were included with land use and betweenness (the best performers from the previous approach) to show general vulnerabilities in the Indian – Nepali – Bhutanese grid system. As stated, natural hazards did not play a role in the July 2012 blackout; however, this information was useful for assessing general substation vulnerability. This dissertation framework also presented methods for situation-based approaches that can take warnings and predict substation vulnerabilities based on those warning areas. Additionally, natural hazards often have seasonal occurrences. For instance, in India the cyclone season begins in April and ends in June, and there is another period of activity from the end of September through the beginning of December. During these months the cyclone activity might be weighted greater to indicate the heightened threat of activity.

Two analyses were conducted for natural hazards: one with even weights and a second with uneven weights. For the weighted analysis, land use was weighted as 0.410, betweenness was weighted as 0.410, and each individual hazard was weighted as 0.059 each. Each analysis had the exact same rankings for the substations impacted on July 30th and 31st as the betweenness and land use even-weighted analysis. This may be due to the fact that there were very few places with very high vulnerabilities to any one of the particular hazards. For cyclones, the majority of the highly and very highly vulnerable locations were located along the eastern coastline, in areas that were already identified as vulnerable. With landslides, there were even fewer locations with high or very high vulnerabilities (located in the northern and eastern-most regions), and those that were highly or very highly vulnerable were areas that have very low vulnerabilities with regards to land use and betweenness. Thus, these very high vulnerabilities would not have as large of an impact on the score. A similar situation evolved with earthquakes, as the majority of the high and very high vulnerabilities occur in areas with low to land use and betweenness vulnerabilities, such as Nepal and western Gujarat.

Figure 4.16 depicts the geographic distribution of the vulnerabilities for both analyses. The results varied little from the geographic distribution of the land use and betweenness evenly weighted analysis depicted in Figure 4.12.

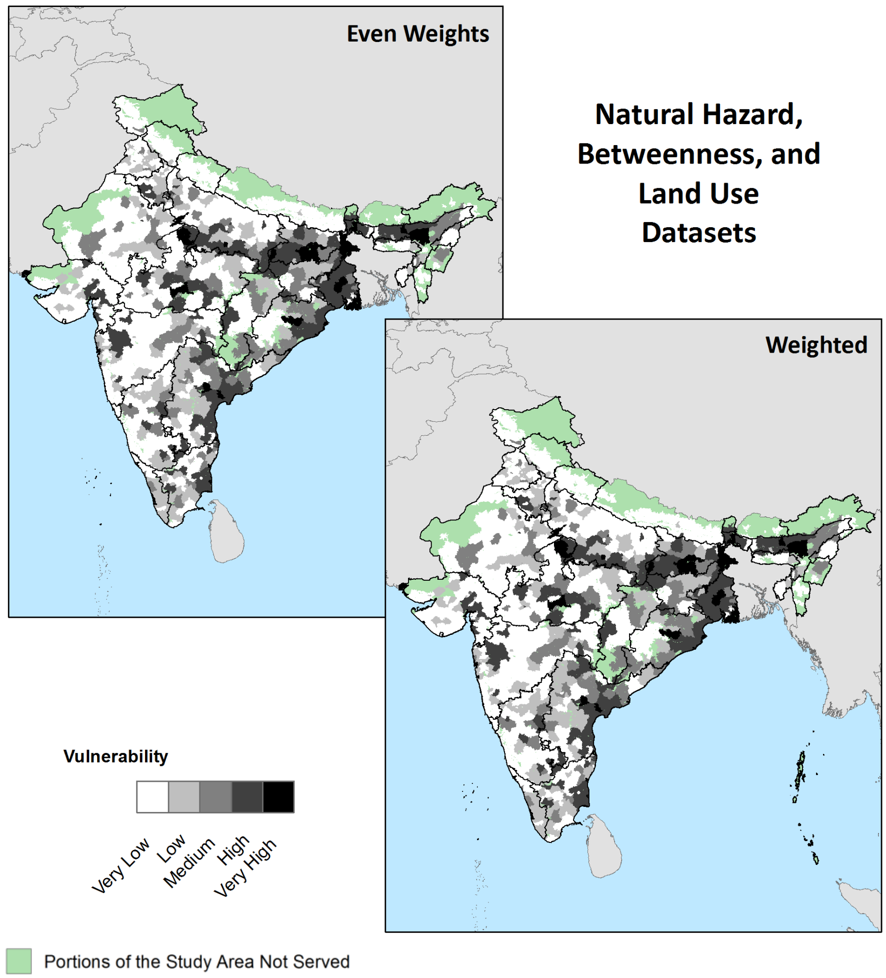


Figure 4.16: Natural hazard analysis vulnerability distribution.

4.5.3. Climatic Extremes

The inclusion of climate change variables in models dealing with critical infrastructure vulnerability is of interest to researchers, as it has not been widely attempted. The approach including climatic extremes built upon the previous analysis including betweenness and land use. An additional combination with the natural hazards data was included for analysis as well.

The inclusion of the temperature extremes data was helpful in identifying additional vulnerable nodes from the July 30th blackout. However, including temperature extremes also identified several nodes (two nodes) that were considered vulnerable in the land use and betweenness analysis as having only moderate vulnerability, while adding three nodes to the high or very high vulnerability ranking. This yielded 51.7% of the substations impacted on July 30th being identified as having high or very high vulnerability (Table 4.22). There was no difference in the results with the addition of the natural hazards datasets, and the performance of the weighted analysis was the same as the land use and betweenness combination (Table 4.23). The weights for the land use, betweenness, and temperature extremes combination were 0.47, 0.47, and 0.06, respectively. With the addition of the natural hazards datasets, the weights were 0.39 for land use, 0.39 for betweenness, 0.08 for temperature extremes, and 0.05 each for the individual natural hazards datasets (cyclones, landslides, and earthquakes).

Table 4.22: Even-weight vulnerability rankings for combinations with temperature extremes.

	Vulneraiblity Rankings - Even Weights			
Substation	**Betweenness**	**Land Use - Betweennness**	**July 30 Temps with Land Use - Betweenness**	**July 30 Temps with Land Use - Betweenness and Natural Hazards**
Agra (PG)	2	2	2	2
Balia	1	3	3	3
Ballabarh (BBMB)	1	2	2	2
Bareilly	1	1	1	1
Bawana	2	2	2	2
Bhinmal	3	3	4	4
Biharshariff	4	5	5	5
Bina (PG)	3	3	3	3
Chittorgarh	1	1	1	1
Gorakhpur (PG)	4	4	4	4
Gwalior (PG)	5	4	4	4
Heerapura	4	4	4	4
Jaipur (PG)	3	3	4	4
Jamshedpur	4	4	4	4
Kanpur (PG)	4	4	4	4
Kota	5	4	4	4
Malanpur	2	5	4	4
Meerut	3	4	3	3
Muzaggarpur	3	4	4	4
Patna	1	4	4	4
PG Bisra	3	3	3	3
Rishikesh	1	1	4	4
Roorkee	2	2	2	2
Rourkela	4	4	3	3
Samaypur	2	2	2	2
Sanchore	1	1	1	1
Sasaram	3	5	5	5
Tehri	1	1	4	4
Udaipur	1	1	3	3
Percent High/Very High	27.60%	44.80%	**51.70%**	**51.70%**

Table 4.23: Weighted vulnerability rankings for combinations with temperature extremes.

	Vulneraiblity Rankings - Weighted			
Substation	**Betweenness**	**Land Use - Betweennness**	**July 30 Temps with Land Use - Betweenness**	**July 30 Temps with Land Use - Betweenness and Natural Hazards**
Agra (PG)	2	2	2	2
Balia	1	3	3	3
Ballabarh (BBMB)	1	2	2	2
Bareilly	1	1	1	1
Bawana	2	2	2	2
Bhinmal	3	3	3	3
Biharshariff	4	5	5	5
Bina (PG)	3	3	3	3
Chittorgarh	1	1	1	1
Gorakhpur (PG)	4	4	4	4
Gwalior (PG)	5	4	4	4
Heerapura	4	4	4	4
Jaipur (PG)	3	3	3	3
Jamshedpur	4	4	4	4
Kanpur (PG)	4	4	4	4
Kota	5	4	4	4
Malanpur	2	5	5	5
Meerut	3	4	4	4
Muzaggarpur	3	4	4	4
Patna	1	4	4	4
PG Bisra	3	3	3	3
Rishikesh	1	1	2	2
Roorkee	2	2	2	2
Rourkela	4	4	4	4
Samaypur	2	2	2	2
Sanchore	1	1	1	1
Sasaram	3	5	5	5
Tehri	1	1	2	2
Udaipur	1	1	1	1
Percent High/Very High	27.60%	**44.80%**	**44.80%**	**44.80%**

The analysis of the substations impacted on July 31, 2012, yielded different results. For all analyses including temperature extremes, the performance was poorer than the land use and betweenness combination. One possible reason for this is the second day of the blackout was indicative of an already degraded system, which may just be indicating additional vulnerabilities in the system. The weighted and the evenly weighted combinations performed exactly the same, identifying 31.1% of the substations impacted on July 31, 2012. The weights for the weighted combinations are the same as the ones used for the July 30th analysis. The first day of the blackout (July 30th) was probably more indicative of what vulnerabilities were present in a system that is operating at a fairly normal capacity.

As stated in previous sections, the accuracy percentage metric helps to explain omission errors, while commission errors in this type of predictive model are difficult to assess. Traditional commission error calculations increase as the number of variables increase (from approximately 5% for betweenness alone to over triple the error for the combination of betweenness and land use). While these commission errors increased, it is not an accurate metric, as for the commission errors indicate incorrect assessments of a cell's classification (in this case incorrectly identifying a cell as vulnerable when it is not). The validation data is only available for one event, and just because a particular substation was not impacted by this particular event, does not mean that it is not vulnerable. While the event likely exploited existing vulnerabilities in the Indian Grid, it may not have exploited them all, which makes it easier to calculate an error of omission versus and error of commission for this dissertation model.

4.6. Statistical Analyses

The majority of these analyses were conducted based on pairwise comparisons and visual interpretation of how each combination performed. How did the results of the combinations results compare with statistical tests? Was the reduction of the variables consistent with a factor analysis? Are any of the models statistically different from using betweenness alone? These questions can be answered with two statistical tests: a Principal Component analysis and an Analysis of Variance (ANOVA).

4.6.1. Principal Component Analysis

Factor analyses indicated where there were redundancies in the data used. A Principal Component Analysis was conducted on the variables of betweenness, land use, additional critical assets, population, natural hazards (earthquakes, landslides, and cyclones), and temperature extremes. There were three resulting components, with the first component explaining 29.8% of the variance, the second explaining 22.8% of the variance, and the third explaining 15.9% of the variance. The first component was most heavily loaded on the temperature extremes, landslide frequencies, and earthquake frequencies, making this the “natural hazard and climate change” component. The second component was most heavily loaded on land use, additional critical assets, and population, making this the “people” component. The final component was most heavily loaded on cyclone frequency and betweenness, making this the “structural demand” component. Table 4.24 is the component matrix. This analysis corresponds to the highest performing combination of variables: land use (“people”), betweenness (“structure”), and climate extremes (“natural hazard and climate change”).

Table 4.24: Component matrix for the Principal Component Analysis.

	Component 1	Component 2	Component 3
Betweenness	-0.103	0.277	0.710
Land Use	-0.196	0.618	-0.306
Add'l Critical Assets	-0.089	0.825	0.077
Population	0.096	0.807	0.213
Temperature Extremes	0.853	-0.130	0.026
Cyclones	-0.056	-0.138	0.773
Landslides	0.917	-0.017	-0.068
Earthquakes	0.865	-0.016	-0.132

4.6.2. ANOVA

For the ANOVA, all of the substations involved in both days of the blackout (July 30 and 31st) were combined for one analysis for a total of 58 substations. There was a statistically significant difference between variable combinations as determined by a one-way ANOVA ($F(8,504) = 3.11$, $p = 0.002$). Of most interest was if the best performing models providing statistically significant differences from using betweenness alone. A Tukey post-hoc test revealed there were no statistically significant differences between any of the best performing combinations (betweenness alone; land use and betweenness; betweenness, land use, and population; betweenness, land use, population, and additional critical assets; betweenness, land use, and temperature extremes; betweenness, land use, and natural hazards; and betweenness, land use, natural hazards, and temperature extremes) and betweenness alone. However, the best performing combination (betweenness, land use, and temperature extremes from July 30th) was the most significant of the not statistically significant combinations (p = 0.247).

4.7. Sensitivity Analyses

When performing any kind of spatial or weighted analysis, a sensitivity analysis is a necessary part of understanding how the model behaves and should be used. This analysis conducted sensitivity analyses with regards to scope and weights.

4.7.1. Weight Sensitivity Analysis

Each combination included weighted and evenly weighted variables. Interestingly, in most cases, the performance of the weighted combinations was either poorer or not better than the evenly weighted combinations. With regards to the two variable combinations, weights reduced the performance with regards to the case study. However, with the three and four-variable combinations, the performance increased, though not by many percentage points with regards to the accuracy metric. Also, this increase in accuracy did not provide the same level of accuracy that the land use – betweenness (best performing) comparisons provided.

When performing analyses with AHP, the weights of the variables must add up to 1. This makes performing traditional sensitivity analyses difficult. To counter this, pairwise comparisons had their weights varied in five weighting schemes. The weighting schemes included weights of 0.1 and 0.9 weights, 0.166 and 0.833 weights, and 0.5 and 0.5 weights (even). This includes large changes in weights (0.5 to 0.1, for instance) and small changes in weights (0.1 to 0.166, for instance). The performance was determined by the accuracy metric (percentage of substations affected by the July 30, 2012, blackout that were ranked as having high or very high vulnerabilities by any given weight and variable combination) (Figure 4.17). For each of the pairwise comparisons, the greatest percentage of substations identified was with the evenly weighted combinations. In some cases, the highly weighted betweenness or land use had relatively high performance; however, this accuracy still did not have greater accuracy than the evenly weighted combinations. This analysis proved that weights did make a difference in the performance of this dissertation model, and the model was affected by changes in these weights. Interestingly, the evenly weighted combinations typically outperformed its weighted counterparts, which would take out much of the subjectivity on the part of the user when trying to determine weights for the model.

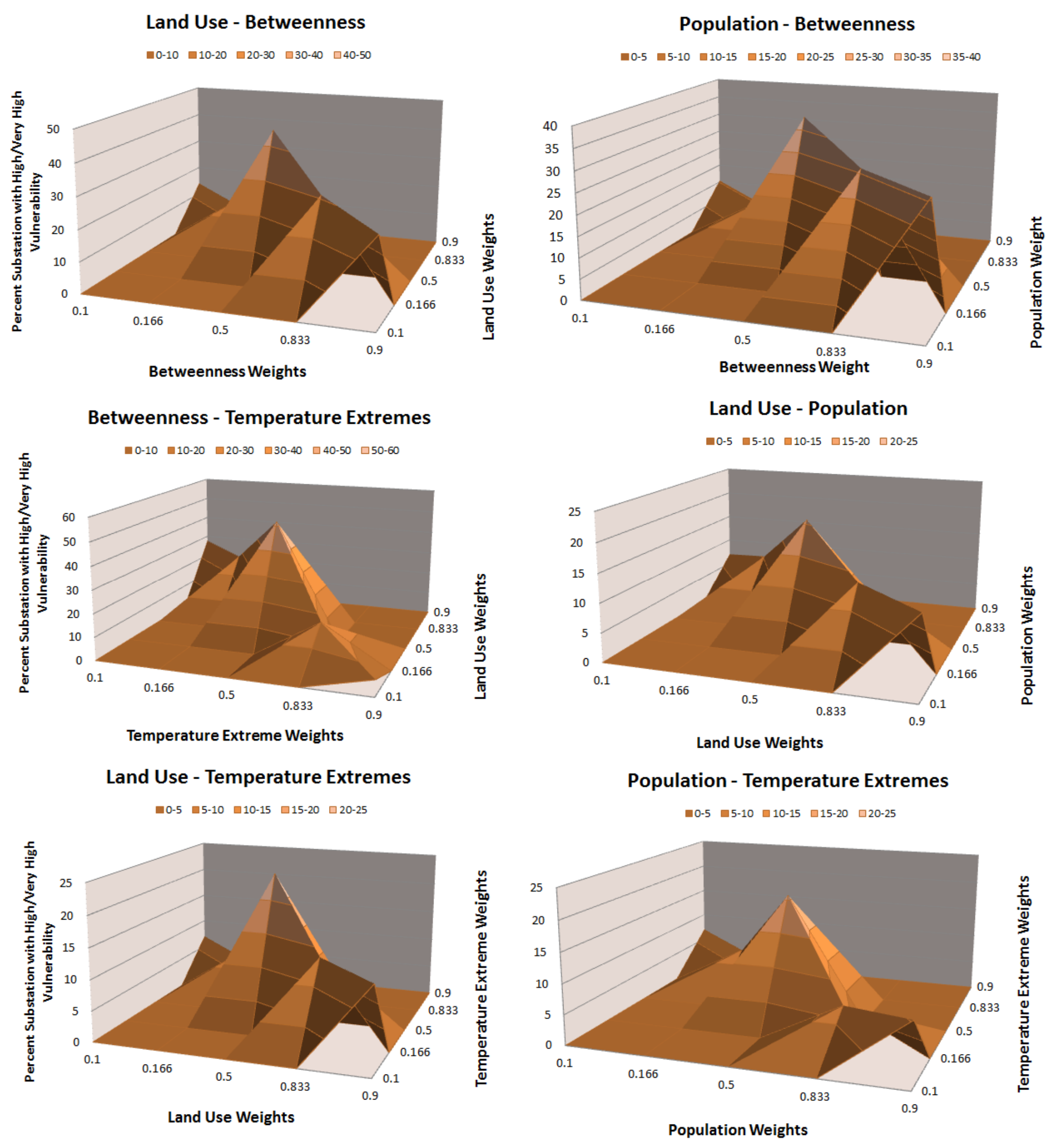


Figure 4.17: Graphic depiction of the variation in performances when weights are varied.

4.7.2. Spatial Sensitivity Analysis

As stated earlier, in addition to assessing the sensitivity of the weights to changes, it was also important to assess changes in scope. To accomplish this, the northern electrical grid region of India was analyzed separately from the entire grid. The grid was resampled to include only the substations and transmission lines within the region; however connections to the other regions were included. All other data was also reanalyzed to see differences in the performance of the best performing combinations. The combinations analyzed were land use and betweenness; population and betweenness; land use, population, and betweenness; land use, betweenness, and natural hazards; Land use, betweenness, and temperature extremes; and land use, betweenness, natural hazards, and temperature extremes. These combinations were evenly weighted, for as shown by the sensitivity analysis in the prior section, evenly weighted combinations had the highest accuracy.

The results of the best performing combinations for July 30, 2012, from the analyses of the entire grid performed similarly overall for only the northern grid. While the overall percentages were similar, the individual rankings of the substation service areas varied. Table 4.25 describes the number and percentage of substations that remained the same. The percentage for the high and very high vulnerability substations that remained the same was calculated by taking the number of substations that remained in either the high or very high vulnerability ranking at both spatial scales and dividing that by the total number of high or very high vulnerability substations found in the northern region only analysis. The amount of substations that retained the same ranking as the entire grid analysis ranged from about 16 percent to about 37 percent (Table 4.25). Substations that had high or very high vulnerability rankings with the entire study area were not considered vulnerable when analyzing the northern region alone. The amount of substations that retained a high or very high vulnerability ranking also varied from 0 percent to about 83 percent (Table 4.25). The best performing combinations were the betweenness and land use combination (52.6%); betweenness, land use, and natural hazards combination (68.4%); and betweenness, land use, and climatic extremes combination (57.9%). All of the substitutions with high and very high vulnerabilities in the betweenness and land use combination were also ranked with high and very high vulnerabilities in the other two better performing combinations. Of course the combination including natural hazards was inclusive of the most substations; however, there was one substation identified as high or very highly vulnerable in the climatic extremes combination that was not identified as such in the natural hazards combination (Table 4.26).

Table 4.25: Consistency between analyzing the Northern Region Grid for India and the entire study area.

		Betweenness	Land Use - Betweenness	Population - Betweenness	Land Use - Population - Betweenness	Natural Hazards - Land Use - Betweenness	Temperature - Land Use - Betweenness	Temperature - Natural Hazards - Land Use - Betweenness
30-Jul	**# of Substation Ranking the Same**	5 (26.0%)	4 (21.0%)	3 (15.8%)	7 (36.8%)	4 (21.0%)	6 (3.16%)	7 (36.8%)
	# of High/Very High Vulnerability Rankings remaing in High/Very High Category	2 (33.3%)	3 (30.0%)	2 (25.0%)	0 (0%)	3 (30.0%)	6 (54.5%)	5 (83.3%)
31-Jul	**# of Substation Ranking the Same**	12 (42.9%)	7 (25.0%)	8 (32.0%)	15 (53.6%)	7 (25.0%)	5 (17.8%)	15 (53.6%)
	# of High/Very High Vulnerability Rankings remaing in High/Very High Category	3 (33.3%)	6 (30.0%)	6 (37.5%)	3 (30.0%)	6 (30.0%)	7 (31.8%)	7 (31.8%)

Table 4.26: Vulnerability rankings for the Northern Region for July 30, 2012.

	Vulneraiblity Rankings - Northern Region						
Substation	Betweenness	Land Use - Betweennness	Population - Betweenness	Land Use - Population - Betweenness	Land Use - Betweenness - Natural Hazards	July 30 Temps with Land Use - Betweenness	July 30 Temps with Land Use - Betweenness and Natural Hazards
Agra (PG)	2	3	2	2	3	2	2
Balia	1	4	4	4	4	4	3
Ballabarh (BBMB)	2	4	3	3	4	3	2
Bareilly	2	4	2	2	4	4	3
Bawana	5	5	5	4	5	5	3
Bhinmal	1	1	1	1	3	3	3
Chittorgarh	2	2	2	2	2	2	2
Gorakhpur (PG)	2	3	3	2	3	4	3
Heerapura	5	5	5	3	4	5	4
Jaipur (PG)	5	5	5	3	4	5	4
Kanpur (PG)	5	5	5	5	5	5	4
Kota	3	3	3	2	3	2	2
Meerut	5	5	5	4	5	5	4
Rishikesh	1	1	3	2	5	4	5
Roorkee	3	4	4	3	4	4	3
Samaypur	4	4	4	2	4	3	2
Sanchore	1	1	1	1	4	1	3
Tehri	1	1	2	2	5	4	5
Udaipur	1	1	2	2	1	3	2
Percent High/Very High	31.60%	**52.60%**	42.10%	21.10%	**68.40%**	**57.90%**	31.60%

Similar results were observed in the July 31, 2012 rankings. Once again, the performance was relatively similar to the performance of the combinations when applied to the entire study area. The overall performance was better than that of the entire grid, identifying more substations as having a high or very high vulnerability. The only difference between July 30 and July 31 was that the natural hazards and climate extremes had the same performance with regards to substations identified as having a high or very high vulnerability (Table 4.27).

Table 4.27: Vulnerability rankings for the Northern Region for July 31, 2012.

Substation	Vulneraiblity Rankings						
	Betweenness	Land Use - Betweennness	Population - Betweenness	Land Use-Population - Betweenness	Land Use - Betweenness - Natural Hazards	July 30 Temps with Land Use - Betweenness	July 31 Temps with Land Use - Betweenness and Natural Hazards
Agra	5	5	5	5	5	5	4
Agra (PG)	2	3	2	2	3	2	2
Allahabad	3	5	5	5	5	5	4
Balia	1	4	4	4	4	4	3
Ballabarh (BBMB)	2	4	3	3	4	4	2
Bhiwadi	5	5	5	4	5	5	4
Dausa	3	3	4	2	4	4	4
Debari	1	1	2	2	1	1	2
Greater Noida	1	4	4	4	4	4	2
Gorakhpur (PG)	2	3	3	2	3	3	3
Jaipur (PG)	5	5	5	3	4	4	4
Jodhpur	4	5	5	4	5	5	4
Kaithal	4	5	4	3	4	4	3
Kaithal (400 kV)	3	4	3	2	4	4	3
Kankroli	1	1	1	1	1	1	2
Kanpur (PG)	5	5	5	5	5	5	4
Kishenpur	4	4	4	2	5	5	5
Koteshwar	1	1	2	2	5	5	5
Maler Kotla	3	4	3	2	4	4	3
Maharani Bagh	3	5	5	5	5	5	4
Mandaula	5	5	5	4	5	5	4
Panki	3	5	4	4	4	4	3
PG	4	5	4	3	4	4	3
PG Kankroli	2	2	3	2	2	2	3
Rewari	3	4	3	2	4	4	3
Samaypur	4	4	4	2	4	4	2
Suratgarh	2	2	3	2	2	2	2
Wagoora (PG)	2	4	2	2	5	5	4
Percent High/Very High	35.80%	**71.40%**	57.10%	35.70%	**78.60%**	**75.00%**	42.90%

The spatial sensitivity analysis identified the importance of the spatial scope with which the analysis is conducted. The vulnerabilities of the substations change with the scope they are being analyzed at (Figure 4.18). It was important to analyze the entire grid to know the vulnerabilities of the system as whole, but local vulnerabilities were also important. Since the Indian electrical grid is a regional grid, comprised of five grids nationally, and the fact that the electrical grids were developed at the state level, make it important to analyze the grids at these local scales. Analyzing the grids at regional scales allow decision-makers to determine what vulnerabilities there are in these systems before they were linked to the other states and regions to create a national grid. This sensitivity analyses demonstrates the need for analyzing electrical grids at a variety of different levels.

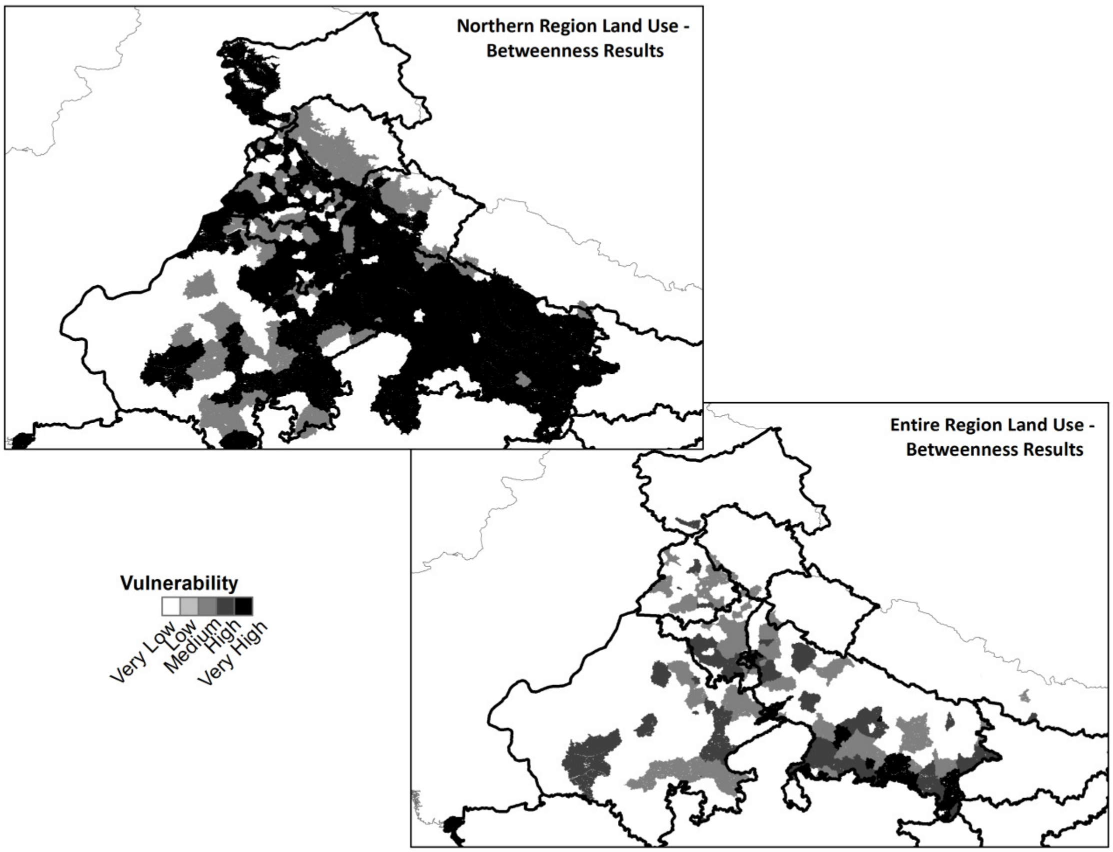


Figure 4.18: Comparison of only looking at the northern region to looking at the entire study area grid system.

4.8. Summary

This research developed and tested a modeling framework for energy grid vulnerability on the Indian grid, utilizing the July 30 and 31, 2012 blackout as a case study. The case study involved three approaches: 1) using the betweenness metric only;

2) using betweenness, land use, population, and additional critical assets; and 3) utilizing land use, betweenness (best two performing variables), and climatic data. Overall, this case study indicated that betweenness was not adequate on its own to identify vulnerabilities in the electrical grid. The best performing combination from the second approach, including pair-wise comparisons, three-way variable comparisons, and all four variables, was the combination of land use and betweenness, evenly weighted. The land use and betweenness combination was able to identify 44.8% and 46.6% of the substations involved in July 30 and 31st blackout as high or very highly vulnerable, versus 27.6% and 24.4% respectively for the betweenness only measure. The addition of temperature extremes improved the performance to 51.7% of the substations involved on the first day (July 30th) as having a high or very high vulnerability ranking. While there were no statistically significant differences between the mean values of the results of each combination, the results do indicate that variables other than just betweenness metric value can help indicate substation vulnerability.

Additional analyses were conducted on the sensitivity of the results to various changes in parameters including sensitivity of the weights and sensitivity of the model to changes in spatial scope. Both analyses show that these changes do make a difference in how the model performs. However, the results did show that not weighting the variables, or leaving the weights even, had the highest performance. The variations in model performance at various spatial scales also indicate the importance of running the model at various spatial scopes. This is true in any network or spatial analysis; when analyses are performed at a different spatial scope, different vulnerabilities may arise than were present when performing the analyses for the entire study area.

A smaller case study was also performed on the Miami area. This additional analysis shows the applicability of the model both to developed countries and less developed countries.

Chapter V

Conclusions

The abundance of blackouts occurring around the world suggests a lack of understanding, modeling, and mitigation for electrical grid vulnerability. Whether the lack of understanding of grid vulnerability is due to a lack of adequate data or a lack of appropriate methods is still unknown. This research developed a new framework for identifying electrical grid vulnerabilities.

Research Objectives

One of the major objectives was to identify the relevant representations of the factors, the relationships between these factors and the appropriate data model to represent them. This research utilized a variety of different characteristics including: population, betweenness, land use, number of critical assets, temperature extremes, and natural hazard frequency (specifically earthquakes, landslides, and cyclones). These factors had raster representations, and were reclassified into raster datasets of similar units (vulnerability units) at the service area level (appropriate unit of analysis). By giving the factors a raster representation and the same units, comparing between the attributes to see which combinations were most useful for identifying vulnerabilities in the electric grid was much easier.

Research Questions

The developed electrical infrastructure vulnerability modeling approach was implemented an used to answer the dissertation research questions. Each question is revisited in this section and the outcomes noted.

1. What are the differences that arise from analyzing network vulnerability using the new integrated framework versus graph theory alone?

There were definitive differences in the performances of the model framework with the addition of geographic variables. The use of betweenness alone in the model enabled approximately 28% and 25% of the substations to be 'correctly' identified as vulnerable on July 30th and July 31st , respectively. The best factor performers overall were the land use and betweenness combination (accounting for approximately 45% of the substations impacted by the blackout) and the land use, betweenness, and climatic extremes (accounting for approximately 50% of the substations impacted) (Table 5.1). Despite the increase in the percentage of substations identified as having high or very high vulnerabilities, none of these differences between betweenness alone and the factor combinations were statistically significant. It is interesting to note that all of the highest performing combination of factors included betweenness, which shows that betweenness is important for identifying vulnerable nodes. However, as stated in the literature, betweenness is misleading when viewed without ancillary information. This is evidenced by the fact that the combinations including these ancillary factors outperformed using betweenness alone.

Table 5.1: Summary of combination results for the 2012 Indian Blackout.

Combination	Percent of Substations with High or Very High Vulnerability on July 30, 2012	Percent of Substations with High or Very High Vulnerability on July 31, 2012
Betweenness	27.60%	24.40%
Land Use - Population	20.70%	17.80%
Land Use - Other Critical Assets	13.80%	15.60%
Land Use - Betweenness	**44.80%**	**46.60%**
Population - Other Critical Assets	10.30%	15.60%
Population - Betweenness	37.90%	42.20%
Other Critical Assets - Betweenness	27.60%	28.90%
Land Use - Population - Other Crtical Assets	3.40%	11.10%
Land Use - Population - Betweenness	20.70%	17.80%
Land Use - Other Critical Assets - Betweenness	17.20%	22.20%
Population - Other Critical Assets - Betweenness	17.20%	24.40%
Betweenness - Land Use - Population - Other Critical Assets	27.60%	24.40%
Betweenness - Land Use - Natural Hazards	**44.80%**	**46.60%**
Land Use - Betweenness - July 30 Temperature Extremes	**51.70%**	31.10%
Land Use - Betweenness - Natural Hazards -July 30 Temperature Extremes	**51.70%**	31.10%

2. What are the vulnerable nodes in Southeastern Asia?

Using the land use and betweenness combination, the most vulnerable substations in the region were concentrated near the urban areas and the Eastern part of the country. There were a total of 90 substations (9.5%) in the study region that were in the very high vulnerability category. Substations near cities such as New Delhi, Chennai, Mumbai, Bangalore, and Hydrabad were also ranked with very high vulnerability. The substations with the highest vulnerability scores were Biharshariff, Malda, Samaguri, Gwalior, Tandikonda, and Kalwa. Importantly, Gwalior, one of the substations involved in the initiating event on both days of the blackout (loss of the Gwalior – Bina transmission line) was one of the substations having the highest vulnerability. There is also a belt from west to east in northern India of vulnerable substations that can be explained by the abundance of irrigated cropland, urban areas, and susceptibility to climate change. There were no highly vulnerable substations in Bhutan or Nepal. This may be due to the fact that neither country serves their entire population with regards to the electrical grid. In fact, Bhutan has been working toward rural electrification, but the terrain makes it difficult to develop the infrastructure (ADB 2012). Additionally, the grid networks in Bhutan and Nepal are not nearly as complex or extensive as India's grid.

There were also geographic differences in the manifestation of vulnerabilities on the landscape. Figure 5.1 shows a comparison of betweenness vulnerability with the highest performing combination: betweenness and land use. The betweenness vulnerability is spottier and less geospatially based, as it takes into account only the structure of the grid. This was a stark contrast to the betweenness and land use combination, where there were clear patches of very high vulnerability corresponding to population centers and areas where irrigated cultivation is prevalent. Interestingly, though the majority of combinations in the scenarios showed Maharani Bagh to be vulnerable, the betweenness alone scenario showed that this substation was not vulnerable. Maharani Bagh is a substation serving the Delhi area, and was one of the substations impacted on July 31, 2012. This was an important difference between the betweenness-only vulnerability rankings and the betweenness and land use combination vulnerability rankings. All of the substations in and around the Delhi metropolitan area were considered to have a low vulnerability in the betweenness-only rankings; however, with the addition of land use, those substations became vulnerable. This was an instance where there would be a major difference in policy decisions between the two models. If policy-makers were only viewing the betweenness-only vulnerability rankings, the Delhi area would not be seen as an area needing additional protection or care; however, in the betweenness and land use rankings, a policy-maker would be more willing to invest in protection mechanisms for the Delhi area.

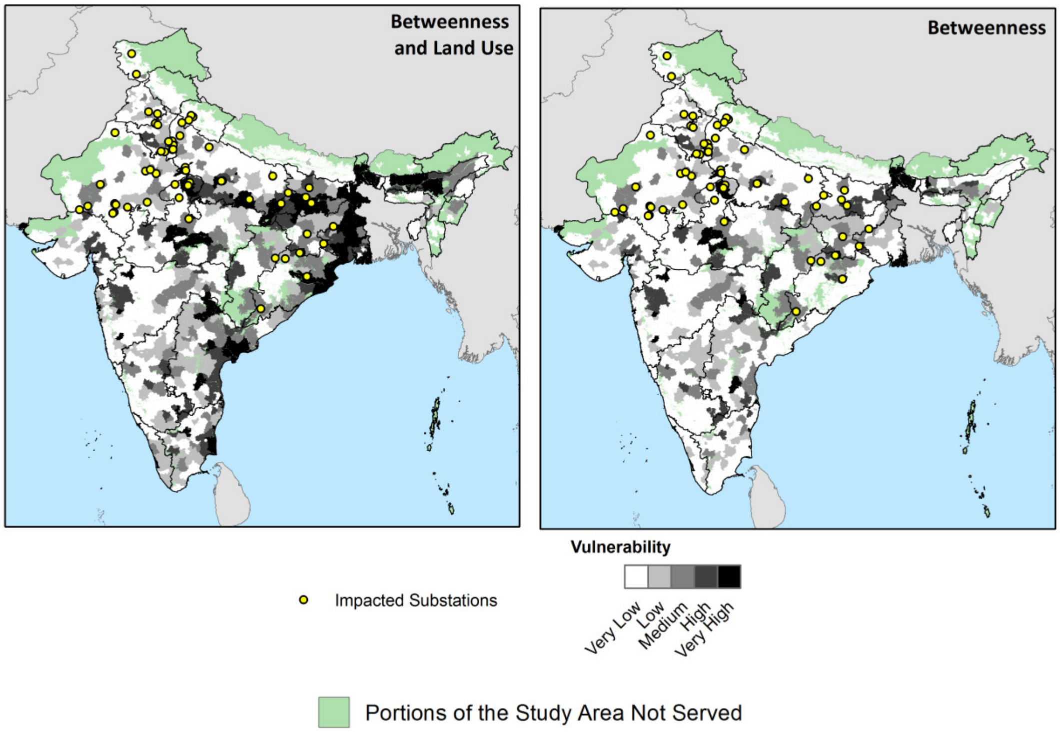


Figure 5.1: Comparison of results from betweenness metric alone analysis and the betweenness and land use combination.

Another instance where this type of holistic framework could be of use to policy makers was with regards to those substations whose vulnerability rankings became high or very high only after the addition of temperature extremes or natural hazard frequency, such as Tehri and Rishikesh. Results from the research in this dissertation may bring critical infrastructure vulnerabilities to climate change to the forefront of policy-makers decisions with regards to upgrading and protecting their assets that are vulnerable in these contexts.

3. What types of evaluation methods are applicable?

Most vulnerability models of electrical infrastructure (and perhaps most vulnerability models in general) are presented without a validation or even suggesting methods for evaluating model performance. This research presented various mechanisms for evaluating this framework, including accuracy assessments, sensitivity analyses, and comparisons to real world events (2012 Indian Blackout).

Contributions

While drawing on prior research from a variety of disciplines, this research makes considerable contributions to both the Geographic Information Science (GIScience) and CIP literatures. These contributions range from new, improved tools to new frameworks. PoDiuM, an advancement in how service areas are calculated in CIP modeling, reduces processing time, and streamlines the data flow in a GIS. PoDiuM improves on past iterations of service area calculation algorithms by improving processing time, while still including important attribute information (supply and demand) and integrating it within a GIS environment. Prior iterations of electrical service area models were calculated outside of a GIS and required laborious transformation before being usable in spatial analyses. PoDiuM also exhibits an acceptance by federal agencies as demonstrated by its use in emergency management modules for helping determine how many customers are impacted by electrical outages by a natural hazard.

Additionally, the integrated framework is the first introduction of geographic variables for critical infrastructure vulnerability assessments. This dissertation research identifies the inability of current graph metrics to identify all the vulnerabilities in electrical infrastructure. Hines et al. (2010) indicated that one must be careful in viewing electrical infrastructure vulnerabilities only in the context of the structure of the graph, as ancillary data (e.g. population, critical facilities other than electrical) are also important in analyzing vulnerabilities in these networks. This research shows the importance of this ancillary data, especially land use, climatic extremes, and natural hazards. Most of these data have been absent from previous analyses of critical infrastructure, and this research indicates that more should be done to incorporate these factors into such analyses for a better understanding of the electric grid. This research also demonstrates the importance of place and geography in analyzing critical infrastructure. The importance of geospatial aspects of critical infrastructure vulnerability is often absent from CIP research. These results, highlighting the importance of these ancillary variables, can help inform policy decisions should they be implemented by government agencies.

This dissertation research also makes contributions to the GIScience literature. This framework improves on previous models, as when implemented, is faster than other models, as it is more cohesive. The user is not required to use various different models in different systems, which can reduce preprocessing and time spent entering resulting data from one model into another model. Additionally, this research also highlights weaknesses in existing frameworks in the CIP literature. One additional advancement of this dissertation research to the GIScience open source community is the development of a framework for developing open source energy network datasets.

This work is also being shared among the critical infrastructure protection communities, from federal agencies to state governments to academics. The framework, software, and data are incredibly sharable and easily transferred from organization to organization. Federal agencies particularly interested in this research include the Department of Defense and the Department of Energy.

Future Work

The research findings hihglight the importance of geospatial information in the analysis of critical infrastructure vulnerabilities, but more research must be conducted. Despite the increase in performance over betweenness alone, the highest accuracy was only 51%, which is not an incredibly high accuracy rate. It is inevitable that this dissertation research did not cover the entire expanse of possible factors that could be included in a geospatial vulnerability model for critical infrastructure. More research should be conducted regarding the inclusion of different factors for understanding the dynamics of critical infrastructure vulnerability. Some examples of factors that might contribute to electrical infrastructure vulnerability include ownership of the substation and socio-cultural factors. Ownership of the substation may be indicative of loyalties, upkeep, and a variety of other concerns of extreme importance to a substation's vulnerability.

This research also indicates the importance of exploring climatic variables related to climate change and their influence on critical infrastructure vulnerabilities. This dissertation research clearly illustrates the importance of temperature extremes on assessing grid vulnerability; however, there are likely a number of additional climatic variables also of importance such as precipitation and drought. Future research should assess these climatic indicators and climate change in relation to grid vulnerability and how best to incorporate them into CI models. This dissertation research offers a beginning framework, but just a beginning.

More extensive work should be done to assess the viability of this research on other study areas including both developed and less developed countries. Also, testing this model in different areas can help determine if there are unique differences in what factors contribute to grid vulnerability in different countries. With access to electric company service area data the PoDiuM model could be validated.

Additionally, research should be conducted on the impact of spatial resolution on the results of this framework as finer spatial resolution data becomes available. Tests utilizing critical infrastructure networks other than electrical networks would also be an interesting addition to the literature to determine if all critical infrastructure networks have similar results.

Improvements could also be made to the framework to extend its usability. Firstly, it could also be modified to make it more dynamic. Right now the model is static and provides a snapshot in time. Making this framework into a dynamic model would enable the users to remove the most vulnerable node and re-run the model to see how the grid vulnerabilities change with the loss of that asset. Another improvement to the framework would be the integration of a risk component. The framework in its current form only addresses substation vulnerability, not risk. This framework could be extended as described in Figure 5.2 to incorporate the concept of risk.

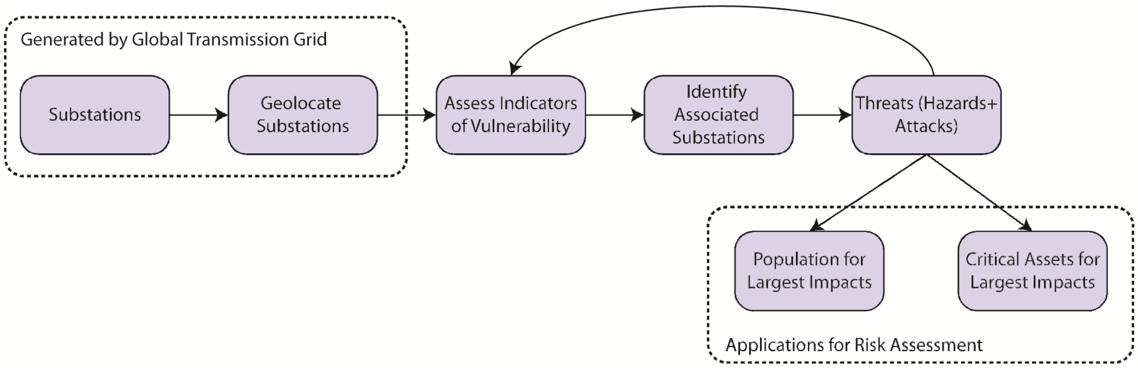


Figure 5.2: Incorporation of risk to the vulnerability framework.

Additional improvements to the tools developed by this research could also be explored in the future. As stated in Chapter 3, PoDiuM could be extended to involve the evolution of the grid to help accommodate current un-served demand or future increases in demand. The AHP tool could also be improved to allow the user more flexibility in how to utilize the weights generated by the pairwise comparisons. Currently, the tool only accommodates a weighted linear combination, or weighted sum approach, but other methods may be exploited in the future.

References

Akabane, K., K. Nara, Y. Mishima, and K. Tsuji. 2002. Optimal Geographical Allocation of Power Quality Control Centers by Voronoi Diagram. *Power Systems Computational Conference,* 24-28 June 2002, Seville. Session 16, Paper 4, 1-6.

Arianos, S., E. Bompard, A. Carbone, and F. Xue, 2009. Power grid vulnerability: A complex network approach. *Chaos,* 19.

Asian Development Bank, 2010. Bhutan-Energy Sector. Available from: http://www.oecd.org/countries/bhutan/47177609.pdf [Accessed 11 April 2013].

Bent, R., R. Djidjeva, B. Hayes, J. Holland, H. Khalsa, S. Linger, M. Mathis, S. Mniszewski, and B. Bush, 2009. Hydra: A Service Oriented Architecture for Science Simulation Integration. Los Alamos National Laboratory, LA-UR-08-7830.

Bhaduri, B. 2013. Foundation level data for the Advanced Grid Modeling Program. *Advanced Grid Modeling Workshop*, 5-6 February 2013, Knoxville, TN.

Bhaduri, B., n.d. LandScan Global Population Databases. Oak Ridge National Laboratory. Available from: http://computing.ornl.gov/cse_home/about/LandScan%20long.pdf [Accessed 11 April 2013].

Birkmann, J., 2006. Chapter 1: Measuring vulnerability to promote disaster-resilient societies: conceptual frameworks and definitions. *In:* J. Birkmann eds. *Measuring Vulnerability to Natural Hazards: Towards Disaster Resilient Societies.* New York: UNU Press, 7-54.

Blaikie, P., T. Cannon, I. Davis, B. Wisner, 1994. *At Risk: Natural Hazards, People's Vulnerability, and Disasters.* London: Routledge.

Boroushaki, S., and J. Malczewski, 2008. Implementing an extension of the analytic hierarchy process using ordered weighted averaging operators with fuzzy quantifiers in ArcGIS. *Computers & Geosciences,* 34: 321-410.

Brown, M., and R. Sedano, 2004. *Electricity Transmission: A Primer.* National Council on Electric Policy.

Bush, B., P. Gigure, J. Holland, S. Linger, A. McCown, M. Salazar, C. Unal, D. Visarraga, K. Werley, R. Fisher, S. Folga, J. A. Kavicky, M. McLamore, E. Pontante, and S. Shamsuddin, 2003. Interdependent Energy Infrastructure Simulation System –IESISS: Version 1.1 Technical Reference Manual. National Infrastructure Simulation and Analysis Center.

Cardona, O.D., M. K. vanAalst, J. Birmann, M. Fordham, G. McGregor, R. Perez, R. S. Pulwarty, E.L.F. Schipper, and B. R. Sinh, 2012. Determinants of risk: exposure and vulnerability. *In:* C.B. Field, V. Barros, T.F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D Mastrandrea, K. J. Mach, G. K. Plattner, S. K. Allen, M. Tignor, and P. M. Midgley eds. *Managing the Risk of Extreme Events and Disasters to Advance Climate Change Adaptation.* New York: Cambridge University Press, 65-108.

Central Intelligence Agency, 2013. Bhutan. *The World Factbook.* Available from: https://www.cia.gov/library/publications/the-world-factbook/geos/bt.html [Accessed 11 April 2013].

Chartrand, G. and P. Zhang, 2005. *Introduction to Graph Theory.* McGraw-Hill Companies.

Chen, Y., J. Yu, K. Shahbaz, and E. Xevi, 2009. A GIS-based Sensitivity Analysis of Multi-Criteria Weights. 18th World IMACS/MODSIM Congress, Cairns, Australia, July 13-17.

Curtin, K. M., 2007. Network Analysis in Geographic Information Science: Review, Assessment, and Projections. *Cartography and Geographic Information Science,* 34(2), 103-111.

Cutter, S. L., 2006. Vulnerability to environmental hazards. *Progress in Human Geography,* 20(4), 529-539.

Davidson, R. A., H. Liu, I. K. Sarpong, P. Sparks, and D. V. Rosowsky, 2003. Electric Power Distribution Systems Performance in Carolina Hurricanes. *Natural Hazard Review,*4(1), 36-45.

Delgado, M. G. and J. B. Sendra, 2004. Sensitivity analysis in multicriteria spatial decision-making: a Review. *Human and Ecological Risk Assessment: An International Journal*, 10(6): 1173-1187.

Demers, M., 2002. *GIS Modeling in Raster.* Wiley, 320 p.

Department of Homeland Security, 2008. *Risk Steering Committee DHS Risk Lexicon.* Available from: http://www.dhs.gov/xlibrary/assets/dhs_risk_lexicon.pdf [Accessed 1 May 2013].

Desmar, U., O. Spatenkova, and K. Virrantaus, 2007. Centrality measures and vulnerability of spatial networks. *ISCRAM*, 201-209.

Desmar, U., O. Spatenkova, and K. Virrantaus, 2008. Identifying Critical Locations in a Spatial Network with Graph Theory. *Transactions in GIS,* 12(1), 61-82.

Drobne, S. and Lisec, A., 2009. Multi-attribute decisionanalysis in GIS: Weighted linear combination and ordered weighted averaging. *Informatica*, 30, 459-474.

Esri, 2012. *Create Thiessen Polygons(Analysis).* ArcGIS 10 Desktop Help. Available from: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00080000001m000000 [Accessed 20 June 2013].

Esri, 2007. *Cost distance algorithm*. ArcGIS 9.2 Desktop Help. Available from: http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Cost_Distance_algorithm [Accessed 8 March 2013].

Fenwick, J. and L. J. Dowell, 1999. Electrical substation service-area estimation using cellular automata: an initial report. *Proceedings of the 1999 ACM symposium on Applied Computing*, February 28 – March 2, San Antonio, TX.

Gnansounou, E., 2008. Assessing the energy vulnerability: Case of industrialised countries. *Energy Policy* 36(10), 3734-3744.

Golden, M. and B. Min, 2012. Theft and loss of electricity in an Indian State. *University of Oxford.* Available from: http://www.theigc.org/sites/default/files/Golden%20and%20Min_Electricity%20theft.pdf [Accessed 11 April 2013].

Goodchild, M., B. Parks, and L. Steyaert, 1993. *Environmental Modeling in GIS.* New York: Oxford University Press.

Goetz, S. J., P. Jantz, and C. A. Jantz, 2009. Connectivity of core habitat in the northeastern United States: parks and protected areas in a landscape context. *Remote Sensing of Environment,* 113, 1421-1429.

Green, M., and R. Flowerdew, 1997. New evidence on the modifiable areal unit problem. In *Spatial Analysis: Modeling in a GIS Environment.* Wiley, 400.

Gross, J. L. and J. Yellen, 2003. *Handbook of Graph Theory.* CRC.

Hayes, B., 2005. *Infrastructure: a field guide to industrial landscape.* W. W. Norton & Company, New York.

Hijmans, R. J., S. E. Cameron, J.L. Para, P.G. Jones, and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25: 1965-1978.

Hines, P., E. Cotilla-Sanchesz, and S. Blumsack, 2010. Do topological models provide good information about electrical infrastructure Vulnerability? *Chaos,* 20.

Holmgren, A., 2007. A Framework for Vulnerability Assessment of the Electric Power Systems. In *Critical Infrastructure*, Springer, 31-55.

Jankowski, P., 1995. Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Systems,*9(3): 251-273.

Kang, S. H., 2010. Tight coupling ArcGIS for simulating inundation depth in densely area. *Natural Hazards and Earth System Sciences*, 10, 1523-1530.

Kim, C. J., and O. B. Obah, 2007. Vulnerability assessment of power grid using graph topological indices. *International Journal of Emerging Electric Power Systems,* 8(6): 1-15.

Koger, W., 2008. Critical infrastructure at risk: A need for a new conceptual approach and extended analytical tools. *Reliability Engineering and System Safety*, 93, 1781-1787.

Kuensel, 2012. Bhutan shares energy through cross border interconnections. *The Santos Republic.* Available from: http://thesantosrepublic.com/2012/09/bhutan-shares-energy-through-cross-border-interconnections/ [Accessed 11 April 2013].

Langford, M. and G. Higgs, 2006. Measuring Potential Access to Primary Healthcare Services: The Influence of Alternative Spatial Representations of Population. *The Professional Geographer,* 58(3), 294-306.

Lemon, D., 2004. A Methodology for the Identification of Critical Locations in Infrastructures. Thesis. Massachusetts Institute of Technology.

Li, J., H. Chen, Z. Liu, and C. Zou, 2010. Integrated GIS/AHP-based flood disaster risk assessment and zonation: a case study of Henan Province, China. *Remote Sensing for Agriculture, Ecosystems, and Hydrology XII*, September 20, Toulouse, France.

Ma, J., N.R. Scott, S.D. DeGloria, A.J. Lembo, 2005. Siting analysis of farm-based centralized anaerobic digester systems for distributed generation using GIS. *Biomass and Bioenergy,* 28(6).

Maguire, D. J., M. Batty, and M.F. Goodchild, 2005. *GIS, Spatial Analysis and Modeling.* ESRI Press, Redlands, 480 pp.

Mahmud, K. 2012. Power grid interconnection of South Asian Region to retain sustainable energy security and figure out the energy scarcity. *Global Journal of Researches in Engineering Electrical and Electronics Engineering,* 12(10): 15-19.

Malczewski, J. 2006. GIS-based Multicriteria Decision Analysis: a Survey of the Literature. *International Journal of Geographical Information Science*, 20(7), 703-726.

Malczewski, J., 2000. On the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches. *Transactions in GIS*, 4(1), 5-22.

Marinoni, O., 2004. Implementation of the analytical hierarchy process with VBA in ArcGIS. *Computers and Geosciences,* 30: 637-646.

McMorrow, D., 2011. *Impacts of Severe Space weather on the Electric Grid.* Available from: http://www.fas.org/irp/agency/dod/jason/spaceweather.pdf [Accessed 11 April 2013].

Memmott, M. 2012. Indignant to India: Blackouts have Millions ‘Fuming.’ *National Public Radio*, 01 August, 2012, International.

Ministry of Power, 2012. Report of the Enquiry Committee on Grid Disturbance in NorthernRegion on 30th July 2012 and in Northern, Eastern and North-Eastern Region on 31st July 2012. Available from: www.powermin.nic.in/pdf/GRID_ENQ_REP_16_8_12.pdf [Accessed 11 April 2013].

Moreira, M.P., C.J. Dupon, and M. Vellasco, 2009. PROMETHEE and Fuzzy PROMETHEE Multicriteria Methods for Ranking Equipment Failure Modes. *Intelligent Systems Applications to Power Systems.* Curitiba, IEEE, 1-6.

National Research Council, 2006. *Facing Hazards and Disasters*. Washington, D.C.: National Academies Press.

Netwon, K. and D. Schirmer, 1997. On the methodology of defining substation spheres of influence within an electric vehicle project framework. *Environmental Systems Research Institute User Conference*, 1997, San Diego.

Ohta, K., G. Kobashi, S. Takano, S. Kagaya, 2007. Analysis of the geographical accessibility of neurosurgical emergency hospitals in Sapporo City using GIS and AHP. *International Journal of Geographic Information Science*, 21(6): 687-698.

Okabe, A., T. Satoh, T. Furuta, A. Suzuki, and K. Okano. 2008. Generalized network Voronoi diagrams: Concepts, computational methods, and applications. *International Journal of Geographic Information Science*, 22(9): 965-994.

Omitaomu, O. and S. Fernandez, 2009. A methodology for enhancing emergency preparedness during extreme weather conditions. Proceedings of the 3rd Annual AFIT/WSU Mid-West CBRNE Symposium, Wright-Peterson Air Force Base, September 22-23.

Parry, M. L., O.F. Canziani, J.P. Palutikof, P. J. van der Linden, and C.E. Hanson, 2007. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007. New York: Cambridge University Press.

Pearson, I. L., 2011. Smart grid cyber security for Europe. *Energy Policy*, 39, 5211-5218.

Philpott, T. and M. Jones, 2012. Did Drought Cause India's Power Outage? Available from: http://www.howmany.org/newsarc/did-drought-cause-indias-power-outage/ [Accessed 11 April 2013].

Quirk, M. D. and S. J. Fernandez, 2005. Infrastructure Robustness for Multiscale Critical Missions. *Journal of Homeland Security and Emergency Management*, 2(2), 1-24.

Rahman, M. R., Z. H. Shi, and C. Chongfa, 2009. Soil erosion hazard evaluation—an integrated use of remote sensing, GIS and statistical approaches with biophysical parameters towards management strategies. *Ecological Modelling,* 220 (13-14): 1724-1734.

Rocco, C. M., J. E. Ramierz-Marquez, and D. E. Salazar, 2011. Some metrics for assessing the vulnerability of complex networks: An application to an electric power system. *Advances in Safety, Reliability, and Risk Management*, CRC Press, 2556-2561.

Saaty, T.L., 2008. Decision making with the analytic hierarchy process. *International Journal of Services Sciences,* 1: 83-98.

Saaty, T.L., 1994. How to make a decision: the analytic hierarchy process. *Interfaces*, 24(6): 19-43.

Saaty, T.L., 1977. A Scaling Method for Priorities in Hierarchical Structures. *Journal of Mathematical Psychology,* 19: 57-68.

Sarkar, D., 2012. Indian trade with energy starved Nepal under trouble. *The Economic Times.* Available from: http://articles.economictimes.indiatimes.com/2012-05-06/news/31597528_1_nepalese-chambers-indo-nepal-nepal-electricity-authority [Accessed 4 April 2013].

Schintler, L.A., S. Gorman, R. Kulkarni, and R. Stough, 2007. Moving from Projection to Resiliency: A Path to Security Critical Infrastructure. In: *Critical Infrastructure*. Springer, 291-307.

Seck, P., 2007. Links between Natural Disasters, Humanitarian Assistance and disaster Risk Reduction. *Human Development Report 2007/2008.*

Sener, S., E. Semer, B. Nas, R. Karaguzel, 2010. Combining AHP with GIS for Landfill Site selection. *Waste Management*, 30(11): 2037-2046.

Sevtsuk, A. and M. Mekonnen, 2012. Urban Network Analysis: A Toolbox for ArcGIS10/10.1. Available from: http://cityform.mit.edu/projects/urban-network-analysis.html [Accessed 11 April 2013].

South Asia Regional Initiative for Energy (SARI/E), 2012. Nepal: Energy Sector Overview. Available from: http://www.sari-energy.org/PageFiles/Countries/Nepal_Energy_detail.asp [Accessed 11 April 2013].

Southern Load Dispatch Center, n.d. Introduction. Available from: http://www.srldc.org/Downloads/SRLDC_Introduction.pdf [Last Accessed 11 April 2013].

Tate, E., 2013. Uncertainty Analysis for a Social Vulnerability Index. *Annals of the Association of American Geographers*, 103(3): 526-543.

Thirumalaivasan, D. and M. Karmegam, 2001. Aquifer Vulnerability Assessment Using Analytic Hierarchy Process and GIS for Upper Palar Watershed. 22nd Asian Conference on Remote Sensing, November 5-9, Singapore.

United States Agency for International Development (USAID), 2009. *South Asia Transmission Utility Regional Network (SATURN): Transmission Utilities Characterization Report.*

United States Government, 2003. *The National Strategy for the Physical Protection of Critical Infrastructure and Key Assets.*

Wang, C., X. Niu, H. Wang, and H. Ye, 2001. A Group Multi-Criteria Decision-Making Method for Power System Skeleton Restoration. *Power and Energy Engineering Conference (APPEEC) Asia-Pacific*, IEEE, 1-4.

Warner, K., C. Ehrhart, A. deSherbinin, S. Adamo, and T. Chai-Onn, 2006. In Search of Shelter: Mapping the Effects of Climate Change on Human Migration and Displacement. *2009 Climate Negotiations.* Bonn, Germany: FSC.

Westervelt, J. D., 2002. Geographic information systems and agent-based modeling. In *Integrating geographic information systems and agent-based modeling techniques for simulating social and ecological processes,* ed. H.R. Gimblett. Oxford: Oxford University Press, 83-103.

White, S., and P. Smyth, 2003. Algorithms for estimating relative importance in networks. *Ninth ACM SIGDD International Conference*, Washington, D.C., 266-275.

Wilbanks, T., S. Fernandez, G. Backus, P. Garcia, K. Jonietz, P. Kirshen, M. Savonis, B. Solecki, and L. Toole, 2012. *Climate Change and Infrastructure, Urban Systems, and Vulnerabilities.*

Xu, Y., J. Sun, J. Zhang, M. Zhang, and X. Lia, 2012. Combining AHP with GIS in synthetic evaluation of environmental suitability for living in China's 35 major cities. *International Journal of Geographic Information Science,* 26(9): 1603-1623.

Young, B., 2009. Dynamic modeling of components on the Electric Grid, ORNL/TM-0000/00, Oak Ridge National Laboratory.

Zio, E. and G. Sansavini, 2011. Component Criticality in Failure Cascade Processes of Network Systems. *Risk Analysis,* 31(8), 1196-1