2014

Developing an integrated spatially explicit scale dependent modeling framework for wind farm suitability assessment in Iowa

Tesfay G. Russell
University of Northern Iowa

DEVELOPING AN INTEGRATED SPATIALLY EXPLICIT SCALE DEPENDENT MODELING FRAMEWORK FOR WIND FARM SUITABILITY ASSESSMENT IN IOWA

An Abstract of a Thesis

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

Tesfay G. Russell

University of Northern Iowa

August 2014

ABSTRACT

Wind energy development is occurring rapidly in the United States due to the drive for energy independence and to mitigate environmental concerns. Wind is a clean, abundant, and entirely renewable source of energy and the most promising source of alternative energy. Among the top wind energy producers in the nation, the state of Iowa is experiencing tremendous growth and it's projected to grow. However, despite the vast development, the contributing factors and spatial decision principles for optimal wind farms placement are not yet well understood. This research advanced an empirical methodology for building site suitability assessment framework for the state of Iowa. Employing the information on existing turbines locations, along with environmental factors (slope, wind power class, elevation, land cover, proximity to neighboring turbine, population density, and distance to transmission line, city, highway, railroad, airport, and river), the study analyzed the contributing factors, their relative importance and regional manifestations. This research developed a spatially explicit scale dependent modeling framework for wind farm suitability assessment based on Iowa context. The framework is based on multiscale empirical module derived from spatial lag regression and machine-learning algorithm coupled with normative component (regulations and policies). The empirical model derived from the spatial lag logistic regression and machine-learning algorithm (Maxent) identified statistically significant factors at different scales. The multiscale spatial lag logistic regression significantly improved modeling compared to standard logistic regression because it accounted for spatial autocorrelation due to the spatial clustering of turbines. Scale's impact on factors importance were examined. At

the Macroscale (statewide) model indicated a good fit to the model with Nagelkerke R square of 0.861. Slope, wind power class, elevation, and distance to transmission line, city, airport, and highway as significant factors that contribute at the Macroscale level. Mesoscale 1 model (regional level) also indicated a good fit with Nagelkerke R squared of 0.801 which identified wind power class, elevation, and distance to transmission line, city, airport, highway and population density as significant factors that impact site suitability at this scale. Mesoscale 2 model (micro-regional) with Nagelkerke R square of 0.794 identified wind power class, elevation, distance to city, river, and transmission line as predictors for site suitability. Microscale model with Nagelkerke R square of 0.784 identified elevation, distance to river and city as significant for predicting suitable site at the scale. As results illustrated, difference in scale of wind development does impact factors importance and changes their significance as well. Overall, elevation, proximity to neighboring turbine, and distance to city are the most important factors that were not impact by scale while the remaining factors displayed scale dependence. Empirical model was coupled with normative factors at a regional scale and the model accuracy of 0.88 indicates a good fit. The framework accounted for the complex technical, environmental, and social constraints to identify suitable sites in Iowa with high accuracy. Ultimately, the framework allows for improved resource characterization to maximize resource utilization. Even though the framework developed is in the context of Iowa, it can be modified for other geographic locations.

DEVELOPING AN INTEGRATED SPATIALLY EXPLICIT SCALE DEPENDENT MODELING FRAMEWORK FOR WIND FARM SUITABILITY ASSESSMENT IN IOWA

A Thesis

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

Tesfay G. Russell

University of Northern Iowa

August 2014

This Study by: Tesfay G. Russell

Entitled: DEVELOPING AN INTEGRATED SPATIALLY EXPLICIT SCALE DEPENDENT MODELING FRAMEWORK FOR WIND FARM SUITABILITY ASSESSMENT IN IOWA

has been approved as meeting the thesis requirement for the

Degree of Master of Arts in Geography

Date	Dr. Andrey N. Petrov, Chair, Thesis Committee
Date	Dr. Bingquing Liang, Thesis Committee Member
Date	Dr. Patrick P. Pease, Thesis Committee Member
Date	Dr. Michael J. Licari, Dean, Graduate College

ACKNOWLEDGMENTS

I would like to express the deepest gratitude of appreciation to the National Science Foundation under Grant Number EPS-1101284 and the Iowa Space Grant Consortium for all of the funding they provided me in order to make this thesis possible.

I am most grateful to my committee chair, Dr. Andrey N. Petrov, who showed me the road and helped to get started on the path to this degree. His encouragement and advice throughout my graduate studies have been extremely helpful. I would like to thank my committee members, Dr. Patrick P. Pease and Dr. Bingquing Liang for their suggestions to improve the research and being accessible at all times.

I would also like to acknowledge and thank Mrs. Donna Jacob for being available at all times and the tremendous work proofreading my thesis. Thank you.

Most importantly I want to say thank you to my family who have given me their unequivocal support and encouragement in every endeavor including my graduate studies.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

To reduce dependence on fossil fuels and mitigate environmental concerns, wind energy development has accelerated in the last decade (American Wind Energy Association [AWEA], 2012a; 2012b; Aydin, Kentel, & Duzgun, 2010; U.S. Department of Energy [DOE], 2008). Wind energy development offers positive impacts in terms of greenhouse gas (GHG) reductions, water conservation, and energy security (DOE, 2008). Wind is a clean, abundant and entirely renewable source of energy and the most promising source of alternative energy in the state of Iowa (AWEA, 2012a). According to the National Renewable Energy Laboratory (NREL) resource assessment, Iowa's wind resource is 7th best in the nation, and it has yet to be fully harnessed (AWEA, 2012b). As untapped wind energy resources coupled with improved wind energy technology, the cost of wind power per kilowatt hour (kWh) is making wind energy competitive with other electricity producing sources (Blair, Hand, Short, & Sullivan, 2008; DOE, 2008; AWEA, 2012a; 2012b).

In 2011, 24.5% of Iowa's electricity was generated from wind which ranks first in the nation in this category (AWEA, 2012a). This production is an equivalent of 1.3 million average Iowan homes being powered by wind energy (AWEA, 2012a). Iowa is also first in wind production capacity per sq. mile, third in wind power installed per capita, and third in the number (3,198) of utility scale wind turbines installed (AWEA, 2012a). Tremendous wind resources being harnessed every year, Iowa will be among the leaders in the nations in wind energy production (AWEA, 2012a). According to NREL, 75% of Iowa is suitable for wind energy development with estimated wind resources of 570, 000 megawatt hours (DOE, 2008). Wind industry is well established including manufacturing, transporting, and installing wind turbines (Halvatzis & Keyser, 2013). Iowa's wind industry is projected to grow which makes the state an ideal study site for empirically driven approach to assess site suitability.

However, contributing factors and spatial decision principles for optimal wind farm placement are not yet well-understood. Optimal placement based on a spatially integrated nuanced predictive model in Iowa context is not developed. Therefore, this research advances an empirical approach to analyze contributing factors, their relative importance, scale dependency and regional manifestations. Recent studies from various regions try to fill this knowledge gap demonstrating the increased importance of determining the optimal placement of wind turbines to maximize resource use (Grady, Hussaini, & Abdullah, 2005; Mann, Lant, & Schoof, 2012; Marmidis, Lazarou, & Pyrgioti, 2008; Mosetti, Poloni, & Diviacco, 1994). Despite advancement in the studies of optimal placement, there has been a gap in the research of scale dependence, factors scale manifestation, and the use of empirically driven methods. However, results from multiscale analysis can very across geographic scales due to the Modifiable Areal Unit Problem (MAUP) which refers to the potential result inaccuracies due to scale (Dark & Bram, 2007; Openshaw, 1983, 1984; Openshaw & Taylor, 1979).

This study has important intellectual merit because it reveals new evidence that can be used to identify suitable locations for wind energy development. It can also be used to evaluate wind projects and make science based recommendations to developers, policy makers, industry leaders and other stakeholders. The model framework developed in this study aims to improve resource characterization and maximize resource utilization in Iowa. In addition, the research identifies the contributing factors and their relative importance, and scale dependence using existing wind turbines in the state.

1.1 Research Goal and Objectives

The goal of this research is to develop an integrated spatially explicit scale dependent modeling framework to assess wind farm site suitability in Iowa. Therefore filling the knowledge gap about contributing factors, their relative importance, scale dependency and regional manifestations in Iowa. Thus, the following questions and objectives will be addressed in this research.

Objectives

1. Using locations of existing wind turbines to identify contributing factors and their relative importance in Iowa.
2. Identifying differences among contributing factors and their relative importance at the different scales.
3. Developing spatially explicit scale dependent modeling framework for wind farm site suitability assessment.

1.2 Thesis Structure

Chapter 2 of this thesis provides a literature review that will reveal wind energy development in the United States and the state of Iowa. Also, suitable site identification methods, suitability factors for wind energy development, and Modifiable Areal Unit Problem (MAUP) in spatial analysis are outlined. Chapter 3 provides detailed description of the environmental and physical characteristics of the study area along with spatial dataset used in the analysis. Also, this chapter describes analysis methodologies based on multiscale empirical models derived from spatial lag logistic regression and machine-learning algorithm coupled with normative component. Chapter 4 presents the results. Chapter 5 provides discussion and overarching interpretation of the results along with case study application and testing of the modeling framework. Chapter 6 discusses conclusions, limitations and future directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Wind Energy Development in the United States

Wind power is the fastest growing renewable energy source in the world with an annual growth rate of approximately 35% (Sathyajith & Philip, 2011). The United States has over 40,000 megawatts (MW) of installed wind power capacity at the end of 2010 and as of 2011, the Unite Stats is only second to China in cumulative wind energy capacity installed at 43.5 GW (DOE, 2011; 2012). In 2010, wind supplied 94.7 billion kilowatt-hours (kWh) of electricity which is only 2.3 percent of national consumption but a five-fold increase since 2005 (DOE, 2008).

Many reports and studies predict and verify that the United States has the potential to generate 20% of its electricity from wind by 2030 (DOE, 2008; Hand et al., 2008). As untapped wind energy resources are identified coupled with turbine technology advancement, the cost of wind energy production per kilowatt hour (kWh) has been declining making it competitive with other electric producing sources (DOE, 2008). Current installations are spread among more than 25 states and the vast majority of capacity is concentrated in Texas, Iowa, and California. Many states use specific polices to encourage renewable energy development but these policies can vary widely among states both in scope and implementation. However, wind development is heavily dependent on federal policies that incentivize renewable energy investment. At the federal level, the Renewable Energy Production Tax Credit (PTC) has helped to make much more cost effective (AWEA, 2012a; 2012b). Under this policy, producers receive

2.2 centers per kWh for renewable energy including wind power. In addition, states have implemented renewable energy standards (RPS) to open the energy market at the state level in order to incentivize renewable energy producers.

2.2 Wind Energy Development in Iowa

Iowa was one of the earliest states to have renewable portfolio standards enacted by the state legislature in the early 1983 (Hurlbut & NREL, 2008). Iowa's installed wind capacity has been growing steadily over the past decade due to federal and state policies. In 2012, Iowa was first in per capita and second in total production of wind energy. The increased integration of wind energy into state's energy portfolio has made Iowa ideal location for increased investment in wind energy development (AWEA, 2012a). Actually, Iowa possesses abundant wind resources and among the leading states in wind energy production and manufacturing (AWEA, 2012b). According to NREL, 75% of Iowa is suitable for wind energy development with estimated wind resource of 570, 000 Megawatts. As of 2012, Iowa produced 24.5% of electricity from wind energy and it was the first state in the nation to reach this threshold (AWEA, 2012a; 2012b; DOE, 2008).

As stated earlier, the primary benefits of expanding wind energy are reduction in CO_2 emissions since turbines do not release atmospheric emission, generating domestic generating power, renewable source, and cost competitiveness among other electric producing energy sources (DOE, 2008). However, various studies reveal that wind energy has an impact on both society and ecology but the advantages outweigh the disadvantages (Acker, Williams, Duque, Brummels, & Buechler, 2007; Griffiths &

Dushenko, 2011; Rodman & Meentemeyer, 2006; Van Hoesen & Letendre, 2010). The physical, socio-economic, technical, and environmental factors need to be thoroughly assessed to maximize the wind energy resource potential (DOE, 2008).

This can be done by establishing a framework to evaluate physical, environmental and human impact factors to assess suitability (Menz & Vachon, 2006; Wiser, Namovicz, Gielecki, & Smieth, 2007). The framework can be applied to other regions, and the information can be used by developers and policy makers to predict the extent wind energy can be developed based on land availability (Rodman & Meentemeyer, 2006). As wind energy production rapidly expands, the social and landscape factors are shifting toward mutually beneficial partnership between communities and wind energy developers (Slattery, Johnson, Swofford & Pasqualetti, 2012; Sowers, 2006; Swofford & Slattery, 2010). In Iowa, there is tremendous support for wind energy development from landowners, state and federal legislatures (Sowers, 2006).

2.3 Wind Resource Assessment and Site Identification

Utility scale wind resource assessment is a multiphase process that incorporates the environmental, physical, and socioeconomics constraints to maximize production and efficiency while keeping the cost low (New York State Energy Research and Development Authority [NYSERDA], 2010). Figure 1 summarizes the utility scale wind energy project lifecycle which contains five main areas: wind resource assessment, permitting, financing, constructing, and operation and decommissioning (NYSERDA, 2010).

The first stage of the wind resource assessment process identifies potential development sites. Site selection requires comprehensive assessment on factors to determine the suitability of the area. The complex array of critical factors is drawn from physical, demographical, economic, and environmental along with policies and regulations which are all major components of site assessment (Bennui, Rattanamanee, Puetpaiboon, Phukpattaranont, & Chetpattananondh, 2007). So, identifying suitable site for wind energy development is a gradual multi-stage process. Furthermore, correctly estimating the potential energy availability is essential to the successful development and economic viability of the wind farm project (AWS Scientific, Inc., 1997). Ultimately, the overall site feasibility process should include comprehensive site study consisting of the social, environmental, economic and human impacts.

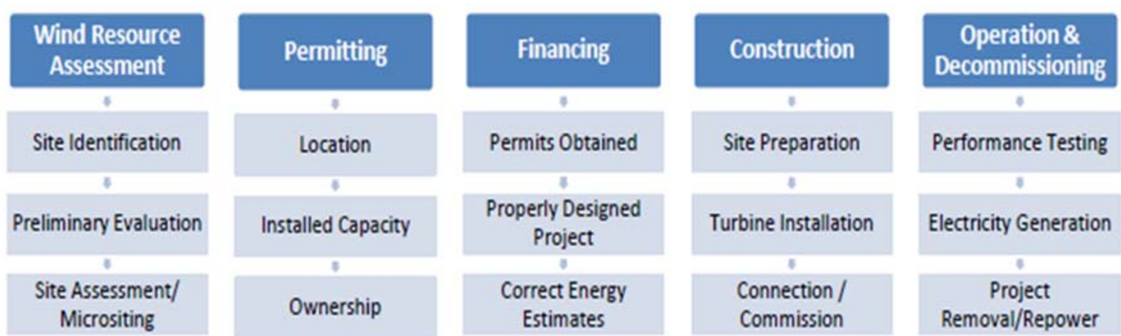

Figure 1: Utility scale wind energy project lifecycle. (NYSERDA, 2010)

Site Identification

As a first step, a geographic database that contain terrain data, project boundary, water bodies, land cover data, environmental sensitive areas, transmission lines, buildings, pipelines, exclusion areas, road networks, permit requirements, and airport restriction should be compiled in a Geographic Information Systems (NYSERDA, 2010). Geographic Information Systems (GIS) has become an integral part of the site selection process and a very useful tool to analyze and determine the most effective locations to install monitoring towers for wind resource assessment and placing turbines.

The second stage involves characterization of the wind resource at the site. Monitoring towers are installed, and the primary objective for monitoring wind is: (i) determining or verifying whether sufficient wind resource exists at the site to justify project continuation, (ii) Enables wind resource comparisons with various potential sites, (iii) estimating the potential economic viability of the project.

The third stage of wind resource assessment includes a detailed evaluation of wind resource at the chosen site. At this point, wind resource is characterized as accurately as possible at all relevant temporal and spatial scales with the primary objective being an accurate estimation of energy production and the optimal placement of turbine within the project area. Characterizing the observed wind resource and data validation process is analyzed to generate an estimated hub-height wind resource. Estimating a wind turbine's energy production often requires extrapolating the measured data to the turbine hub height and analysis of information about the site including the local meteorology, topography, and land cover.

2.4 Iowa Site Suitability Factors

The factors and criterion used for wind farm siting from previous studies are summarized in Table 1. Environmental, technical, and social constraints are included in the comprehensive site study. The environmental and economic assessments are essential for successful wind energy development (Griffiths & Dushenko 2011; Josimovic & Pucar 2010; Leung & Yang 2011). Environmental sensitive areas consist of wetland, wildlife preserves, and federal land (Acker et al., 2007). Developers should consider local, state and federal regulations in order to comply with the law.

Table 1: *Factors and criteria used for wind farm siting in past studies (a)*

Study	Year	Wind Speed	Elevation	Slope	LU-LC: Forests	LU-LC : Protected Areas	Urban Areas	Population	Highways	Airports	Power Grid
Ouammi et al.	2011			<10%		constraint	1000 m buffer		within 1500 m	2500 m buffer	within 1000 m
Janke	2010	weighted			weighted	weighted	weighted	weighted	weighted		weighted
Mari et al.	2010	4 m/s				constraint	constraint				
Sliz-Szkliniarz & Vogt	2010	5 m/s			200 m buffer	500 m buffer	500 m buffer		100 m buffer	3000 m buffer	200 m buffer
Tegou et al.	2010	4 m/s		<30%		500 m buffer	1000 m buffer		within 10,000 m	constraint	
Van Hoesen & Letendre	2010	weighted	600-1050 m	<60°	weighted						
Aydin et al.	2009					250-1000 m buffer	2000 m buffer			2500 m buffer	
Simao et al.	2009				weighted	weighted	weighted		weighted		
Yue & Yang	2009	5 m/s			250 m buffer	250 m buffer	500 m buffer		constraint		
Dutra & Szklo	2008	6 m/s				constraint					
Lejeune & Feltz	2008				200 m buffer	100-2000 m buffer	constraint		weighted	5000 m buffer	150 m buffer
Bennui et al.	2007	weighted	<200 m	<15%		2000 m buffer	2500 m buffer		500 m buffer	3000 m buffer	
Bravo et al.	2007			<10%	constraint	constraint	constraint				
Byrne et al.	2007	4 m/s									
Ramierez-Rosado et al.	2007	weighted	weighted	weighted	500 m buffer	500 m buffer	500 m buffer		weighted	constraint	weighted
Acker et al.	2006	weighted		<20%	50% constraint	constraint	constraint			3000 m buffer	within 10 mi
Nguyen	2006	weighted	weighted		500 m buffer	500 m buffer	2000 m buffer	weighted	100 m buffer	2500 m buffer	
Hansen	2005	weighted			300 m buffer	300 m buffer	500 m buffer		150 m buffer	5000 m buffer	200 m buffer
Rodman & Meentemeyer	2005	4.5 m/s	weighted	<40°	constraint	constraint	constraint				
Ramachandra & Shruthi	2004	weighted									
Krewitt & Nitsch	2003	4 m/s			500 m buffer	500 m buffer	500 m buffer		500 m buffer		
Baban & Parry	2001	5 m/s	constraint	<10%	500 m buffer	1000 m buffer	2000 m buffer		within 10,000 m		within 10,000 m
Hillring & Krieg	1998	weighted			constraint	constraint	constraint				
Voivontas et al.	1998	6 m/s	<1000 m	<60%	constraint	2000 m buffer	1000 m buffer			2500 m buffer	
Criteria Usage in Studies		79%	29%	42%	67%	88%	83%	8%	54%	46%	33%

(a)Values indicate the constraints for wind turbine placement. The term weighted indicates if the study used a custom scale to weight criterion layers in the suitability study.

However, procedures and regulations differ state by state which makes it very difficult for wind energy developers to speed up project developments (Geiszler, Koppel & Gunther, 2013). There is also a push to make a standardized system and singular regulator body to make the process swift and efficient. The benefit of this is to set up a "one-stop" permit process to increase wind energy development (Portman, Duff, Koppel, Reisert, & Higgins, 2009).

The most important factor that determines site suitability is wind speed. Based on the extensive literature review, 79% of studies have identified wind speed as a critical factor. Other important criterion identified are elevation, slope, land cover, protected areas, urban area and distance to airport, power grid, and highway.

Wind Power Class (Wind Speed)

The primary factor for wind farm development is the availability of good wind resources which is essential for the economic viability of the project (AWS Scientific, Inc., 1997). Wind is intermittent and varies over time and over vertical and horizontal height. The vertical profiles of wind speed and wind direction vary by location, and the vertical profile of wind speed is strongly dependent on landscape's roughness (Toke, Breukers, & Wolsink, 2008). The average annual data at referenced height above the ground level has horizontal spatial resolution of 2.5 km. The extrapolation of the wind speed at different heights is affected by the local conditions and significantly influences the wind shear in the first 200-300 m (Grassi, Chokani & Abhari, 2012). Wind shear is known as the variation in speed with height. In most places, wind speed increased with increasing height and this is refer to as wind shear. The shear is typically measured using simultaneous speed measurements at more than one height on a mast (Grassi et al., 2012). Extreme wind shear can cause extra wear and tear on turbine components as well as losses in energy production so there is a minimum and maxim ideal wind speed rate to operate turbines.

Accurate wind resource assessment requires collecting precise wind data over a period of 1- 2 years at the proposed site (AWS Scientific, Inc., 1997). Estimates of the wind resource are expressed in wind power classes ranging from class 1 to class 7, with each class representing a range of mean wind power density or equivalent mean wind speed at specified height above the ground (Table 2). Wind power class is defined by the upper limits of mean wind power density and mean wind speed at 10 m (33 ft.) and 50 m (164 ft.) above ground level (AWS Scientific, Inc., 1997). Grid cells designated as class 4 or greater are generally considered very good wind conditions, but due to improved technology and taller turbine towers, even wind power in class 3 are suitable for utility scale development.

NREL has established a composite of the best available data to classify wind power class which represents an annual average wind speed at 10 meters above the surface and vertically extrapolated wind speed to 50 meters based on the 1/7 power law(Table 2). The annual average wind speed is the primary basis mentioned as a way to rate or rank wind project sites (NYSERDA, 2010). Most wind project development are occurring at sites with a mean wind speed of over 6.5 m/s at 50 m hub height

(NYSERDA, 2010). The mean wind speed is based on Rayleigh speed distribution of equivalent mean power density (Table 2). Each wind power class spans two power densities which means wind power density ranges between 150 W/m² and 200 W/m² in this case for Wind Power Class = 3. However, to determine or verify whether sufficient wind resources exists within the area requires accurate, reliable and multi-year climatic data (AWS Scientific, Inc., 1997). Wind with reasonable speed is not the primary determinate of wind energy development for practical and economic reasons so the potential site has to be thoroughly investigated and the wind speed profile and density accurately calculated (Mohandes, Rehman, & Rahman, 2011).

Table 2: *NREL Wind Power Class Classification*

Wind power classes at 10 m and 50 m elevation (a)				
Wind Power class	10 m		50 m	
	Wind speed (m/s)	Power Density (W/m2)	Wind speed (m/s)	Power Density (W/m2)
1	0-4.4	0-100	1-5.6	0-200
2	4.4-5.1	100-150	5.6-6.4	200-300
3	5.1-5.6	150-200	6.4-7.0	300-400
4	5.6-6.0	200-250	7.0-7.5	400-500
5	6.0-6.4	250-300	7.5-8.0	500-600
6	6.4-7.0	300-400	8.0-8.8	600-800
7	7.0-9.4	400-1,000	8.8-11.9	800-2,000

(a) Vertical extrapolation of wind speed based on the 1/7 power law.
(b) Mean wind speed on Rayleigh speed distribution of equivalent mean wind power density.

There are three important factors that determine wind speed resource estimate and the degree of certainty; (1) the abundance and quality of wind data; (2) complexity of the terrain; (3) geographical variability of wind resource (AWS Scientific, Inc., 1997). It's assumed that as height increases, wind speed increases; and concrete measurements of wind speed can be obtained by utilizing clustering algorithm based on neuron-fuzzy method to create a profile up to 100 meters based on the knowledge at highs of 10, 20, 30 meters (Mohandes et al., 2011). Wind speed varies depending on the specific site and in the case of Greece wind speed of 4m/s was identified as the minimum required for turbine installation (Tegou, Polatidis & Haralambopoulos, 2010). Acker et al. (2007) classified wind scale ranging from poor (<5.5 m/s), marginal (5.5-6.3 m/s), fair (6.3-7.0 m/s), good (7.0-7.5 m/s), to excellent (>7.5 m/s) in their wind resource assessment for the state of Arizona. Suitable wind speed depends on the geographical context of the site and acceptable standards vary from state to state. Iowa being located in the Midwest and possessing the 7th most wind resource in the nation (the minimum standard of suitable wind speed of 6.4 m/s at 50 meters) is considered exploitable.

Elevation

Studies have identified elevation as a constraint especially in mountainous regions due to complex terrain where topographic influences are strong (Bennui et al., 2007). One method is to measure the wind at numerous locations within the wind project area. Even with this approach; this will require to extrapolate the observed wind resource to other locations using a wind flow model (NYSERDA, 2010). A study of Thailand identified only areas below 200 meters as suitable location for turbine placement so elevation standards vary by the project and geographic requirements (Bennui et al., 2007). A California study by Rodman and Meentemeyer (2006) aggregated wind data and elevation into a single layer to identify areas with high elevation and low valleys as the most suitable location for wind energy development.

Slope

Slope grade is an important factor that affects the suitability of a site. As highlighted in Table 3, minimal percent slope is required for a site to be considered suitable for wind development. In order to operate the heavy machinery and equipment for installation and maintenance, sites should be less than 20% slope grade (Acker et al., 2007; Tegou et al., 2010). However, there were some studies that had 30%t slope as acceptable due to technology advancement and techniques (Baban & Parry, 2001; Tegou et al., 2010).

Table 3: *Suitable slope for wind energy development*

Percent Slope	Suitability Rating
0 – 7	Excellent
7 - 16	Good
16- 30	Fair
30 - 40	Poor
>40	Unsuitable

Land Cover

The land cover impacts site suitability because there are only limited areas where turbines can be placed due to environmental and economic constraints. Also, land cover can be an economic indicator of how much it will cost the developer to rent or lease the land for wind energy development. Spatial pattern of wind energy development in Iowa displayed row crop as the most dominant land cover type in the state (Mann et al., 2012). Also, land value is determined based on corn suitability rating (CSR) which is comprised of soil type and this determines land value. Therefore, land cover will vary based on geographic location.

Certain land covers are not suitable for wind energy development. Restricted areas are urban areas, forests, wetlands, rivers, lakes, reservations, and parks which are unsuitable for wind energy development (Acker et al., 2007; Rodman & Meentemeyer 2006). Federal protected area like national parks, preserves, and forest especially in the western states are off-limits as well. If identified areas for wind energy development are near these areas, the developer should set a buffer distance ranging from 300-2000 m (Tegou et al., 2010).

Distance to Transmission Line (Power Grid)

A critical component in determining the economic viability and success of a project is the distance to power grid. Keeping the costs down in building a wind farm is minimizing the cost to power grid infrastructure which is to be installed. High-voltage lines can cost thousands of dollars per mile so it’s essential to considerer distance to a power grid as one of the factors for a suitable site. In fact, most studies cite power grid distance greater than 2000 meters as not economically viable (DOE, 2008; Mann et al., 2012). On the other hand, optimal locations are sometime in remote locations which make it difficult to develop (Mann et al., 2012). In order to for wind energy to provide 20% of United States electricity needs by the year 2030, significant upgrade and expansion of power grid is needed. The expansion of the power grid will also reduce the congestion on existing lines and better connect wind energy generation areas with high demand areas (DOE, 2008).

Transportation Accessibility (Distance to Highways)

Transportation accessibility is a factor cited by previous studies as a critical component for economical wind farm development. The distance to the nearest road/highway in relation to the wind farm impacts the cost of transportation and operating heavy machinery needed to install turbines. This can mean the difference between making the project feasible and cost effective or laden with increased transportation cost. Nguyen (2007) suggested that roads should at least be 100 m plus away from the neighboring turbine. Mann et al. (2012) indicates turbines are more likely to be sited further away from highways to minimize distraction for drivers and travelers. Generally, most studies recommend that the location of the wind farm to road/highway be no greater than 2,500 meters, otherwise, its economic viability diminishes

Distance to City

Some studies have examined the visual impact of wind farms near city/urban settlement in order to reduce issues of safety, visibility, noise, annoyance, and economic impacts on the local residents (Moller, 2006; Tegou et al. 2010). Generally, wind farms are developed in rural areas away from heavily populated places. This is importance constraint factor. As such, to minimize the visual impact and noise, turbines should be placed at least 500 meters away from the nearest city/urban settlement (Ramirez-Rosado et al., 2008).

Distance to Airport

Due to safety and visibility reasons, wind turbines are required to have marking and lighting by the Federal Aviation Administration (FAA; 2013) regulations to prevent aircraft collisions or radar interference. This regulation applies to military, commercial, and private airports. Nguyen (2007) cites areas unsuitable for wind development are those sited within the minimum of 2,500 meters of the nearest airport as set by federal law.

Distance to River

Baban and Parry (2001) and Nguyen (2007) classify factors like rivers, water bodies, and woodlands as second-grade factors but important factors which require turbines have a minimum distance of 400 meters as a buffer. Also, being further away from rivers or water bodies provides a solid foundation and reduces environmental impact.

Distance to Railroad

Railroads need to be offset 100 meters away from the nearest wind turbine (Nguyen 2007). Again, this is seen as a secondary factor because it is not really so frequent that developers run into rail track problems.

2.5 Modifiable Areal Unit Problem in Spatial Analysis

Gehlke and Biehl (1934) were the first to point out that simple statistics such as correlation coefficients could vary across scale and zoning systems as a result of grouping and aggregation. The study examined the tendency for correlation coefficients to increases as areal region are aggregated into fewer numbers of larger regions. However, the Modifiable Area Unit Problem (hereafter MAUP) was not fully formulated until Openshaw and Taylor (1979) evaluated systematically the variability of correlation values when different scales were used in the analysis. This is an ongoing issue in many areas of spatial quantitative analyses in geography.

MAUP has been well recognized in a wide range of disciplines such as transportation analysis, physical geography, and political geography due to the increase in quantitative studies within many disciplines (Dark & Bram, 2007; Flowerdew, 2011). A prime example of the MAUP problem exists in demographic analysis where the census data which is collected on fine resolution due to privacy concerns is released only after being spatially aggregated to a coarser resolution (block group, census block, and census tract) which affects the accuracy of the result.

While Openshaw and Taylor (1979) examined the correlation between the percentage of Republican voters in the 1968 congressional election and the percentage of population over sixty in the 99 counties of Iowa to examine the effect of the MAUP on bivariate correlation coefficients. As a result, the study illustrated coefficient becomes broader as the number of zones (areal units) decreased and the spatial autocorrelation and contiguous zoning procedure affects the resulting statistics. A study by Houston (2014) assesses the influence of MAUP in the analysis of built environment exposure on moderate and vigorous physical activity of people during walking periods. The study concludes that buffer or grid based zonal/scale configuration is heavily influenced with MAUP therefore impacting the result (Houston, 2014).

2.6 Summary

The drive to be energy independent and mitigate environmental impact, renewable energy sources have gained tremendous support in the public and private sector. Such support has led to increased wind energy development in the last decade globally. This transformation knows no boundaries as developed and developing nations equally participate in the initiatives to shift from fossil fuel to renewable energy. At this forefront is wind energy which is a clean and abundant resource and the United States is endowed with tremendous wind resources. As one of the leading nations in wind energy development, federal and state polices are spurring development. Wind and other renewable energy sources are also incorporated into the federal and states energy portfolio. At the forefront of this development, Iowa utility scale wind industry is robust and among the top wind energy producing states in the nation. Despite the robust growth and maturity of the wind energy, the important factors, spatial dynamics, and regional manifestation of site suitability are not yet developed.

Factors from the physical, environmental, demographical, and economic components were presented in this section. Policies and regulations were also highlighted as major components of the site suitability for wind energy development. Since, wind energy development is projected to grow in Iowa, the need to identify suitable locations and improve resource characterization based on Iowa context is essential if Iowa is going to contribute to the DOE's 20% wind energy goal by the year 2030. Existing models are insufficient as they are based on a limited number of spatial variables and on traditional approaches for suitability analysis based on weighted overlay and buffer.

Despite being widely used, there are two shortcomings with such an approach. First, the difficulties in handling spatial data inaccuracy, multiple measurement scales and factor interdependency hinder in identifying suitable sites. Second, requirements of prior knowledge in identifying criteria, assigning scores, determining criteria preferences, and selecting aggregation function are solely based on individual expertise rather than on an empirically driven approach.

This research will examine the effect of scale on contributing factors which will give better insight on the spatial dynamics and interactions of the factors. Therefore, this research advances an empirical approach to analyze the contributing factors, their relative importance, scale-dependency and regional manifestations. In addition to identifying the contributing factors at multiple scales, the need for optimal placement based on a spatially-integrated nuanced predictive model adapted to the Iowa context is critical. Many publications concerning various regions in recent years try to fill this knowledge gap demonstrating the increased importance of determining the optimal placement of wind turbines to maximize the benefits of wind energy. Ultimately, the predictive framework will optimize wind turbine placement at different scales from both resource utilization and resource characterization perspectives.

CHAPTER 3

METHODOLOGY

3.1 Study Area

The study area for this research encompassed the state of Iowa. According to U.S. Census 2010, Iowa had a population of 3,046,355 (30^{th} most populous state in the USA) with 99 counties and a total area of 56, 276 sq. mile (145, 743 km^2). Iowa lies within the Central Lowlands region of the United States (Figure 2). Iowa is bordered by Minnesota on the north; Nebraska and South Dakota on the west; Missouri on the south; and Wisconsin and Illinois on the east.

Much of the Midwest surface physiography has been shaped by a series of continental glaciers flattening hills and filling in valleys. Debris carried by glaciers was deposited in moraine features giving some relief to relatively flat areas. The regions river valleys and lakes were formed as a result of this period. Due to the low relief throughout the Midwest region, climatic differences gradually change in latitude (between north and south) and longitude (between east and west). This region is generally perceived as being relatively flat but there is a measure of geographic variation. In particular in the Grate Lake Basin, and northern parts of Wisconsin, Minnesota, and Iowa demonstrate a high degree of topographic variation. Prairies cover most of the states west of the Mississippi River. Precipitation decreases from east to west resulting in different type of prairies (Garland, 1955).

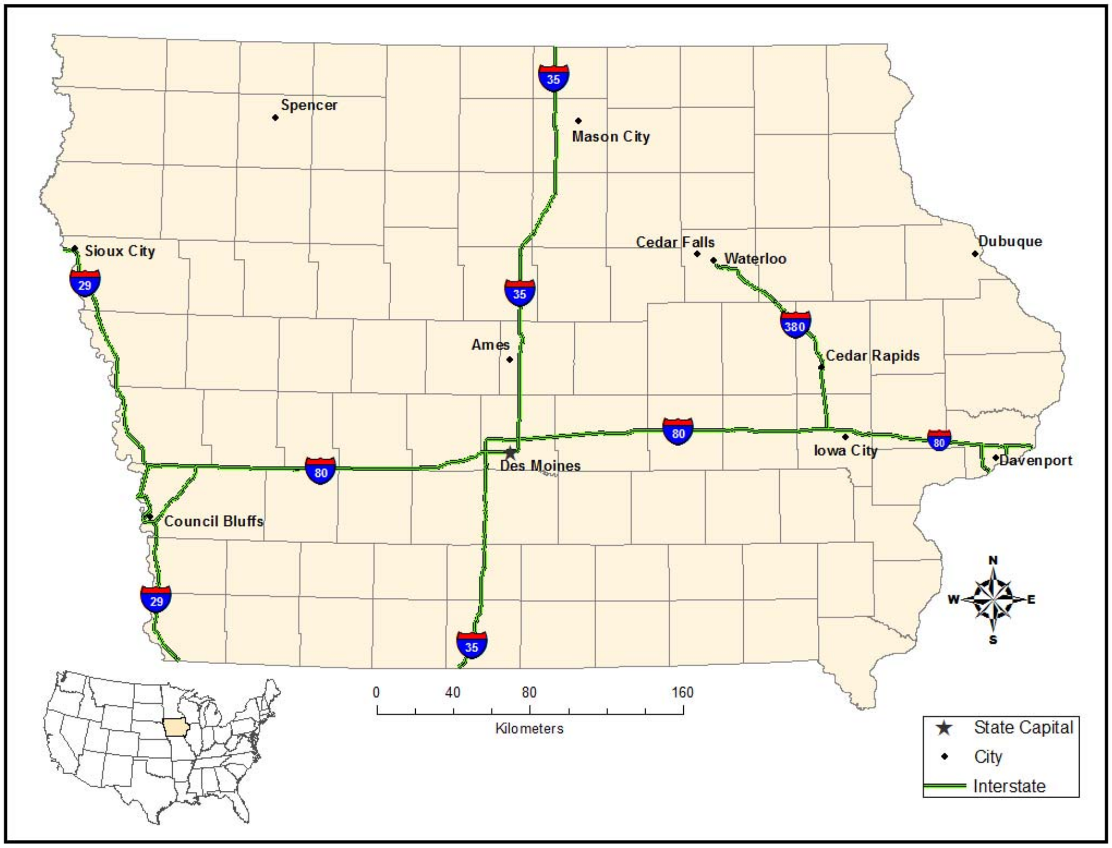

Figure 2: Study area: major cities and interstate network

3.2 Physical and Environmental Characteristics of Iowa

The topography of Iowa is generally flat plains to rolling hills. The glaciers from the last Ice Age shaped the terrain by laying down deposits of drift debris and carving distinct topographic features which are the till plains of mixed clays, sands, gravels and boulders (Freedman, 2010; Nelson, 1967). The deposits of compacted silt and loess cover large areas of the state. There are eight distinct landform features: Des Moines Lobe, Iowa Surface, Loess Hills, Mississippi Alluvial Plain, Missouri Alluvial Plain, Northwest Iowa Plains Paleozoic Plateau and Southern Iowa Drift Plain. These are due to Glaciers from the last Ice Age (Figure 3). The northeast is a hilly area relatively unscathed by the glaciers. High bluff are distinct features along the Missouri and Mississippi Rivers (Freedman, 2010; Nelson, 1967). The Mississippi River forms the eastern boundary of the state while the boundary along the west is formed by the Missouri River south of Sioux City and by the Big Sioux River north of Sioux City(Freedman, 2010; Nelson, 1967)..

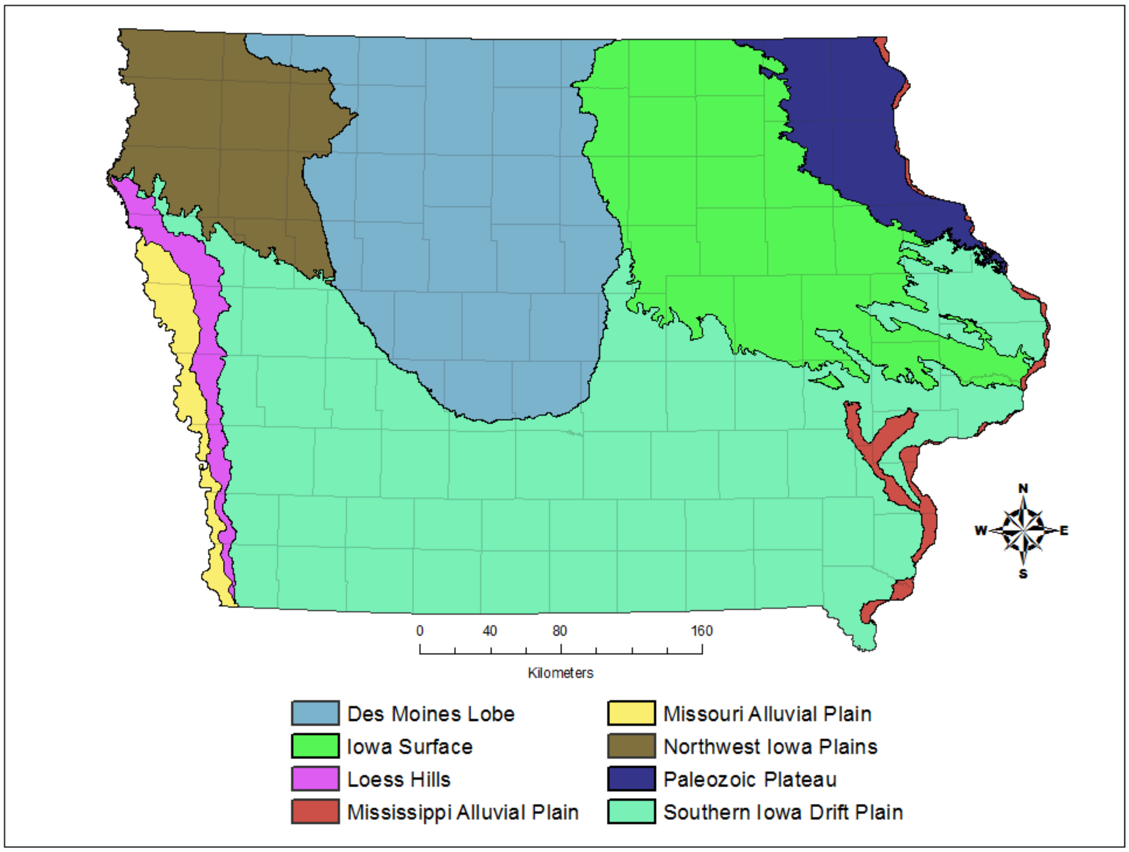

Figure 3: Landform Regions of Iowa

The mean elevation is 340 meters and the highest point in the state is found in the Northwest corner of Iowa with an elevation of 509 meters above sea level; the lowest point is 146 meters above sea level in the confluence of the Des Moines and the Mississippi River near Keokuk in the southeastern part of the state. North central is the flattest part of the state as the result of the last Ice Age while southern and western Iowa consist mostly of rolling to hilly land (Freedman, 2010).

The various landform regions provide rich soils that make Iowa a fertile and agricultural base. The fertile soils tend to be located in the northwest central part of the state. However, Iowa was once comprised of widespread tall-grass prairie. The state was largely converted to an agricultural landscape by the late 1800's following the European settlements (Freedman, 2010). Widespread use of irrigation farming and large-scale farm machinery in the 20th century, coupled with a shift toward a more mass agricultural production, Iowa's landscape was transformed from diverse prairie plants into the large-scale, monoculture farming that are common today (Freedman, 2010).

Row crops are cash cows for the farmers, and each year, approximately 80% or more of Iowa's cropland is planted with corn or soybeans. Soil productivity and agricultural land value assessment is determined using the corn suitability rating index (CSR). This index rates soil types based on their productivity for row-crop prodution. CSR values can range from a high of 100 to a low of 5 index points and this rating is a tool to establishing a cash rate for a parcel of land (Hofstrand, 2010; Miller & Iowa State University, 2012). This is based on the premise that a high CSR means high land prodction of row crop which means high yields that generate large revenue. Land with a high CSR value would have a higher rental rate than land with a low CSR rate. Overall, CSR is an important indicatior of the productivity of farmland for row crop production. The CSR can be used to compute the rental rates for a tract of land. This is computed by dividing the average rental rate by the acreage cropland CSR for the county (Hofstrand, 2010; Miller & Iowa State University, 2012).

3.3 Iowa Wind Resources Characteristics

The North American Interior Plains extends west from the Appalachian Plateau to the Rocky Mountains. Containing large rivers (Mississippi, Missouri, and Ohio), diverse vegetation landscape (variety of grass lands) and climatic conditions vary throughout this region from extreme cold to very humid summers (Johnson, 1985). The meteorology influencing the wind resource is largely controlled by the position and the strength of the upper-level jet stream and turbulences within the jet stream (Johnson, 1985). During the winter, the jet stream positions further south which in turns creates stronger winds than summer. Spring and fall, the position of the jet stream generally lies between the summer and winter positions (EnerNex Corporation & WindLogics Inc., 2004).

The main factor controlling the jet stream position and speeds is the magnitude and location of the temperature gradient. A larger temperature gradient exists in the winter corresponding to a stronger jet stream and summer’s small temperature gradient corresponds to a weaker jet stream (Johnson, 1985). The key factor driving the wind resource in the lowest 100 meters of the atmosphere is the horizontal pressure gradient.

Furthermore, wind systems converge in the mid latitudes where the prevailing westerly winds are the primary force that affects the Midwest region (Johnson, 1985). The Upper Midwest exhibits significant seasonal variability; therefore, extreme seasonal weather variation produces wind speeds that are often very high (EnerNex Corporation & WindLogics Inc., 2004). The surface of the region being relatively flat grassland with hills, valleys, river bluffs, and lakes stirs a complex and variable wind.

Iowa's geographical position creates an abundance of wind due to the position positioning and physical characteristics of the region. Wind speed is the rate at which air flows past a point above the earth's surface, and it can very over time and space. Iowa has seasonal wind strongest in the winter and early spring and weakest in the summer. Daily winds generally are strongest during the afternoon and lightest during the early morning. Figure 4 illustrates the distribution of wind speed in Iowa obtained from the Iowa wind center. The north central and the northwest parts of the state have some of the most fertile soils, highest elevation and wind speed (7.0 - 8.0 m/s on average). The Des Moines Lobe and the Northwest Iowa Plains landforms, and the northwest portion of the Southern Iowa Drift Plain possess the highest elevation, wind speed, and also the most fertile farms (based on CSR). In contrast, the South, Southeaster portion of the state has the lowest elevation but also a drop in wind speed (6.0 - 6.5 m/s). One turbine can be found (Henry County) in the southern portion of the state.

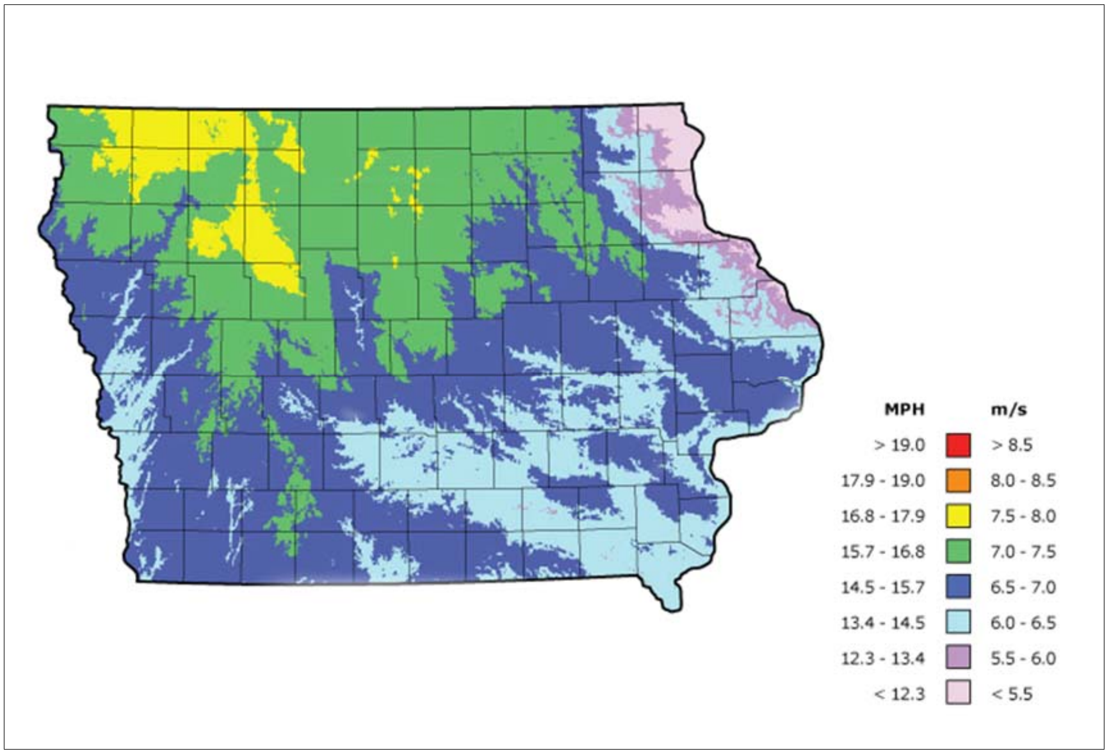

Figure 4: Iowa average wind speed at 50 m height

3.4 Data

The factors (variables) slope, elevation, wind power class (wind speed), land cover, population density km², proximity to neighboring turbine, distance to transmission line, city, highway, airport, river, and railroad were compiled in ArcGIS 10.0 and converted to raster. Table 4 contains a summary of the predictors and sources used to obtain the data. The Iowa Department of Natural Resources Geographic Information Systems Library (2012), Iowa Energy Center (2012) and FAA (2013) were the sources for this data collection. Iowa DNR uses the Universal Mercator; Zone 15 North (UTM Zone 15 N) spatial reference system for all the data and the FAA dataset were set to was in commonly separated values (CSV). The FAA data was spatial referenced to UTM Zone 15 North and converted to feature in ArcGIS 10.0.

Table 4: *Data Description and Source*

Data	Description	Source
Existing Turbines	X Y coordinates	Iowa DNR and FAA
Slope	Derived from DEM (%)	Iowa DNR
Wind Power Class	6 Classifications	Iowa Energy Center
Elevation	10 m DEM	Iowa DNR
Land Cover	Categorical classification	Iowa DNR
Transmission Line	>69 kilovolt	Iowa DNR
City	Incorporated cities	Iowa DNR
Highway	Major roads	Iowa DNR
Railroad	Current RR track polyline	Iowa DNR
Airport	Point feature	Iowa DNR
River	Polyline	
2010 Population density	Census tract	U.S. Census 2010

Turbine Feature

Figure 5 displays existing turbines in Iowa. A vector layer of 3,177 existing wind turbines with X, Y coordinates (NAD 1983 UTM Z 15) was acquired from the DNR (2013) and FAA (2013). To ensure data reliability, only existing turbines as of December 31, 2013 were used, and validation consisted of checking the latest aerial photographs. The FAA dataset (1,222 turbine records) and the DNR dataset (1,955) were combined for a total of 3, 177 existing turbine records used in this analysis. Number of turbines distribution and characteristics are as follows: Figure 5 displays the average turbine height (including blade) from the year 2008 to 2013. The hub height (just the tower) is

80 meters in height but newer turbines are being built with 100 meters hub height. The numbers of turbines constructed in Iowa varies year to year, and it can easily be affected by regulation, policies, and economy. In 2008, 135 turbines were constructed and were followed by 470 turbines (2009), 20 turbines(2010), 690 turbines(2011), 442 turbines(2012), and 121turbines (2013).

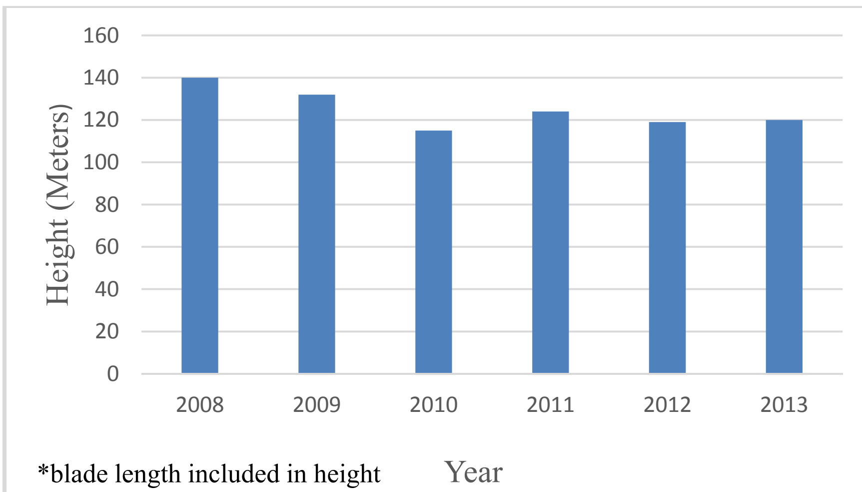

Figure 5: Average turbine height by year

Figure 6 displays the distribution of turbines which tend to be clustered. The majority are located in the north central and the northwest parts of the state. This part of possesses some of the rich soils that make the state an agricultural basket of the country. Current turbines locations average a CSR value of 68.3 out of 100. So, if a farmer or large land owner were to decide to rent out land to wind energy developers, the compensation tends to be connected to the CSR value of the parcel of land. The higher the CSR value, the more valuable the land thus the more compensation the land owner would receive.

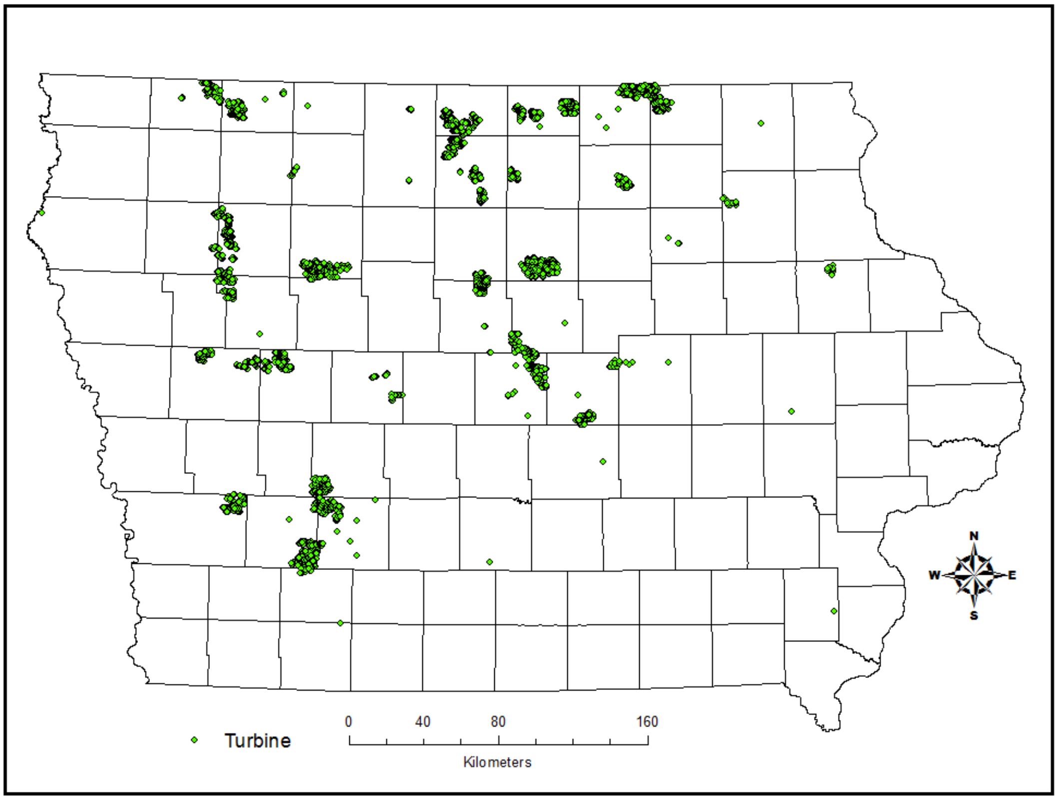

Figure 6: Existing turbines location in Iowa

Wind Power Class (WPC)

Wind Power Class (hereafter WPC) data was obtained from the Iowa Energy Center (2012) which provides annual average wind speeds in Iowa at the 50 meters above ground level (Figure 7). This data was produced by the U.S. Department of Energy's Wind Powering American program and validated by NREL and other wind meteorological consultants. The wind profile power law is often used in wind power assessment where wind speeds at the height of a turbine are estimated from near surface wind observations (~10 meters). Since wind profile of the atmospheric boundary layer up to 200 meters is generally logarithmic in nature, this is used to calculate the wind power density then reclassified to wind power class 1-7, the latter being the highest (Table 3).

Wind speed range 11.5 to 12.5 mph have wind power density 150 – 200 (w/m²) and are classified as wind power class 3 (=< 6.4 m/s). Iowa's wind resource characterization by NREL illustrates Iowa as having the highest wind speed classification in the northwest going southeasterly (Figure 7). The northwest part of the state indicates the strongest wind with wind speed 6 – 7 m/s (wind power class 5-6), and it has a total area of 49,246.903 km² which is almost 30% of the total land. This part of the state is also the least populated. It contains the most fertile soil based on the corn suitability index, and also it's where turbines are built and continue to be built. Wind speed drops in the south and easterly part of the state. However, there is substantial wind available (wind power class 3 – 4), and this area encompasses over 93,000 km². On the other hand, the northeast corner along the Mississippi River and southward, displays very low wind speed (wind power class 1 and 2) which is very weak and there are no turbines built in this area.

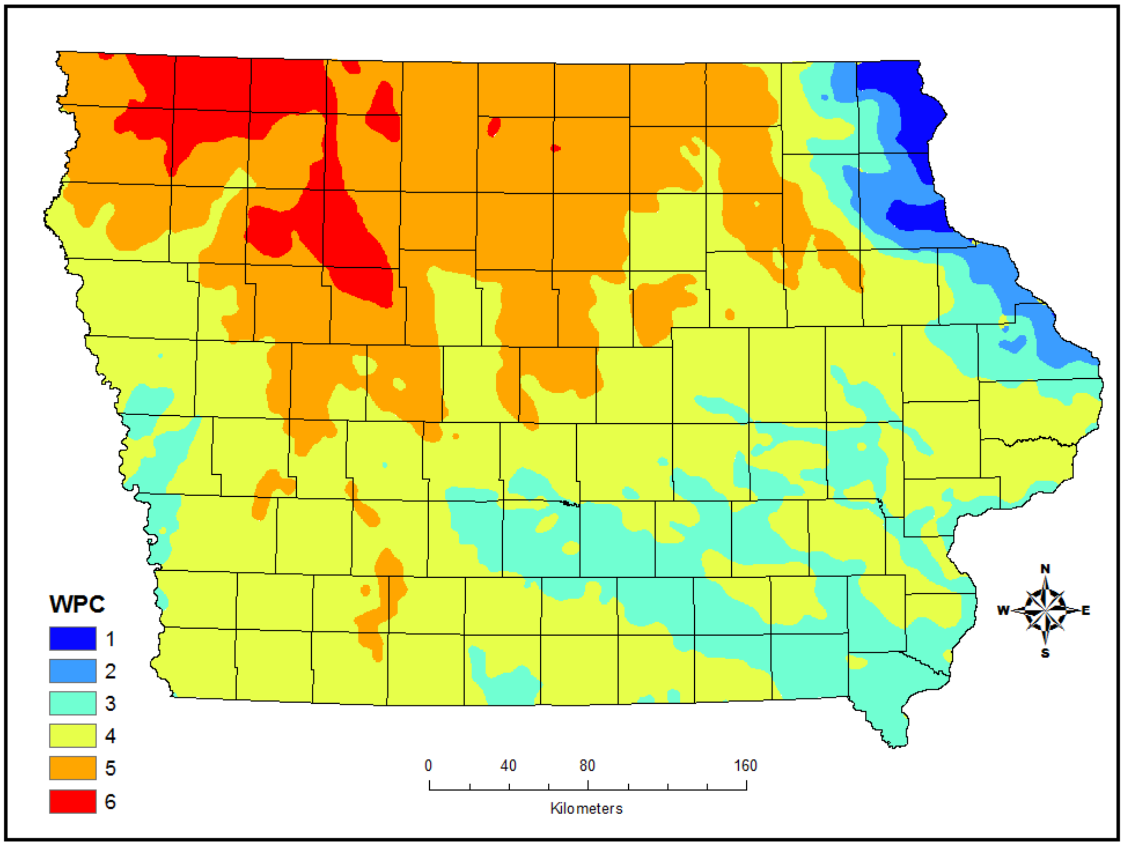

Figure 7: Wind power class distribution in Iowa

<u>Proximity to Neighboring Turbine (NT)</u>

The proximity to nearest turbine variable was used as the spatial lag independent variable. The assumption is that proximity to other turbines is indicative of factors being suitable for turbines to exist. This is appropriate since NT variable directly influences the likelihood of another turbine. The general standard for modern day turbines are spaced 300-500 meters depending on the size of turbines and the configuration of the wind farm. In this study, turbines tend to be sited at 350 meters or greater (greater than 3 blade diameters) apart next to each other and over 1,000 meters (10 blade diameters) in the primary wind direction. Proximity to neighboring turbines was classified as “yes” it has a neighbor or “no” it does not. A turbine within 500 meters was assigned (1) yes and greater than 500 meters was assigned (0) no. NT was used as ordinal variable in the logistic regression models.

Elevation

Digital Elevation Model (DEM) is a ground representation of land surface where every pixel is associated with an elevation value above sea level (Figure 8). Iowa’s elevation ranges from 507.365 meters at the northwest part of the state to 146.222 meters at the southeast along the Mississippi River. The northwest going south along the Missouri Alluvial land contains areas of the highest elevation in the state. Elevation declines going southeasterly towards the Mississippi river. Also, Iowa’s major rivers (Iowa, Cedar, and Des Moines) are clearly visible. In this research, 10 meters DEM was acquired from the Iowa DNR. Spatial resolution is the pixel representation of the surface. In this case, the 10 meters DEM means that each pixel represents a 10 meters by 10 meters area on the ground. Elevation cell values were extracted in ArcGIS 10.1 for each turbine point to obtain elevation and stored as elevation field.

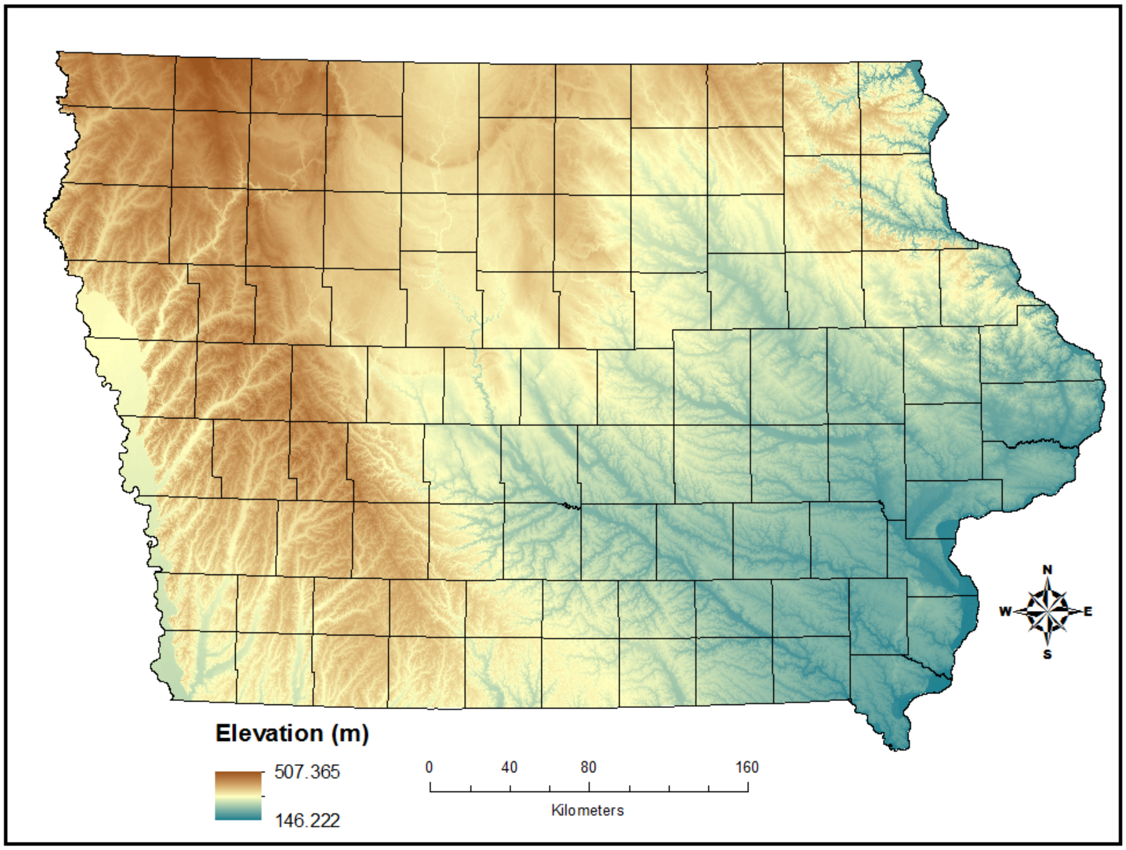

Figure 8: Elevation Profile of Iowa

<u>Slope</u>

High slope affects the ability to operate heavy machinery needed to install wind turbines; therefore, it's important to identify ideal percent slope (Figure 9). The 10 meters DEM was used to create a continuous slope layer using the slope spatial analysis tool in ArcGIS 10.0. The fertile part of the state is also with minimal slope change and this is also the area where majority of turbines are located. The state mostly is below the 30% slope identified in literature as ideal area for wind development. However, the northeast corner of the state contains the highest slope percentage of greater than 30%

slope but this is also the area with the lowest wind power class (1-2) as well. Turbine feature percent slope was extracted from the raster for each turbine point to obtain percent slope values.

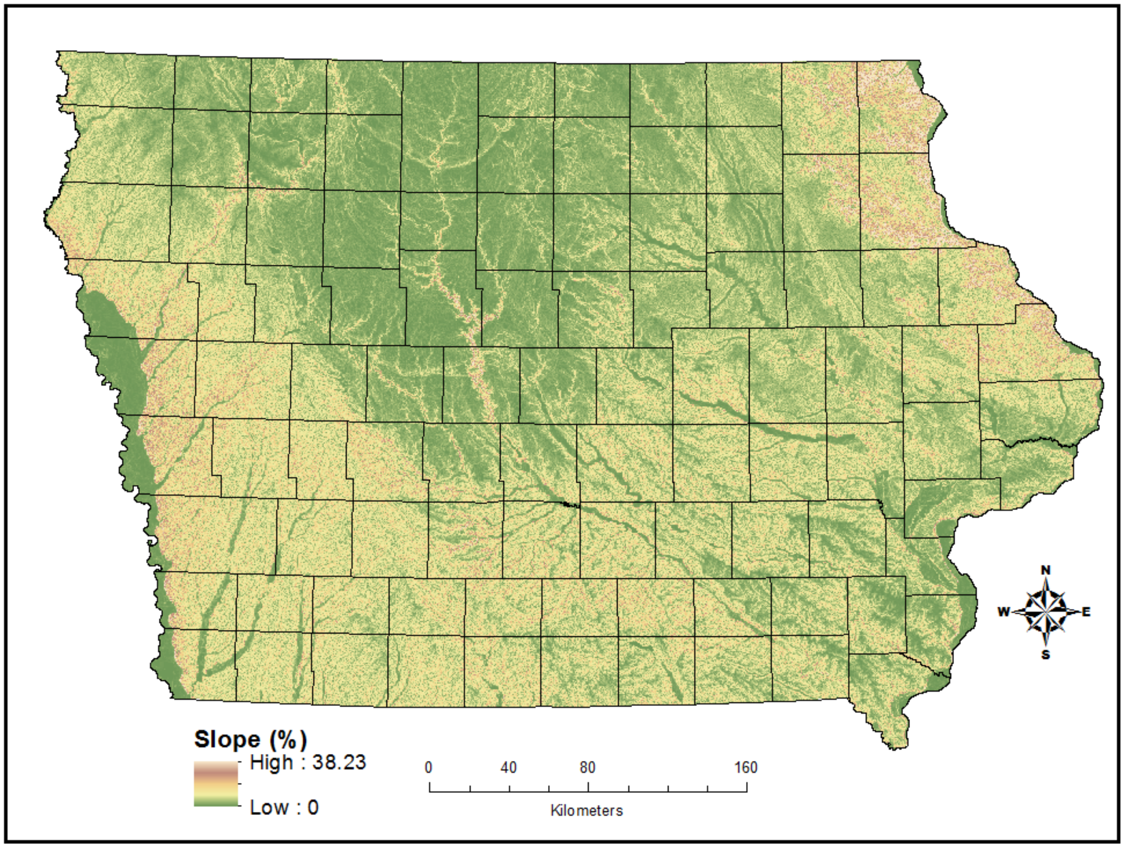

Figure 9: Percent Slope

Land Cover

The Land cover feature for 2002 was obtained from the DNR. The data contained 16 land cover classification but for the purpose of this research, it was reduced to five land cover classification in ArcGIS 10.0 (Figure 10). The most common land cover classifications are cropland and grassland. The north half of the state is mostly cropland and grassland while the southern half contains most of the forest and woodland areas. The blue areas highlight the populated areas and urban settlements.

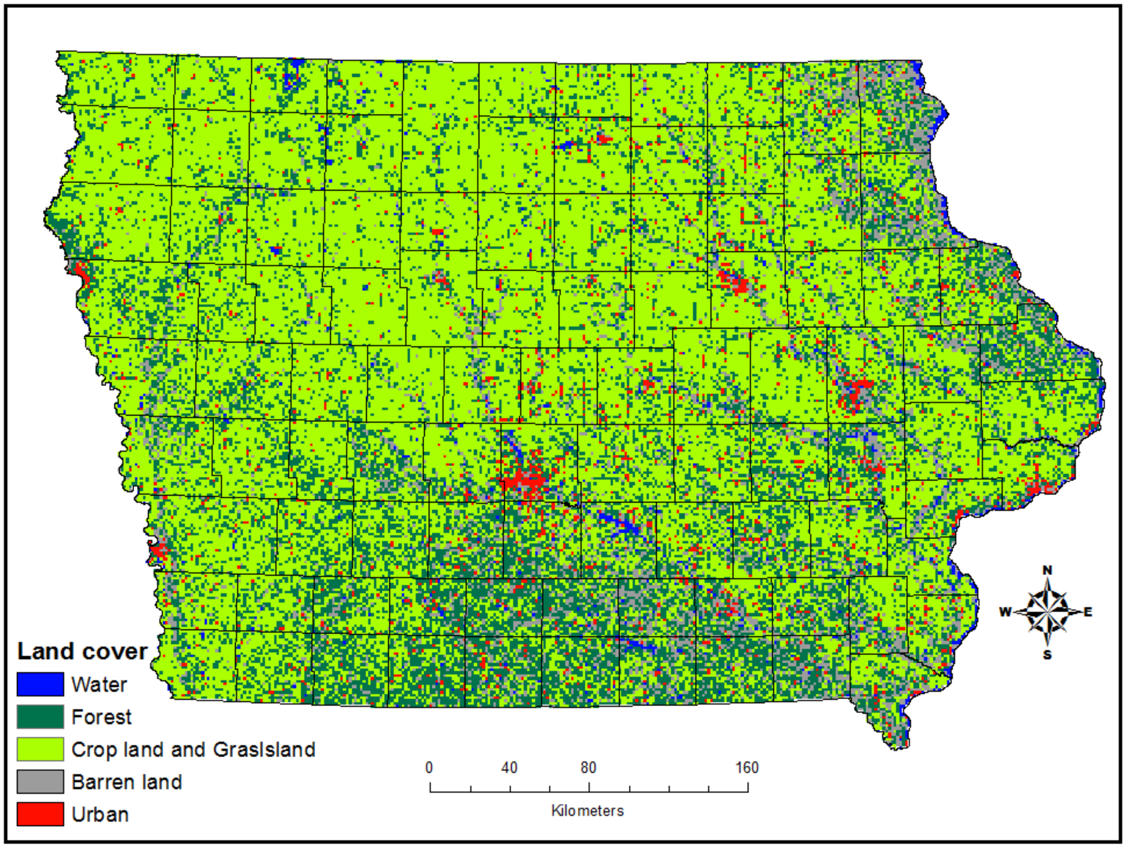

Figure 10: Iowa land cover classification

<u>Population Density (Pop Den)</u>

Population density was used as a factor to measure the effects populated areas influence on wind development and to identify where the development is near highly/low populated areas. Iowa census tract population data from the 2010 census was acquired from the U.S. Census Bureau. The shapefile was processed to contain population density and it was converted to raster with cell values assigned population density from the census tract (Figure 11). The cell values were extracted for each turbine point.

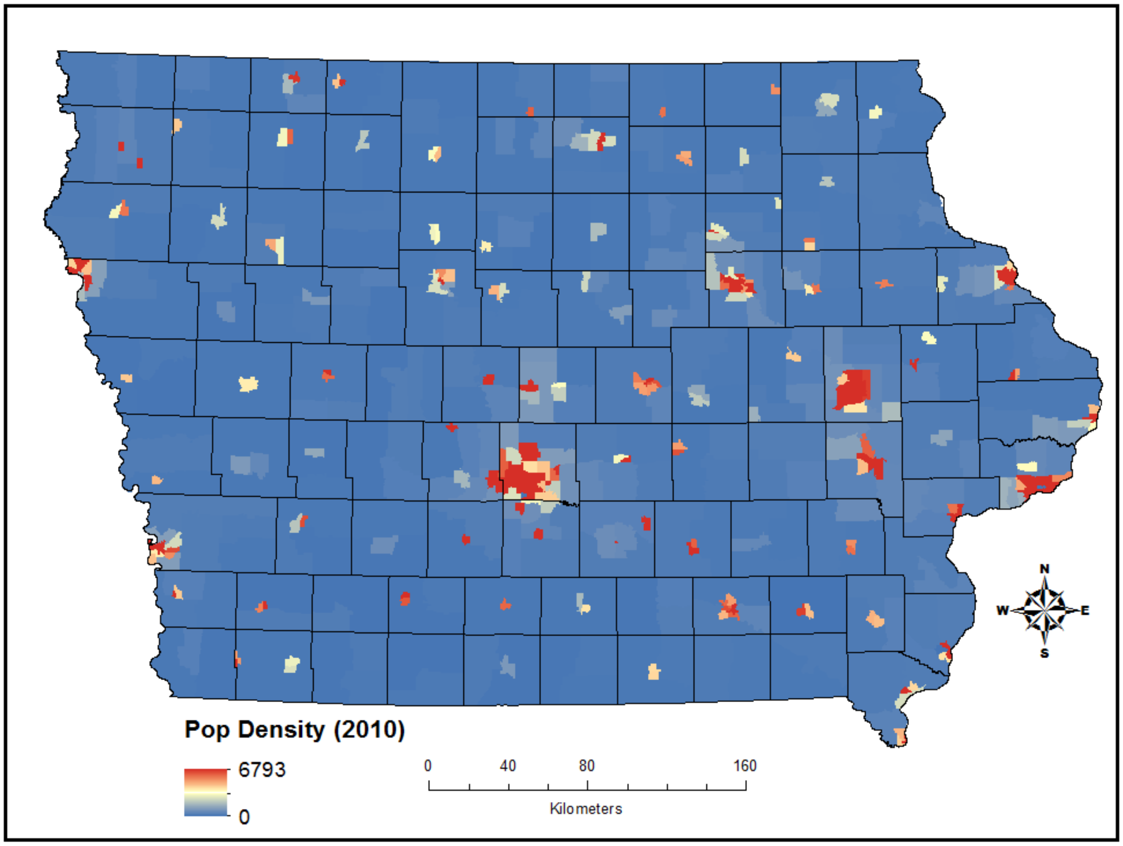

Figure 11: Census track population density

Distance to Transmission Lines (TL)

Connection to the grid is an intrinsic part of wind energy development generally included as part of project feasibility study. Transmission line feature was obtained from the DNR and it contains power lines greater than 69 kv up to 345 kv (Figure 12). The turbines distance to the transmission line is a critical component for the economic viability of the project and to deliver the energy to the market place as well. Distance to TL was measured in km to assess the impact on turbines, e occurrence.

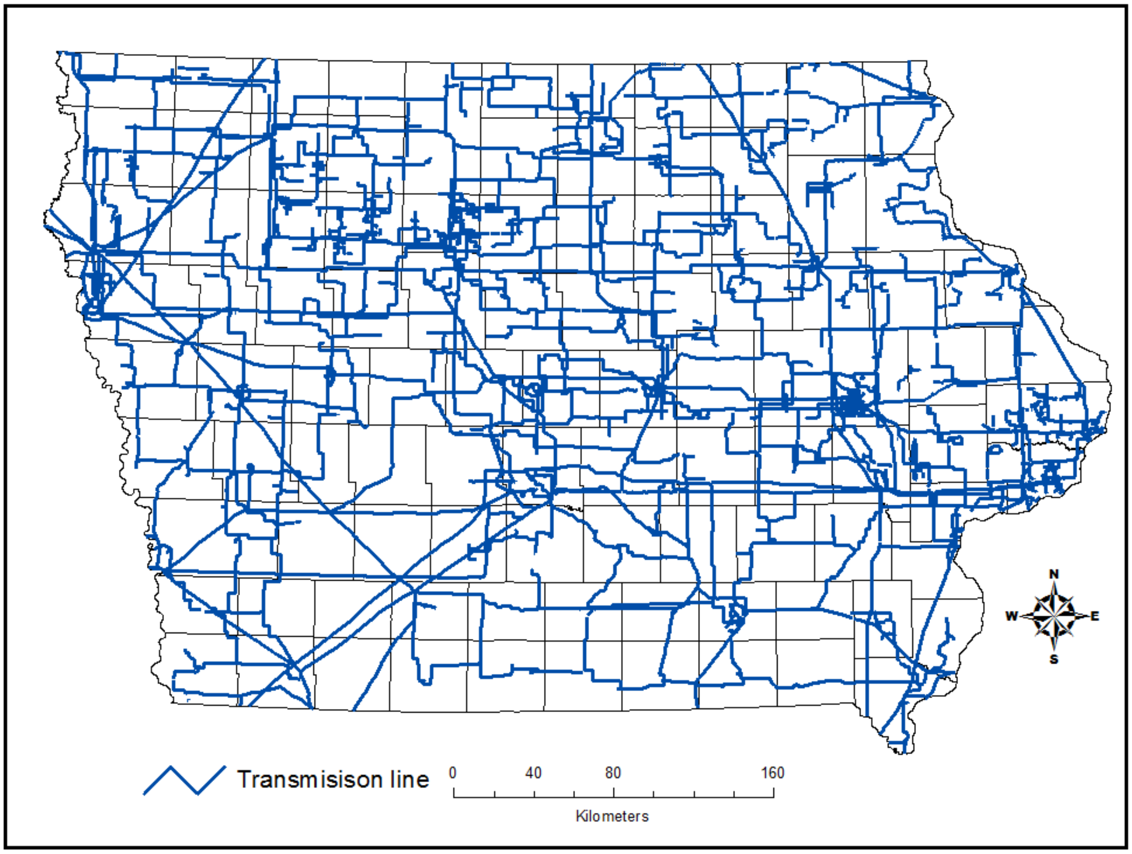

Figure 12: Electric transmission networks in Iowa

Distance to Airport (AP)

Wind turbines represent a risk of collision with low flying aircraft and interfere with radar operations. Developments within a specified radius of major civilian and military airports are subject to mandatory approval by the FAA. Airports point feature was obtained from the DNR GIS Library which contains landing facilities in the state as supplied by the FAA (Figure 13). Distance to airport relative to turbine location was measured in km.

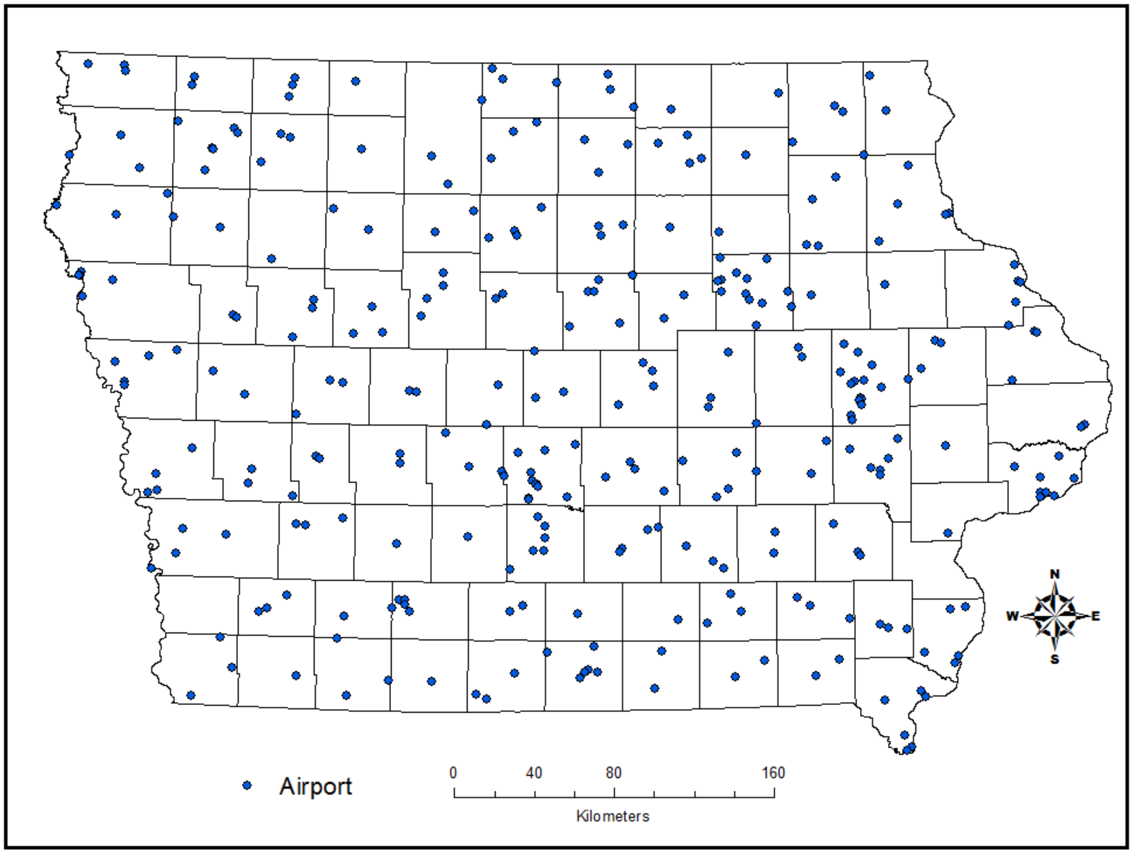

Figure 13: Civilian and military airports in Iowa

Distance to Highways (Hwy)

Literature has indicated the proximity highways need to be within the threshold suggested since accessible roads are needed to operate heavy machinery and equipment need to install and repair turbines. Further details can be found in Section 2.4 which describes the importance and criteria needed for the highways factor. The highway feature was obtained from the DNR, and this feature was used to measure distance to highway from each turbine (Figure 14).

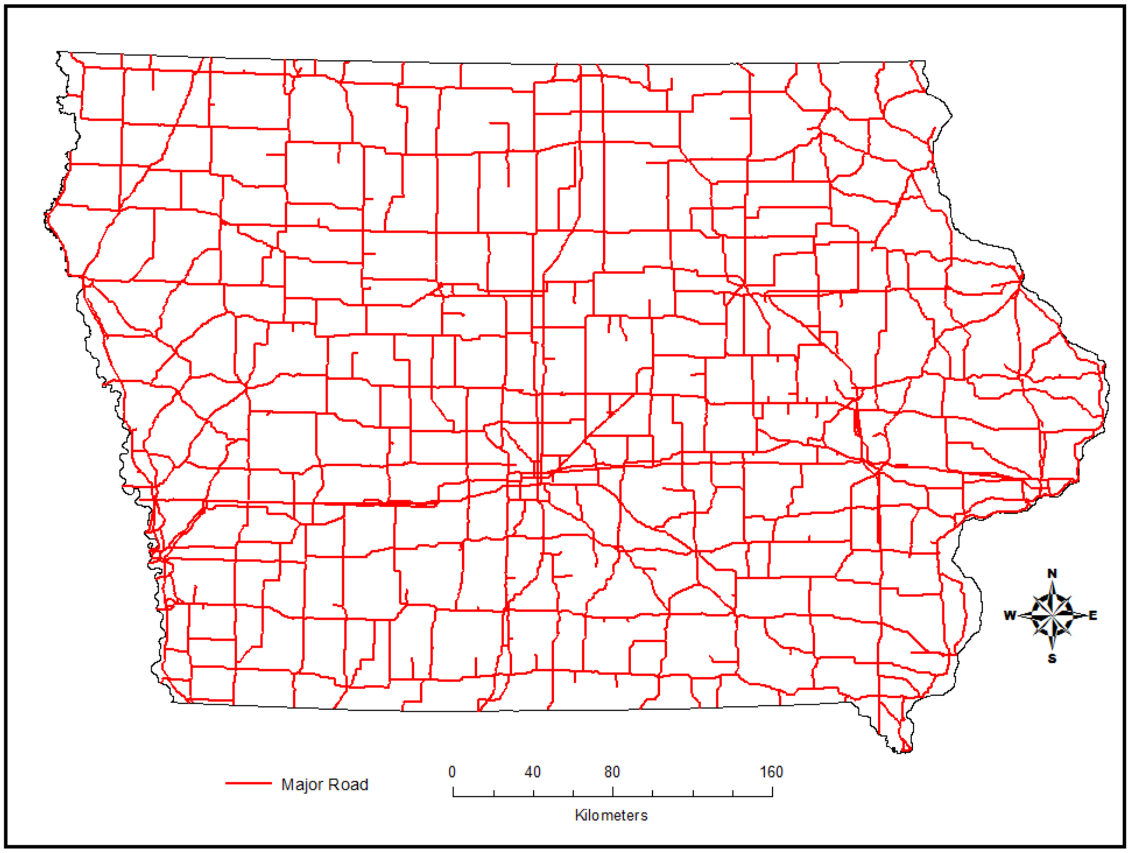

Figure 14: Major road networks in Iowa

Distance to City

A 2010 incorporated cities boundaries were derived from the census shows data obtained from the DNR (Figure 15). This was done to gain insight into whether or not turbines are placed in an adequate distance from populated areas to mitigate aesthetics and noise concerns by people near wind turbines. Further details can be found in section 2.4. Again, the distance to the nearest incorporated city was collected for each turbine point.

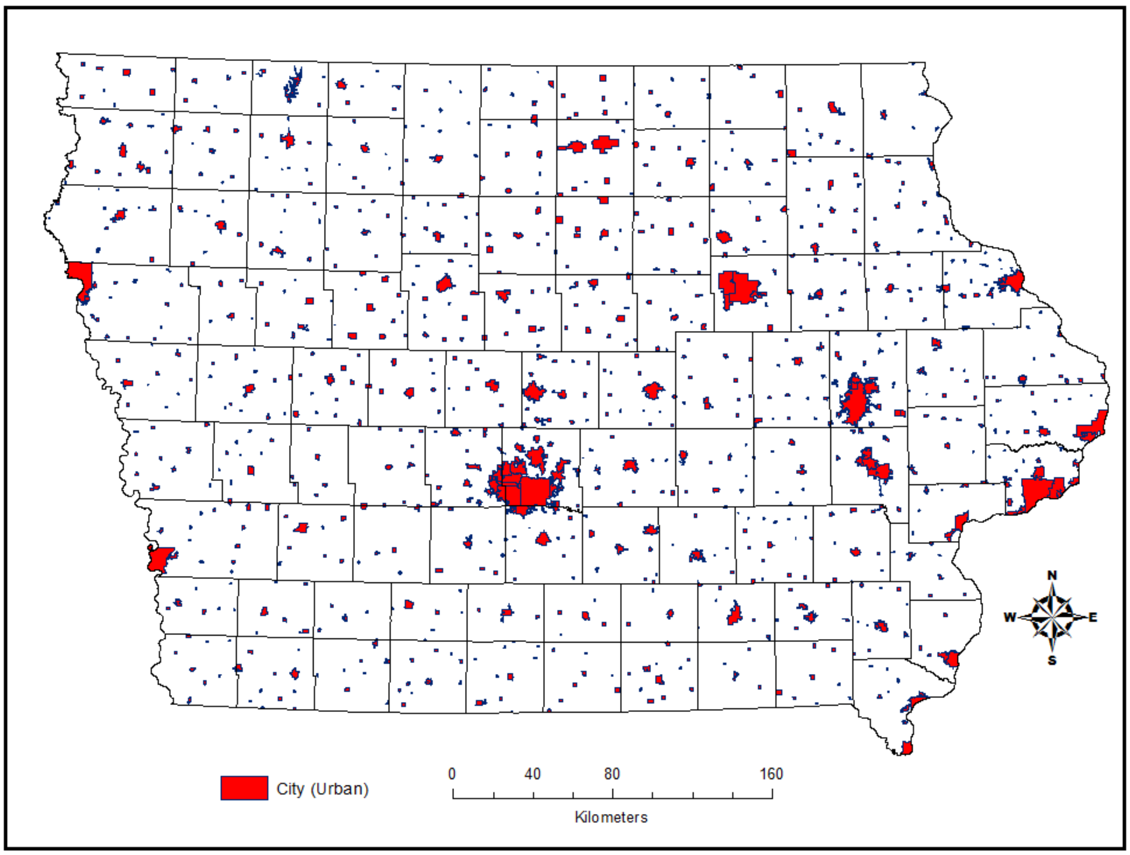

Figure 15: Urban areas in Iowa

Distance to Railroad (RR)

Infrastructure like railroad variable was included to examine whether or not it hinders the development of wind energy (Figure 16). Section 2.4 describes the importance of railroads in the development of wind energy. Railroad feature was obtained from the DNR, and distances to turbine points were measured to be used in the regression analysis to determine the most influential factors in wind energy development.

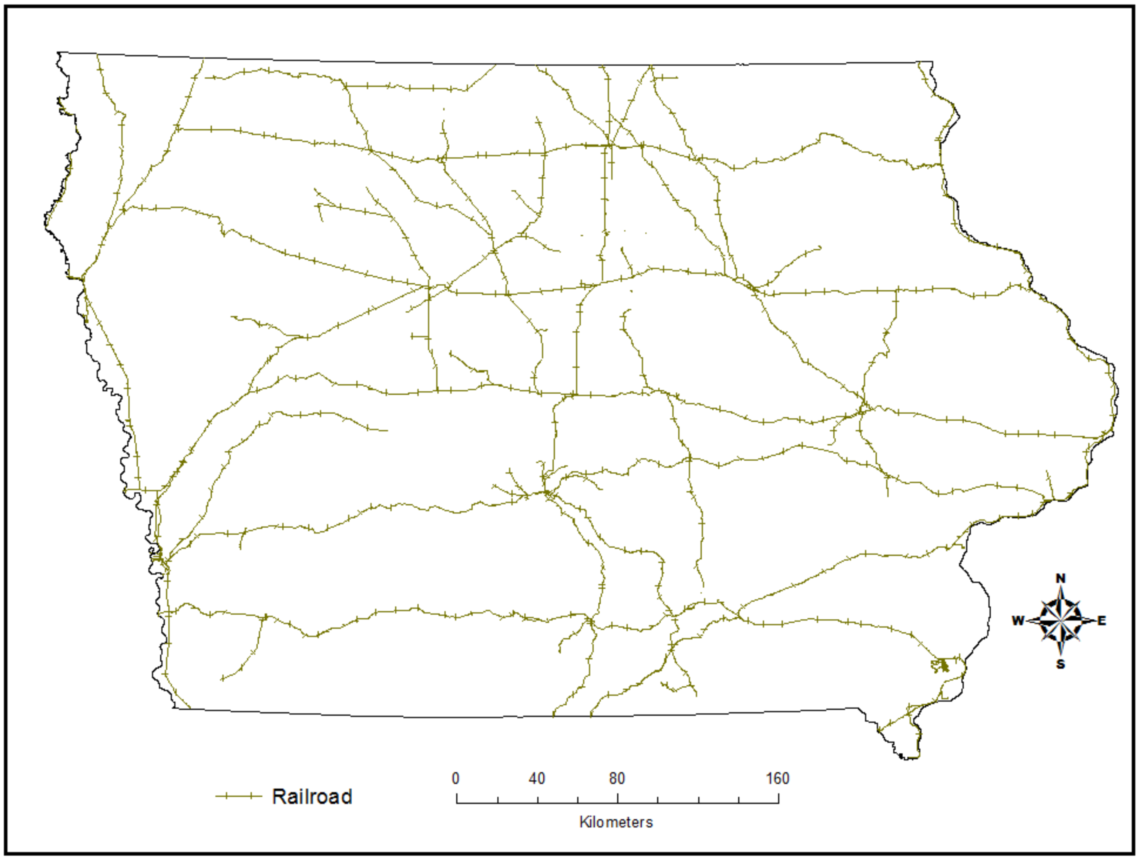

Figure 16: Iowa's railroad networks

Distance to River

General description of the influence of rivers in the development of wind energy development is given in Section 2.4. Ultimately, minimizing environmental hazards is critical so the river feature was included for analysis, and distance to turbine feature was measured in km. The river feature was obtained from the DNR (Figure 17).

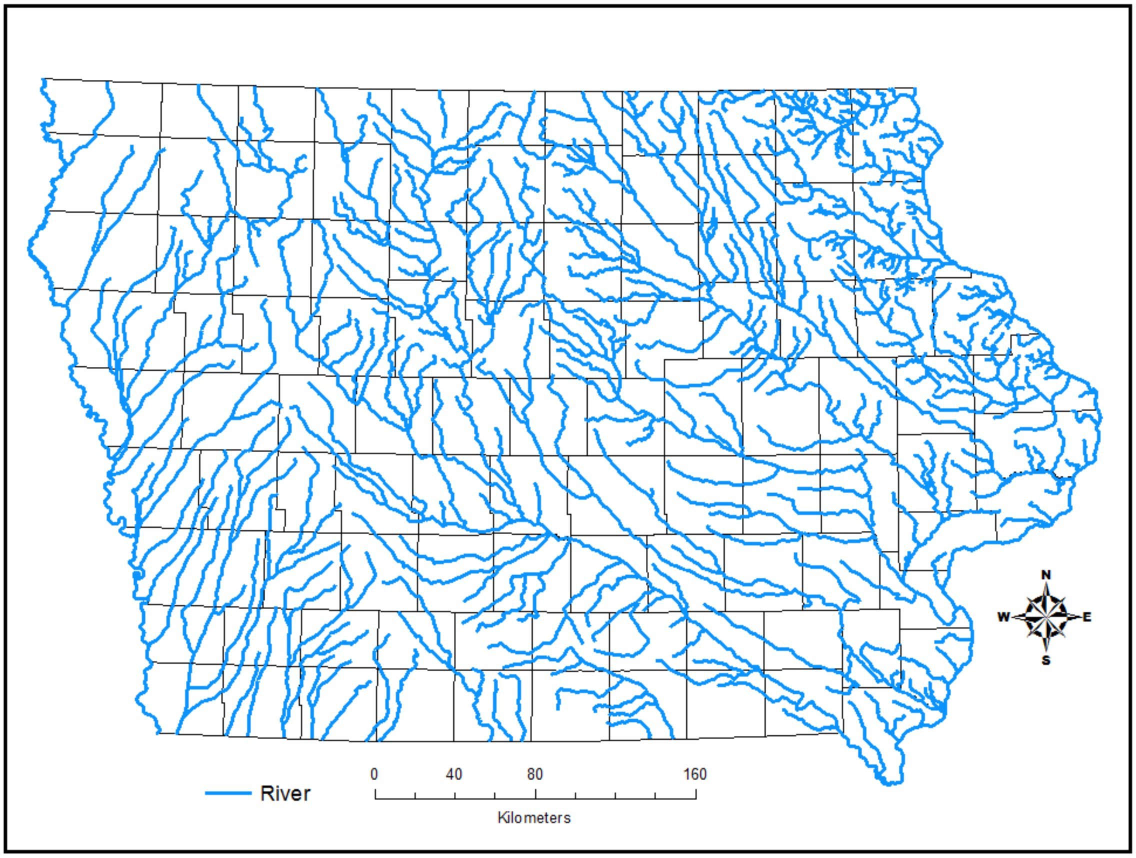

Figure 17: Iowa's rivers

3.5 Empirical Module

Standard Logistic Regression

Previous approaches to wind farm suitability modeling incorporated weighted averages, priority settings, outranking, fuzzy principles, overlay, buffering, and weighted overlay analysis to identify suitable site, or they solely focused on wind turbine placement within the wind farms. Expert knowledge and GIS were used to determine criteria deemed important such as distance from environmental sensitive zone, urban areas, and transmission lines then incorporated as constraints in the model to determine suitability (Aydin et al., 2010; Baban & Parry, 2001; Nadai, 2007; Rodman & Meentemyer, 2006). On the other hand, Mann et al., (2012) utilized mixed modeling that incorporates empirical approach to analyze and identify the spatial patterns of wind energy development in Iowa.

In contrast, this study incorporates spatially explicit empirical modeling framework where existing turbine locations and normative criteria's (i.e., regulation and polices) were used to identify suitable sites. This model was built first identifying the most useful and influential explanatory variables at different scales in wind energy development. Figure 18 displays the spatial scales used in the regression analysis. Table 5 provides attributes of each scale, the number of turbines used to build the models, and total area each scale encompasses.

The spatial characteristics and distribution of wind turbines were examined to gain insight into spatial dimensions. Spatial autocorrelation, which measures the degree to which near and distant things are related was implemented in this study. Specifically, spatial lag regression was used in the empirical model by including statistically significant spatially varying explanatory factors such as wind power class, elevation, proximity to neighboring turbine (spatial lag variable) and other explanatory variables.

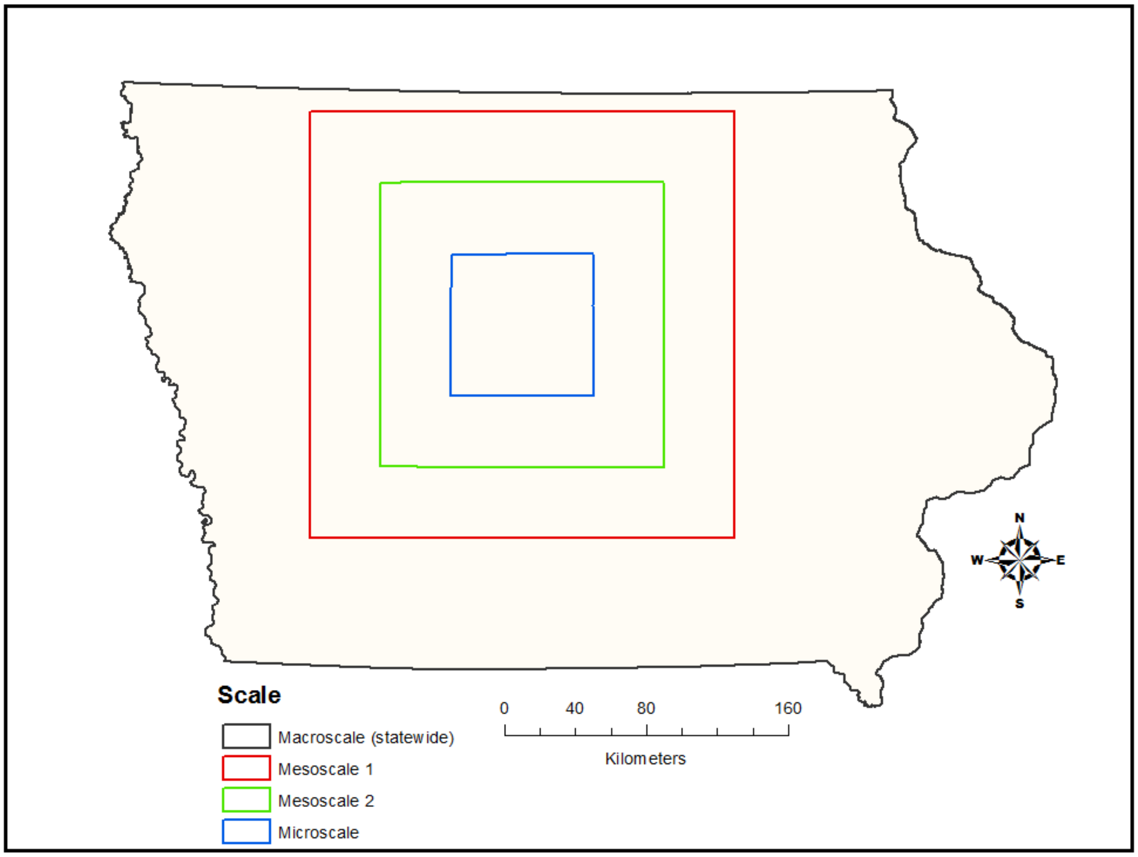

Figure 18: Multiscale study area

Table 5: *Logistic Regression Scales and Attributes*

Scale	Model	# of Turbines	Total Area (km²)
Macroscale: Statewide	L1	3,177	145,745
Mesoscale 1: 240 x 240 km	L2	2,145	57,600
Mesoscale 2: 160 x 160 km	L3	1,750	25,600
Microscale: 80 x 80 km	L4	340	6,400

Spatial Logistic Regression

Logistic regression is a frequently used mathematical modeling approach that can be used to describe the relationship of several variables to a dichotomous dependent variable. Spatial autocorrelation or spatial dependence measure similarity or dissimilarity measure between two values of an attribute that are spatially neighboring (Kissling, & Carl, 2008). Positive spatial autocorrelation tends to cluster in space, and observations of species distribution data often are inherently similar from nearby locations than would be expected on a random basis (Kissling, & Carl, 2008). Moran coefficient (Moran's I) quantifies the spatial autocorrelation measures, and the statistical value are between -1 and 1 when it takes a value of 0, the variable is randomly distributed rather than exhibiting a spatial pattern.

Spatial lag logistic regression incorporates spatial autocorrelation or spatial dependence to capture the influence of the variable on the regression (y). Therefore, spatial lag logistic regression methods are becoming more and more common procedures utilized to explore phenomena in various fields. Spatial lag modeling was utilized to analyze county level homicide rates and whether neighboring counties with high rates affected the neighboring county's crime rate (Baller, Anselin, Messner, Deane, & Hawkins, 2001).

The spatial lag model (SLM) is a general spatial autoregressive model in which explanatory variables include a spatial lag for the dependent variable as well as a set of factoring variables can be expressed as: (Baller et al., 2001).

$$y = (\rho)Wy + x(\beta) + \varepsilon$$

Wy = spatially lagged variable for weights matrix W

x = matrix of observations on the explanatory variable

ε = Vector of error terms

β and ρ are parameters (ρ) is the simultaneous autoregressive coefficient and quantifies the effect of neighbor observation and the direction of that effect

Spatial lag binary logistic regression requires dichotomous outcome so the existing turbine points (3,177) were assigned 1 and to fit model to data, "pseudo-absence" turbine points (3,177) were assigned 0. The next step extracted cell values from elevation, slope, land cover, and wind speed raster layers. Second, the distance to city/urban area, transmission line, river, highway, airport, and railroad were measured in kilometers. Iowa's 2010 census data was used for the population density km ^{2}. Population density was calculated for each census tract and converted to a raster for the Macroscale model M1 (statewide) and Mesoscale models L2, L3, and Microscale (six county region). Multiple scales are used to assess and determine the effect of scales on the importance of factor.

All the fields were normalized, and SPSS (version 15.0) was used for descriptive statics and binary logistic regression. Descriptive statistics were run as exploratory method to identify which explanatory variables matter most and gain statistical inset into the dataset. Correlations and multicollinearity diagnostics were run to examine linear or near linear relationships among the explanatory variables. Multicollinearity in the data causes statistical issues because it inflates the value of the least squares estimators and cause large errors in the output. The method used to check multicollinearity problems is the calculation of variance inflation factor (VIF). Collinearity analysis was conducted for the continuous environmental variables using linear regression in SPSS 15.0 statistical software and a VIF greater than 10 indicates the presence of strong multicollinearity. The categorical variable land cover was omitted from the test.

This research incorporated spatial lagged proximity to neighboring turbine (NT) explanatory variable as a stabilizing variable to reduce the variance in the model and captured spatial dependence of nearby turbine observations influence. Turbine points were assigned 1 if neighboring turbine was less than 500 meters and 0 if the neighboring turbine was greater than 500 meters. Models with spatial lag and models without were compared, and since models with spatial lag performed significantly better, the remaining models included the spatial lag variable (proximity to neighboring turbine). Binary logistic regression was run using the backward stepwise selection method to find out how all the independent variables (predictors) combined affect the dependent variable. All the independent variables (predictors) deemed important from literature were chosen. The complete list of variables included in the logistic regression with the dependent variable Y is shown in Table 6. Unstandardized residuals and probabilities were used to create the probability predictive map and residual map.

Table 6: *Logistic Regression Analysis Factors*

	Variable	Nature of variable
	Dependent	
Y	0 – no turbine; 1 – existing turbine	Binary
	Independent	
X_1	Slope (%)	Continuous
X_2	Wind Power Class (WPC)	Continuous
X_3	Elevation	Continuous
X_4	Land cover (LC)	Categorical
X_5	*Neighboring Turbine (NT)	Binary
X_6	* Transmission line (TL)	Continuous
X_7	* City	Continuous
X_8	*Highway	Continuous
X_9	*Railroad (RR)	Continuous
X_{10}	*Airport	
X_{11}	*River (River)	Continuous
X_{12}	*Highway (Hwy)	Continuous
X_{13}	^Population Density (Pop Den)	Continuous

Assessing Model Performance and Variable Contribution

Standard measure and model performance such as the coefficient of determination (R^2) and Standard Error of Estimate are not applicable for to logistic regression. Therefore, the highest Cox & Snell R square (pseudo R square measure range from 0 to 1) and Nagelkerke R square values were used to assess the amount of variation in the dependent variable explained by the model. "Percent Correct Prediction" statistics was used to assess how well the model predicted the correct category for each case.

Logistic models predictive accuracy and goodness-of-fit were tested, and classification criteria (excellent, good, fair, poor, and fail) were used to interpretation. Model's goodness-of-fit was based on the simultaneous of sensitivity (True positive) and specificity (True negative) for all possible cutoff points. Sensitivity is the percent correct prediction in the reference category of the dependent variable (i.e., 1 for binary logistic regression) and specificity; on the other hand is the percent of correct predictions in the given category of the dependent (i.e., 0 for binary logistic regression).

Model fit was determined according to Baldwin (2009) classification category: .9 – 1= excellent; .80-.90 = good; .70 - .80 = fair; .60 - .70 = poor and .50 - .60 = fail. Explanatory variables contribution or importance were determined using the Wald test with sig. < .05. The coefficient β was used to assess positive or negative relationship (which explanatory variable increase the likelihood of turbine occurrence and which factors decreases it), and the odds ratio EXP (B) indicates change in odds [P(event)/P(no event)] of outcome which resulted from a unit change in the predictor.

If Exp (B) > 1, as predictor increases, odds of outcome increases; positive influence

If Exp (B) < 1, as predictor increases, odds of outcome decrease; negative influence

Model diagnostics was conducted using the residual verses fit plots. The residuals were interpolated in ArcGIS 10.0 and maps were created. The maps were used to detect the correlation between the predicted and residual values by observing the spatial distributions of the negative high residuals and the positive high residuals.

Residual is the difference between the observed value of the dependent variable (y) and the predicted value ($\hat{y}$) forms the residual (e) for each data point ($e = y - \hat{y}$).

Suitability Map

ArcGIS 10.0 Spatial Analyst Interpolation Kriging technique was used to create raster surface from the probabilities and residuals. The Kriging method of interpolation is based on regression against observed z values of surrounding data points, weighted according to spatial covariance values. In order to standardize and compare binary logistic regression and Maxent output surface, cell sizes were set 200 m for Macroscale and 30 meters for Mesoscale 1, 2 and Microscale. All scales suitability map were produced and thorough interpretation of each map was conducted. In doing so, various patters were observed at all scales. Finally, the suitability surface grid from logistic regression was compared to the Maxent suitability surface grid to identify similarities and differences.

Machine-Learning Algorithm

Maxent is an open source and the most commonly utilized Ecological Niche Modelling program (Baldwin, 2009; Elith, et al., 2011; Phillips, Anderson, Schapire, 2006; Phillips, Dudik, & Schapire, 2004). Maxent has traditionally been used to model (predict) the species spatial distribution in a geographic space. It's based on machine learning algorithm designed to make the prediction of species spatial distribution given known environmental characteristics (Baldwin, 2009; Elith et al., 2011; Phillips et al., 2004, 2006). Maxent estimates the most uniform distribution (maximum entropy) of sampling points compared to background locations given the environmental constraints (Baldwin, 2009). The maximum entropy algorithm is deterministic and will converge to maximum probability distribution (Baldwin, 2009; Phillips & Dudik, 2008).

Maxent relies on an unbiased sample; therefore, its critical collecting comprehensive set of presence record (cleaned from errors and duplicates) and dealing with biases are critical. Maxent has an advantage since it allows both continuous and categorical variables; therefore, the output result tends to represent better model fit (Baldwin, 2009). The average value of each environmental variable at the occurrence locations serves as the target value for the probability distribution (Petrov & Wessling, 2014).

Ecologists primarily utilize this program to model species distribution from presence-only records and with associated environmental variables deemed essential for the presence of the species over the study area (Elith et al., 2011). However, Maxent is gaining traction in other areas such as habitat suitability modelling and wind energy site suitability. Petrov and Wessling (2014) utilized Maxent to study site suitability for wind energy development in Iowa at the two scales. At the Macroscale, wind power class and elevation contributed 51.2% and 32.6% respectively. While at the Mesoscale, elevation contributed 79% to the suitability distribution. Parisen and Mortize (2009) used Maxent to identify locations most at risk for wildfires in California while Baldwin (2009) utilized Maxent to identify areas suitable for red spruce forest habitat to better protect the species habitat and to incorporate this model into the restoration plan. Ultimately, Maxent was chosen for this study because it has several advantages to building a suitability model: 1. it only requires presence or occurrence data, 2. it utilizes categorical and continuous environmental predictors, 3. it creates outputs which allow interpretation of contribution of predictors to the model (Phillips et al., 2006).

Environment Variables and Procedure

The environmental variables used are elevation, slope, wind power class, land cover, population density, and distance to airport, river, city (urban areas), transmission line, highway, road, and river. These variables were chosen because as previous studies have shown they are important determinants to site suitability for wind energy development (Acker et al., 2007; Griffiths & Dushenko, 2011; Rodman & Meentemeyer, 2006). Environmental layer were converted to ASCII raster format, common cell size, and extent. Eleven environmental variables were selected for their potential importance based on knowledge and from previous studies.

One of the objectives of this study was to determine whether geographical scale affects the importance of variable and to compare the Maxent output to the binary logistic regression output. Table 7 displays the models used to capture the effects of scale on environmental variables. Scale maps produced to aid in visual aid are provided in Figure 17. The grids for the Mesoscale models M2, M3, and Microscale M4 were created based on the geographical mean center of the turbine distribution in Iowa.

Table 7: *Maxent Scale Attributes*

Scale	Model	# of Turbines	Total Area (km²)
Macroscale: Statewide	M1	3,177	145,745
Mesoscale 1: 240 x 240 km	M2	2,145	57,600
Mesoscale 2: 160 x 160 km	M3	1,750	25,600
Microscale: 80 x 80 km	M4	340	6,400

Macroscale (statewide) layers were resampled to a 200 m cell size in order to increase model processing speed, and processing extent was set using the state of Iowa boundary feature. Mesoscale models M2, M3, and Microscale M4 layers were resampled to 30 m cell sizes and each model's extents were set accordingly. Environmental layers were specified as a continuous except for the land cover layer which was classified as categorical layer. Maxent requires that 'species' presence (occurrence) dataset be in a commonly separated file (*csv). The 'species' presence of turbines with latitude and longitude was created for all models.

Eighty percent of the turbine presence records were used as training data set for all scales. Maxent's built-in method was used to validate the accuracy of the predictive distribution by setting aside 20% randomly selected presence records. To focus on critical features of the model and to avoid over fitting, default regularization options were used in linear quadratic product (.050), categorical (.250), threshold (1.00), and hinge (.500). Ten replicated runs were performed for each scale. 'Species' distribution surface for all scales were imported in ArcGIS 10.0 and suitability maps were produced. In addition, variable contribution table and ROC curve graphs provide by Maxent.

<u>Variable Response and Model Evaluation</u>

The modeled distributions were evaluated using the area under the receiver operating characteristic (ROC) curve plots of (AUC). Sensitivity represents how well the data accurately predicts presence; whereas, specificity provides a measure of correctly predicted absences (Baldwin, 2009). The significance of the curve is quantified by the area under the curve (AUC) and has values that range from 0.5 to 1.0. Values close to 0.5 indicate random prediction while a value of 1 indicates a perfect fit. Classification category: .9 – 1= excellent; .80-.90 = good; .70 - .80 = fair; .60 - .70 = poor and .50 - .60 = fail (Baldwin, 2009).

Maxent outputs an estimate of relative contributions of the environmental variables to the model which identifies the most important and influential variables and jackknife which assess the usefulness of variables when run alone. These outputs were used to identify environment variable influence to the presence of the modeled turbine, and their relative importance was determined using the percent contribution table output for all scales. In addition, the logistic output as an ASCII file format was in ArcGIS 10.0 to create probability of suitable/unsuitable distribution surface map. The logistic output format was selected because it allows for easier and potentially more accurate interpretation over other formats, and the logistic format is recommended given that it provides estimates of the probability of occurrence as predicted by included environmental variables (Baldwin, 2009).

3.6 Normative Module

Previous studies by Baban and Parry (2001); Rodman and Meentemeyer (2006); and Aydin et al. (2010) models overlay criteria and constraints using a fuzzy weighting scheme and incorporated normative models as a stand-alone model to identify suitable site. Mann et al., (2012) included weighted normative component as part of their mixed suitability model. Generally, hybrid models incorporate traditional normative weighting approaches is to assign weights to the criteria layers in addition to those identified by the empirical model, and map algebra was used to aggregate all factors to create a final wind energy suitability surface of Iowa. The shortcoming of the normative component is the arbitrary weighting of the factors based on presumed favorable characteristics for wind farm development. However, this study developed normative components to incorporate with the empirically driven spatial lag model as a final hybrid predictive model which was the basis for the framework.

First, federal regulation prohibits infrastructure built taller than 60.96 meters from the nearest civilian and military airports be at least 2,500 meters. Second, in literature and industry, standards have identified a minimum of 1,000 meters away from the nearest city or urban settlement. These two variables were chosen because they are definitively clear and can be incorporated into the model without subjectivity.

3.7 Study Flowchart

Figure 19 outlines the work flowchart of this study and outlines the methodologies used to develop the framework. First, siting factors are identified from the extensive literature review and data collection and processing based on identified factors are compiled. Spatial lag regression and machine-learning algorithm (Maxent) are identified and multiscale analyses are performed. Empirical coupled with normative components are used to develop the integrated spatially explicit scale dependent framework. Case study area are identified and logistic regression probability formula were used to validate site suitability.

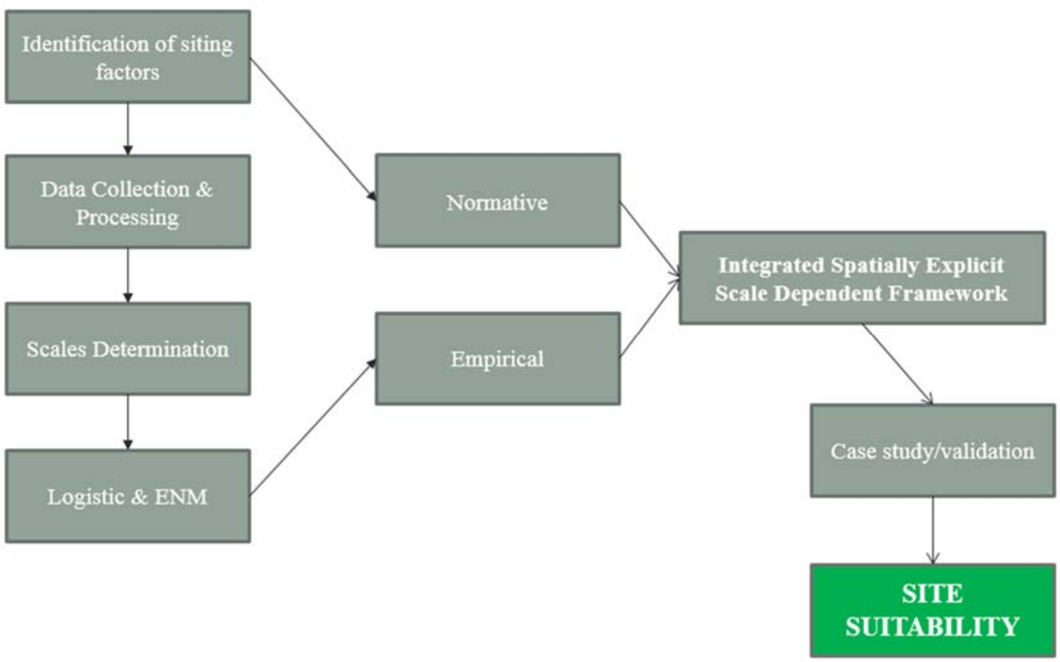

Figure 19: Study flowchart

CHAPTER 4

RESULTS

4.1 Exploratory Data Analysis

The thirteen factors used in the logistic regression were slope, elevation, wind power class, land cover, population density and distance to transmission line, highway, airport, river, and city. Eighty-eight percent of turbines are on land cover type 3 classification (cropland and grassland), and another 10.8% turbine points are in areas classified as forest. The spatial lag factor (independent variable) proximity to neighboring turbine (NT) displayed 55.7% of turbines as not having a neighboring turbine within proximity (< 500 m) while the remaining 44.3% of turbines were classified as having a turbine within the 500 m proximity distance. NREL identified WPC 3 (5.6-6.0 m/s) or greater as minimal requirement for utility scale site wind farm development; and in this case, 98.6% of turbines in Iowa were located in areas with WPC 3 or greater.

Multicollinearity diagnostic was performed, and the variance inflation factor (VIF) and tolerance parameters were used to determine collinearity of factors. VIF and tolerance are both widely used measures of the degree of multicollinearity of factors (independent variable) with the other factors (O'Brien, 2007). There is a wide range of cut off VIFs one can come across in literature, but a cut-off value often used is 4 which has a theoretical basis such that standard errors are doubled at this point thus making it an ideal cut-off point. VIF greater than 4 and tolerance less than 0.10 are indication of multicollinearity in this study. Therefore, since all factors have VIF less than 4 and the tolerance greater than 0.10, collinearity is not a problem in this analysis (Table 8).

Collinearity diagnostics analyses for model L2, L3, and L4 exhibited similar outcomes as well.

Table 8: *Factors multicollinearity diagnostics*

Environmental variable	**Tolerance**	**VIF**
Slope	.733	1.365
Wind power class	.400	2.504
Elevation	.457	2.186
Land cover	.983	1.017
Proximity to neighboring turbine	.669	1.495
Transmission line	.908	1.102
City	.788	1.270
Highway	.799	1.252
Railroad	.791	1.264
Airport	.822	1.215
River	.934	1.070
Population Density	.926	1.080
Tolerance > 0.1 & VIF < 4 = no collinearity Slope (%), Wind Power Class, Elevation (m), Land cover, Proximity to neighboring turbine (NT); Distance (km): transmission line (TL), City, Airport, River, Railroad (RR), highway (Hwy); Population Density (Pop Den) km^2		

The correlation analysis was performed on all factors excluding land cover and neighboring turbine since both are categorical variable (Table 9). Correlation result identifies relationships between predictors (factors) that were expected; however, there are a number of variable relationships worth highlighting. Elevation displayed a strong positive correlation ($r = 0.655$) with wind power class (WPC) which illustrates wind power class will increase as elevation increases. On the other hand, WPC had moderate negative correlation with slope indicating the WPC decreases as slope increases which is not surprising outcome since literature identifies suitable areas with low percent slope. Correlation analyses for models L2, L3, and L4 displayed similar outcome as well.

Table 9: *Correlation table of predicating factors of wind turbine*

	Slope	WPC	Elevation	TL	City	Hwy	RR	Airport	River	Pop Density
Slope	1									
WPC	-.476**	1								
Elevation	-.193**	.655**	1							
TL	.111**	-.128**	-.041**	1						
City	-.010	.080**	.145**	.086**	1					
Hwy	-.038**	.089**	.148**	.090**	.408**	1				
RR	.138**	-.225**	.011	.252**	.187**	.165**	1			
Airport	-.055**	.173**	.258**	.105**	.190**	.231**	.229**	1		
River	-.060**	.213**	.227**	.031*	-.008	.001	.000	.074**	1	
Pop Density	.062**	-.100**	-.119**	-.054**	-.193**	-.106**	-.097**	-.170**	-.041**	1

** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

4.2 Spatial Logistic Regressions

The general description of the empirical models (L1, L2, L3, and L4) are given in section 3.5 (“Empirical Model”). The results illustrated in this section takes a comparative approach between the standard logistic regression and the spatially lagged logistic regression. All spatial lag models contained thirteen factors (slope, wind power class, elevation, land cover, proximity to neighboring turbine, transmission line, city, highway, railroad, airport, river, and population density) while the standard regression models excluded the spatial lag factor (NT).

Standard Logistic Regression Model L1 (Macroscale)

The standard logistic regression Model L1, a full model containing all predictors (except NT) was statistically significant:

$$x^2\ (13, N=6{,}354) = 3{,}640.615,\ p < .0005$$

indicating that the model was able to distinguish between predictors that influenced turbine occurrence and nonoccurrence (Table 10). The model as a whole explained between 43.6% (Cox & Snell R square) and 58.2% (Nagelkerke R square) of the variance and 82.4% of cases were correctly classified validating the model. WPC displayed the strongest positive impact on the model with an odds ratio of 1.722 indicating turbine occurrence is 1.722 more times likely with the WPC factor included in the model given all other variables stay constant. Elevation, distance to city and airport also exhibited a positive odds ratio increasing the likelihood of turbine occurrence. Slope, population density and distance to transmission line indicated a negative relationship signifying the high slope gradient, high population density and greater distance to the transmission lines with the likelihood of turbine occurrence decreases by 0.947, 0.992 and 0.854 restrictively.

Spatial Lag Logistic Regression Model L1.1

In contrast, the spatial lag model L1.1 a full model containing all predictors was statistically significant:

$$x^2 \text{ (15, N=6,354)} = 6{,}593.73,\ p < .0005$$

indicating that the model was able to distinguish between predictors that influenced turbine occurrence and predictors that did not. The model as whole explained between 64.6% (Cox and Snell R square) and 86.1% (Nagelkerke R square) of the variance in turbine status and 92.6% of the cases were correctly classified validating the model. Nine of the factors (predictors) were statistically significant thus impacting the overall model to determining turbine occurrence (Table 10).

The strongest predictor of turbine occurrence was wind power class (WPC) which displayed a positive relationship with the dependent variable (occurrence of turbine) with an odds ratio of 1.711. This indicates that turbine occurrence is 1.711 times more likely with WPC predictor included in the model given all other predictors stay constant. Also, elevation with an odds ratio of 1.017, city with an odds ratio of 1.157, and airport with an odds ratio of 1.1019 are strong predictors of turbine occurrence as well. This suggests that high wind power class, high elevation and distance away from city and airport are parameters more likely associated with the turbine occurrence. Land cover 3 (cropland and grassland) and land cover 5 (barren land) classification were statistically significant confirming that turbines are more likely to be located in cropland and barren land relative to other land use in Iowa. On the other hand, slope had a negative relationship to the dependent variable with an odds ratio of 0.967 indicating that for every one unit increase in percentage slope, the occurrence of turbine is 0.967 times less likely controlling for other factors in the model. Furthermore, distance to transmission line and highway also displayed negative coefficient implying that nearness to these infrastructures is associated with a higher likelihood of turbine occurrence.

Model L1.1 identified wind power class, elevation, and distance to city as the most influential variable in addition to the spatially lagged variable proximity to neighboring turbine (NT). These variables exhibited strong positive impact on the model in comparison with the standard regression. Variables like slope, distance to transmission line, highway, and river showed negative relationship with the dependent variable. While significant variables in the standard regression (Model L1) identified similar variables with weaker odds ratio strength and the overall performance of the model exhibited significantly lower model fit as well. Spatial lag Model L1.1 Nagelkerke R Square (0.861) substantially improved from Model L1 higher classification indicating good fit in comparison to the standard logistic regression which had lower Nagelkerke R square and classification.

Spatial lag model L1.1 overall performance and model fit justified the selection of the spatially lagged model as the most robust and accurate model to use. The model's goodness-of-fit was determined by examining the receiver operator characteristics (ROC) area under the curve of 0.979 with 95% confidence internal (0.977, 0.982) indicated an excellent fit. The area under the curve is significantly different from 0.5 and since p-value is < 0.0005, the spatial-lag logistic regression classified (0.979) the group significantly better than by chance.

Table 10: *Macroscale Logistic Regression*

	Model L1*			Model L1.1** (spatial lag)		
	B	**Sig.**	**Exp(B)**	**B**	**Sig.**	**Exp(B)**
Slope	-.054	.000	.947	-.034	.000	.967
WPC	.543	.000	1.722	.537	.000	1.711
Elevation	.023	.000	1.024	.017	.000	1.017
LC	n/a	.000	n/a	n/a	.005	n/a
LC(1)	-21.548	.998	.000	-20.417	.999	.000
LC(2)	-2.877	.002	.056	-1.514	.187	.220
LC(3)	-.558	.110	.573	-1.007	.029	.365
LC(4)	-.359	.295	.698	-.779	.082	.459
LC(5)	-1.628	.001	.196	-2.553	.000	.078
NT	n/a	n/a	n/a	-5.518	.000	.004
TL	-.158	.000	.854	-.196	.000	.822
City	.073	.000	1.076	.146	.000	1.157
Hwy	n/a	n/a	n/a	-.065	.006	.937
Airport	.081	.000	1.085	.103	.000	1.109
River	n/a	n/a	n/a	-.038	.085	.963
Pop Density	-.008	.001	.992	-.006	.056	.994
Constant	-10.922	.000	.000	-4.857	.000	.008

* Nagelkerke R square 0.582 ** Nagelkerke R square 0.861

The Interpolated map of the probability and residual outputs from the spatial logistic regression have identified areas of suitability with proximity to neighboring turbines. Spatially lagged model L1.1 probability predictive map displays areas with brightest color indicating high probability that conditions are suitable for turbine occurrence (Figure 20). Areas with probability 0.41 – 0.6 highlight areas that possess conditions suitable for turbine occurrence primarily due to the influence of neighboring turbine factors. Figure 21 displays the residuals output which indicate how well the model over or under estimates and in this case, the model under estimates. In this case, negative high residuals (green) indicate areas where conditions are suitable for turbine to exist, but turbines don't exist while positive high residuals (red) indicate where turbine currently exist but might not be suitable sites for the turbines.

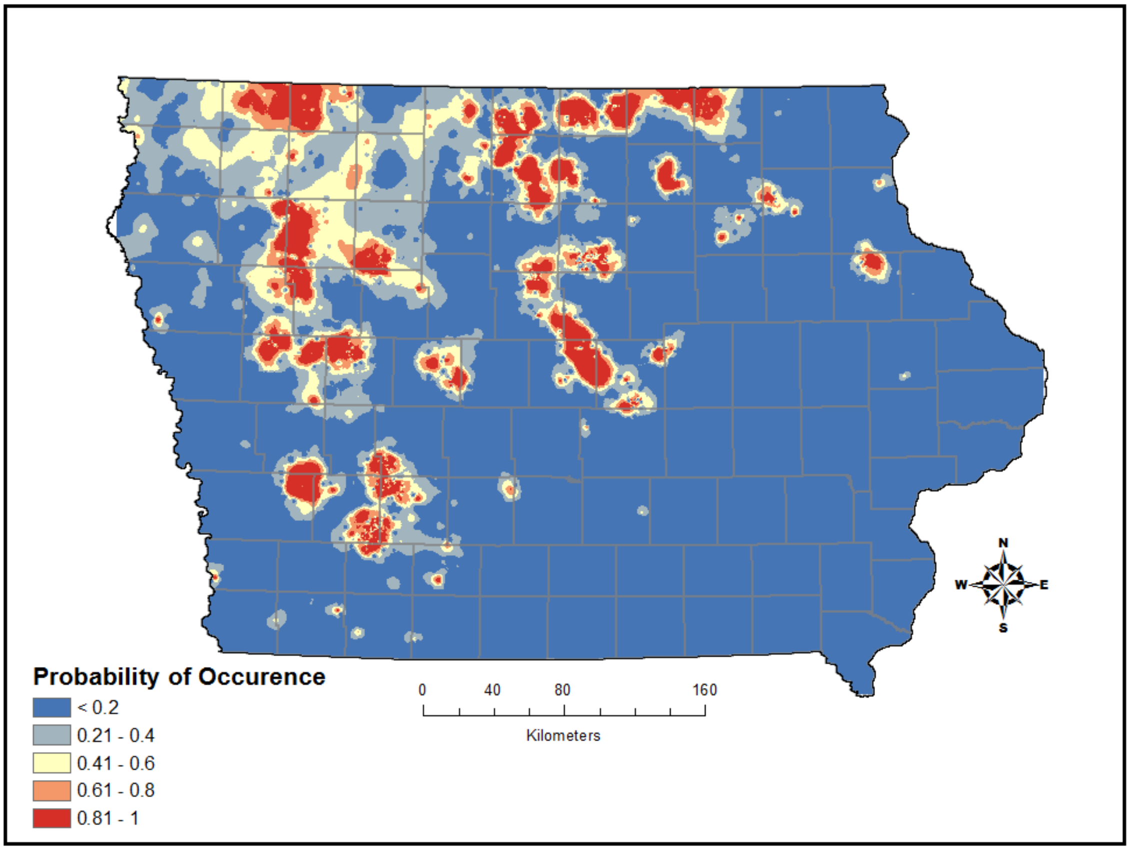

Figure 20: Spatial lag model L1.1 suitable sites for wind turbine occurence

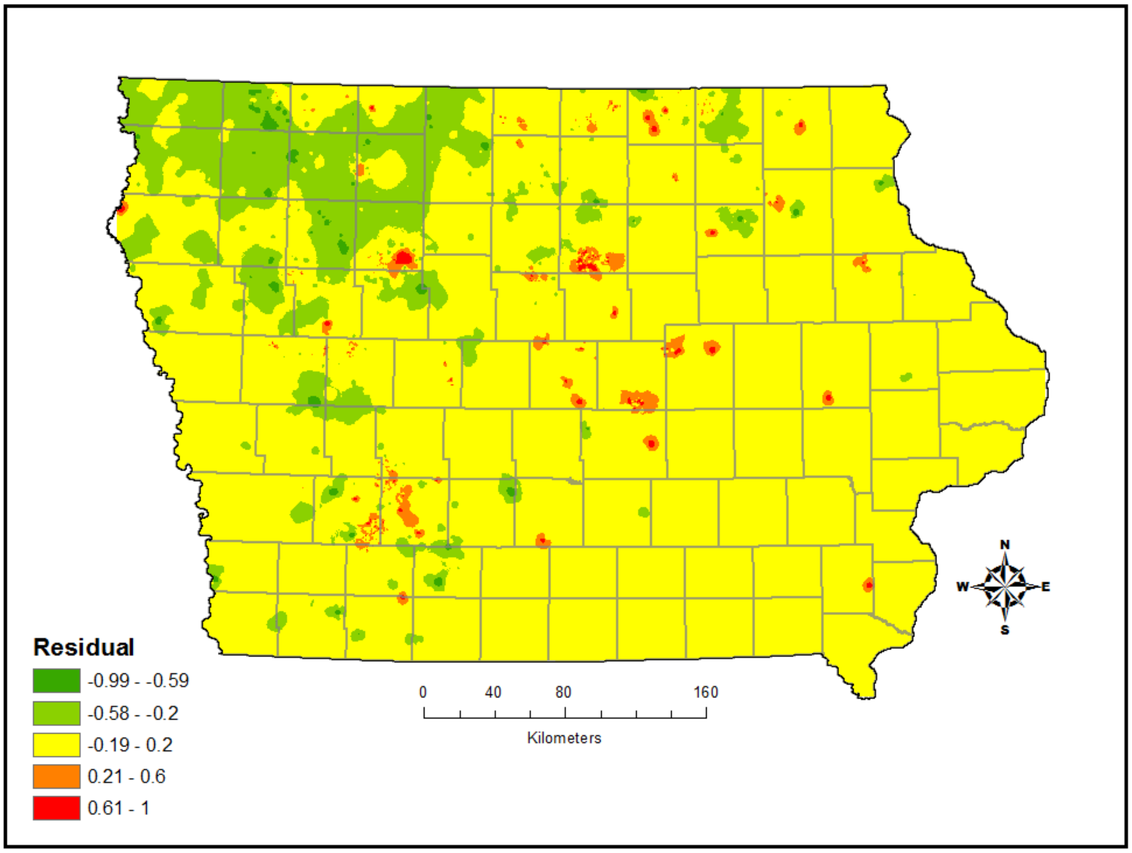

Figure 21: Spatial lag model L1.1 residual diagnostics

Standard Logsitic Regression Model L2 (Mesoscale 1)

Model L2 containing all predictors (except NT variable) was statistically significant:

$$x^2 \text{ (10, N=4,289) = 1625.561, p < 0.000}$$

indicating that the model was able to distinguish between predictors that affected turbine occurrence and those that did not (Table 11). As a result, the model as a whole explained between 31.5% (Cox & Snell R square) and 42.1% (Nagelkerke R square) of the variance in turbine occurrence status. Model L2 also classified 75.4% of cases correctly and the

ROC area under the curve of 0.785 with 95% confidence interval (0.771, 0.799) indicates a good fit. Wind power class (WPC) displayed the strongest positive impact on the model with an odds ratio of 1.516 therefore indicating turbine existence is 1.516 times more likely with this variable included in the model given all other variables stay constant. Distance to city, river, and airport displayed positive relationships which indicates the likelihood of turbine occurrence to be higher when distance from these factors increases. In contrast, distance to transmission line, population density, and highway showed negative relationships. In this case, each illustrated the odds ratio of turbine occurrence to be higher when the distance to transmission line and highway is shorter and population density is lower

Spatial Lag Logistic Regression Model L2.1

In comparison, spatial lag model L2.1 containing all predictors was statistically significant:

$$x^2\ (12, N=4{,}289) = 3{,}931.626,\ p < 0.000$$

indicating that the model was able to distinguish between independent variables that affected turbine occurrence and those that did not. The model as a whole explained between 60% (Cox and Snell R Square) and 80% (Nagelkerke R square) of the variance in turbine status and correctly classified 90.4% of cases (Table 11).

The strongest predictor of turbine occurrence was WPC recording an odds ratio of 2.11 indicating that turbine existence would be 2.11 times more likely if it has high WPC given all other factors stay constant in the model. While proximity to neighboring turbine (NT) indicating that for every km closer to neighboring turbine, the more likely if it has a neighboring turbine given all other factors stay constant in the model. Wind power class, elevation, distance to city and airport showed positive relationships with the turbine occurrence. Distance to transmission line, highway and population density exhibited negative relationships. The shorter the distance to transmission line and highway, the more likelihood of turbine occurrence while the lower the population in the area, the more likelihood of turbine occurrence as well.

The spatial lag goodness-of-fit was determined by examining the receiver operator characteristics (ROC) curve and the area under the ROC curve of 0.957 with 95% confidence internal (0.951, 0.963) indicated an excellent fit (Figure 22). The area under the curve is significantly different from 0.5 and the p-value < 0.000 means that the spatial lag model L2 regression classifies the group significantly better than by chance. The result from model L2.1 spatially lagged and standard logistic regression reveals the similar outcomes, but the overall improved model performance was exhibited in the spatially lagged model.

Table 11: *Mesoscale 1 Logistic Regression*

	Model L2*			Model L2.1** (spatial lag)		
	B	**Sig.**	**Exp(B)**	**B**	**Sig.**	**Exp(B)**
Slope	-.054	.000	.948	n/a	n/a	n/a
WPC	.416	.000	1.516	.747	.000	2.111
Elevation	.024	.000	1.024	.019	.000	1.019
LC	n/a	n/a	n/a	n/a	.017	n/a
LC(1)	-20.632	.997	.000	-19.148	.998	.000
LC(2)	-.444	.325	.642	-1.128	.049	.324
LC(3)	-.342	.442	.711	-1.090	.051	.336
LC(4)	-1.849	.002	.157	-2.720	.001	.066
LC(5)	-.054	.000	.000	-4.868	.000	.008
NT	n/a	n/a	n/a	-.147	.000	.863
TL	-.122	.000	.885	.178	.000	1.195
City	.131	.000	1.140	-19.148	.000	.922
Hwy	-.061	.006	.940	-.082	.001	.000
Airport	.033	.000	1.034	.073	.000	1.075
RR	97.692	.000	2.674	n/a	n/a	n/a
Pop Density	-.015	.056	.985	-.012	.046	.988
Constant	-10.762	.000	.008	-7.438	.000	.001

* Nagelkerke R square 0.421 ** Nagelkerke R square 0.8

Figure 22 displays the spatial lag model L2.1 suitability distribution map. Areas with high probability display where turbines exist; therefore, areas are heavily influenced by proximity to neighboring turbines. The residual diagnostic indicates the model under estimates (Figure 23). Areas with high negative residuals (green) highlight areas where the condition are right but no development has occurred and thus should be considered as future sites for development.

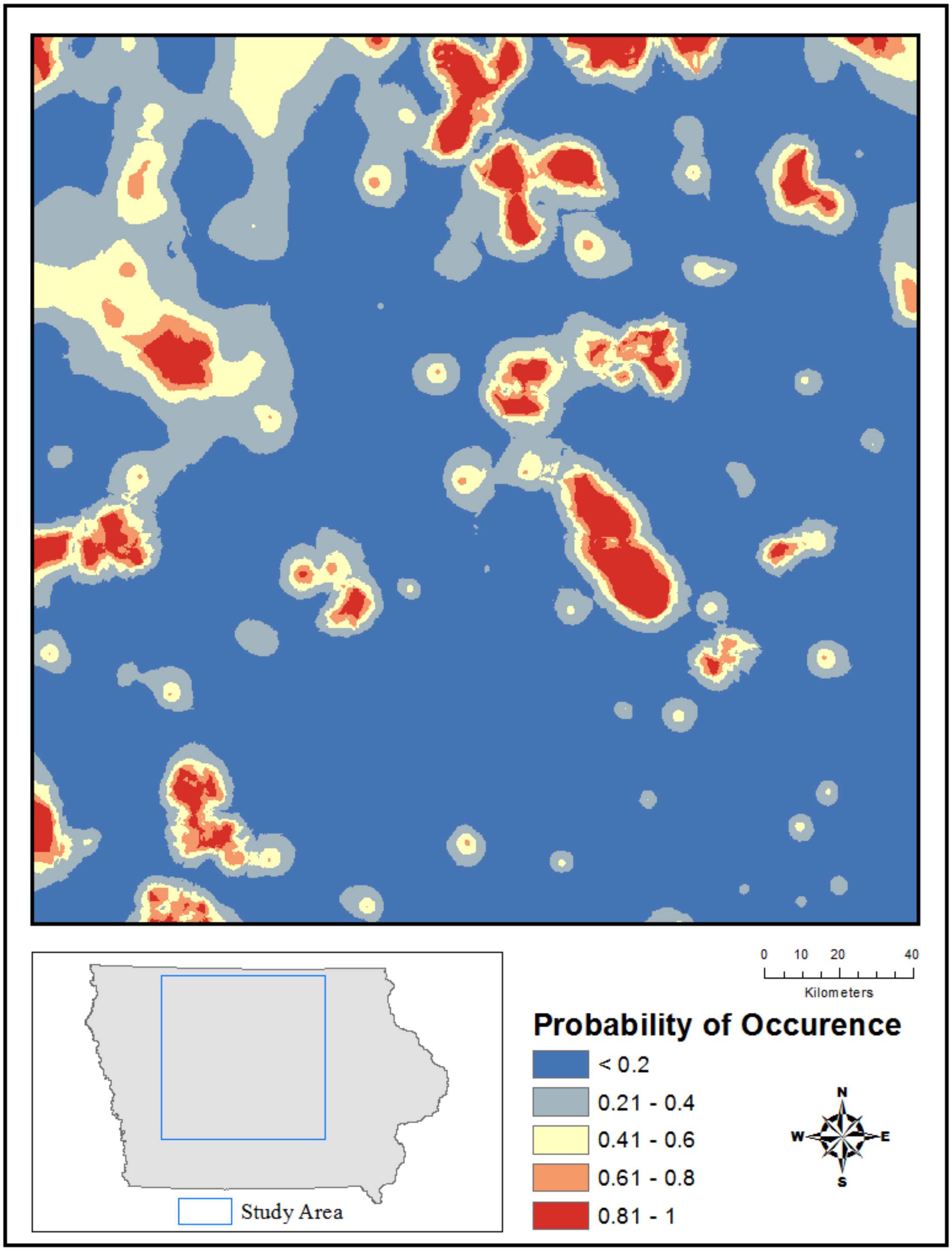

Figure 22: Spatial lag model L2.1 suitable sites for wind turbine occurrence

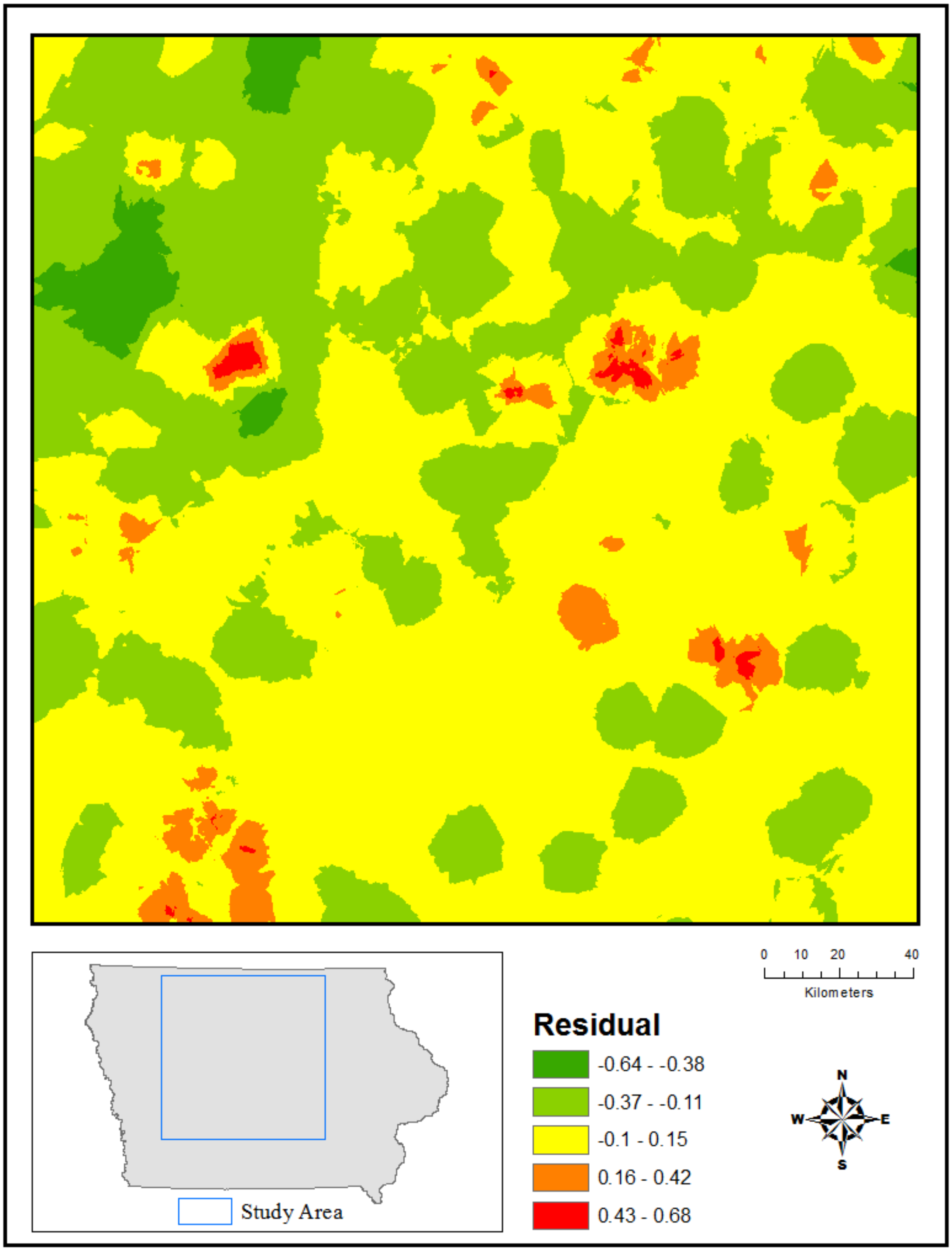

Figure 23: Spatial lag model 2.1 residual diagnostics

Standard Logistic Regression Model L3 (Mesoscale 2)

Model L3, the standard logistic regression full model containing all predictors, was statistically significant:

$$x^2 \text{ (11, N=1,750) = 829.267, p < 0.001}$$

indicating that the model was able to distinguish between predictors that affected turbine occurrence and those that didn't. As a result, the model as a whole explained between 37.7% (Cox & Snell R square) and 50.3% (Nagelkerke R square) of the variance in turbine occurrence status (Table 12). Compared to previous standard logistic regression, model L3 showed significant drop, but compared to (L2), it appears to improve. Model L3 also classified 78.6% of cases correctly but a decline in correct classification continues from previous standard logistic regression models (L1 and L2). The ROC area under the curve of 0.865 with 95% confidence interval (0.848, 0.881) suggests a good fit. WPC showed the strongest positive impact with an odds ratio of 2.344 indicating turbine existence is 2.344 more times likely with this variable included in the model given all other variable stay constant. Elevation, distance to river, city, and airport showed positive influence so the higher the elevation, the further away from river, city, and airport would make the odds of turbine existence more likely, given that all other variables (factors) stay constant. Again, distance to transmission line and population density displayed negative relationships.

Spatial Lag Logistic Regression Model L3.1

On the other hand, spatial lag model L3.1 was statistically significant:

$$x^2 \text{ (11, N=1,750)} = 1{,}582.344,\ p < 0.001$$

indicating that the model was able to distinguish between predictors that affected turbine occurrence and predictors that did not. The model as a whole explained between 59.5% (Cox and Snell R square) and 79.4% (Nagelkerke R square) of the variance in turbine status and correctly classified 90.6% of cases (Table 12). Neighboring turbine, wind power class, elevation, distances to river, city, airport, and transmission line made a unique statistical significant contribution to the model. The strongest predictor of turbine occurrence was WPC recording an odds ratio of 3.284 signifying turbine occurrence 3.284 more times likely with WPC given that all other factors stay constant. Also elevation, distance to city, airport, and river have a positive impact and improve the likelihood of turbine occurrence when elevation is higher and the distances to city, airport and river are further away from the potential site. Proximity to neighboring turbine also increases the likelihood of turbine occurrence as well. Distance to transmission line displayed negative relationship to the dependent variable and similar outcomes are observed in previous models.

The model's goodness-of-fit was determined using the receiver operator characteristics (ROC) area under the curve. The ROC area under the curve of 0.963 with 95% confidence internal (0.954, 0.971) indicates an excellent fit. The area under the curve is significantly different from 0.5; and since p-value is < .000, this means that the spatial lag regression classifies the group significantly better than by chance.

Table 12: *Mesoscale 2 Logistic Regression Analysis*

	Model 3*			Model L3.1**(spatial lag)		
	B	**Sig.**	**Exp(B)**	**B**	**Sig.**	**Exp(B)**
WPC	.852	.000	2.344	1.189	.000	2.111
Elevation	.022	.000	1.022	.020	.000	1.019
LC	n/a	.043	n/a	n/a	.033	n/a
LC(1)	-20.617	.999	.000	-19.754	.999	.000
LC(2)	-1.205	.019	.300	-2.210	.003	.110
LC(3)	-1.234	.013	.291	-2.078	.004	.125
LC(4)	-2.978	.004	.051	-4.643	.021	.010
NT	n/a	n/a	n/a	-4.693	.000	.009
TL	-.255	.000	.775	-.322	.000	.725
City	.267	.000	1.305	.317	.000	1.373
Airport	.102	.000	1.107	.114	.000	1.121
RR	-.046	.012	.955	n/a	n/a	n/a
River	.234	.000	1.264	.215	.000	1.240
Constant	-12.946	.000	.000	-10.219	.000	.000

* Nagelkerke R square 0.503   ** Nagelkerke R square 0.794

The Figure 24 probability suitability map displays areas with high WPC and proximity to existing turbines indicating high probability of turbine occurring. In contrast, it appears that the further away from existing turbines, the probability of turbine occurrence decreases. The residual diagnostic showed slight underestimation by the model (Figure 25). Outcomes were similar as previous models, where negative high residuals, have indicated the right conditions for turbines no yet developed.

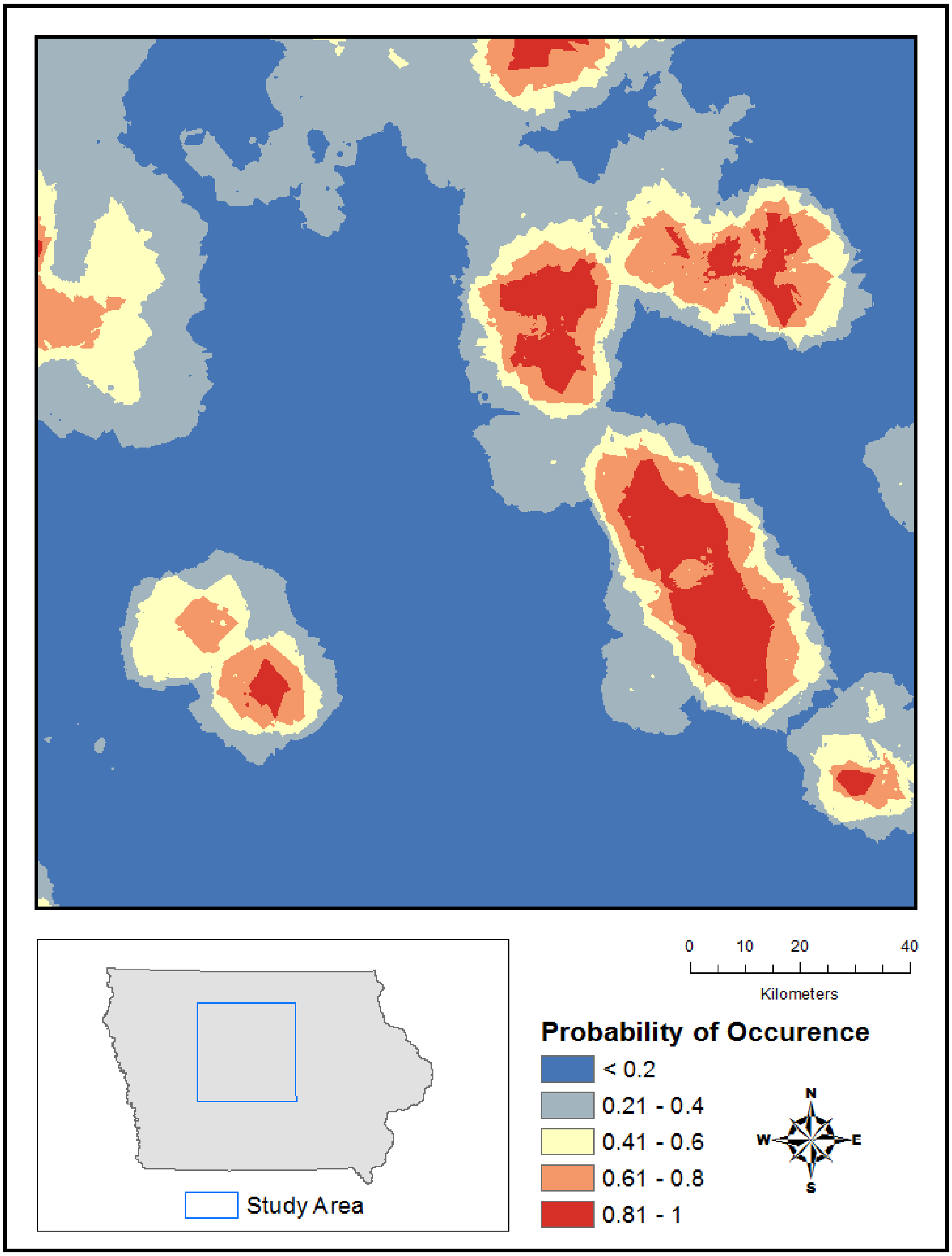

Figure 24: Spatial lag model L3.1 suitable site for wind turbine occurrence

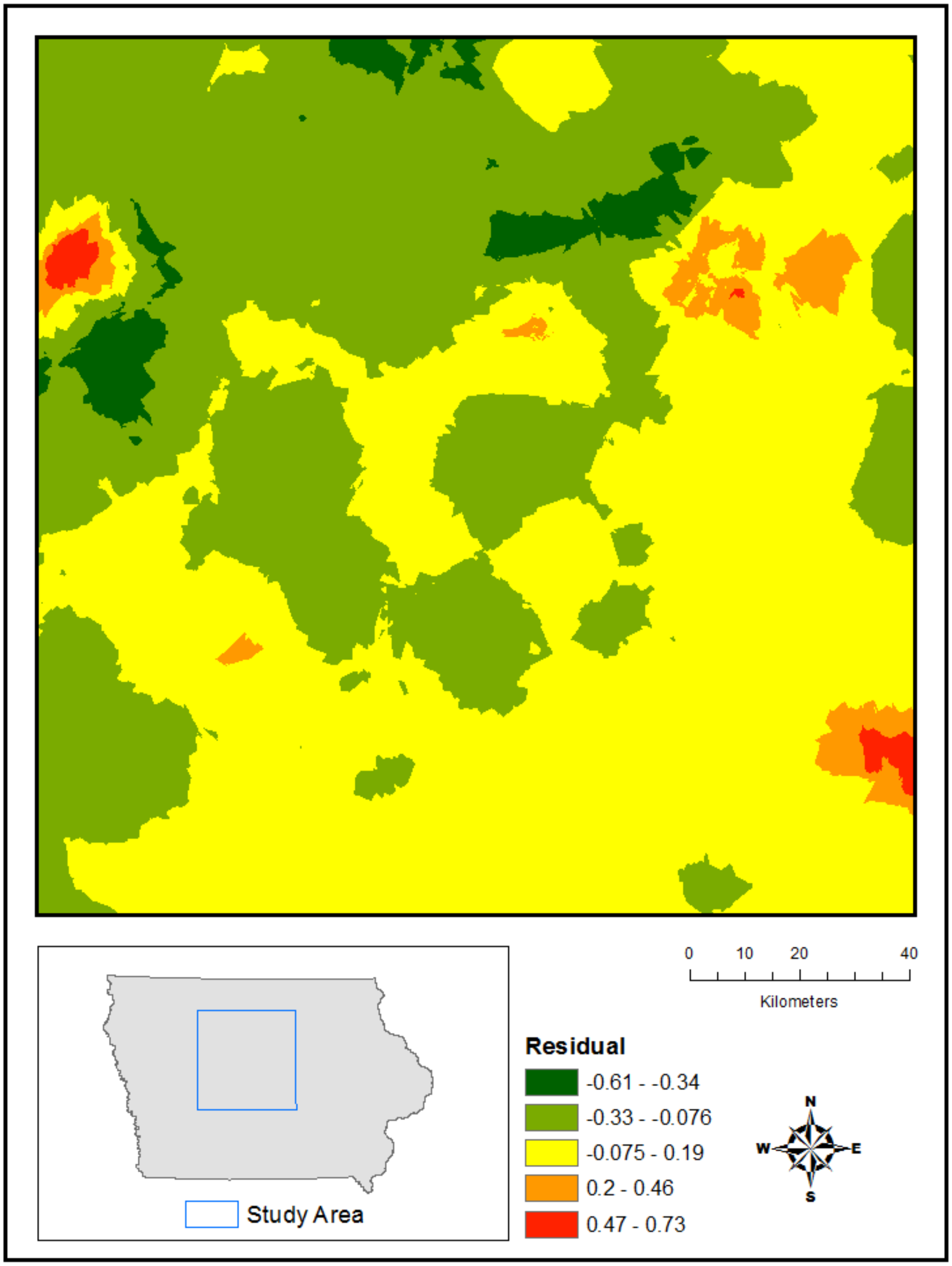

Figure 25: Spatial lag model L3.1 residual diagnostic

Standard Logisitc Regression Model L4(Microscale)

Model L4 was statistically significant:

$$x^2 \text{ (5, N=740) = 395.15, p < .001}$$

indicating that the model was able to distinguish between predictors that affected turbine occurrence and those that didn't. As a result, the model as a whole explained between 41.4% (Cox & Snell R square) and 55.2% (Nagelkerke R square) of the variance in turbine occurrence (Table 13). Compared to previous non-spatial lag model (L1), a significant drop is displayed, but compared to model L2 and L3, it appears to improve. Model L4 also classified 80% of cases correctly but a decline in correct classification continues from previous model L1 and L2 but it's an improvement from model L3. ROC area under the curve of 0.879 with 95% confidence interval (0.855, 0.904) is a good fit. Elevation, distance to river and city displayed positive relationships. In this case, distance to river, the strongest influence, with an odds ratio of 1.353 indicated that the further away from river, the likelihood of turbine occurrence is 1.353 more likely.

Spatial Lag Logistic Regression Model L4.1

Spatial-lag model L4.1 (Microscale) was statistically significant:

$$x^2 \text{ (6, N=740) = 656.08, p < .001}$$

indicating that the model was able to distinguish between predictors that affected turbine occurrence and predictors that did not. The model as a whole explained between 58.8% (Cox and Snell R square) and 78.4% (Nagelkerke R square) of the variance in turbine status and correctly classified 92.7% of cases (Table 13). An improvement from previous models (L1.1, L2.1, L3.1). The strongest predictor variable to incase the likelihood of turbine occurrence appears to be distance to river and city recording an odds ratio of 1.328 and 1.155 respectively. Elevation indicates high likelihood of turbine occurrence at a higher elevation. WPC and distance to TL appears to be insignificant at this scale. ROC area under the curve of 0.960 with 95% confidence internal (0.944, 0.975) indicates an excellent fit. The results display similar trends as in previous models where the spatially lagged models performed significantly better than the standard regression.

Table 13: *Microscale Logistic Regression Analysis*

	Model 4*			Model L4.1** (spatial lag)		
	B	**Sig.**	**Exp(B)**	**B**	**Sig.**	**Exp(B)**
WPC	20.766	.995	251040752	19.341	.995	251040752
Elevation	.052	.000	1.063	.061	.000	1.063
NT	n/a	n/a	n/a	-4.036	.000	.018
TL	n/a	n/a	n/a	-.148	.061	.862
City	.164	.002	1.178	.144	.024	1.155
RR	-.088	.000	1.328	n/a	n/a	n/a
River	.302	.000	1.353	.284	.000	1.328
Constant	-123.097	.994	.000	-116.425	.994	.000

* Nagelkerke R square 0.552 ** Nagelkerke R square 0.784

Figure 26 displays the spatial lag model L4.1 site suitability map. As expected, higher probability areas (< 0.5) display ideal sites. However, the southwest corner of the map displays probability 0.5 - 0.74 which is rather low for areas with existing turbines. The residual diagnostic slightly underestimates the model (Figure 27).

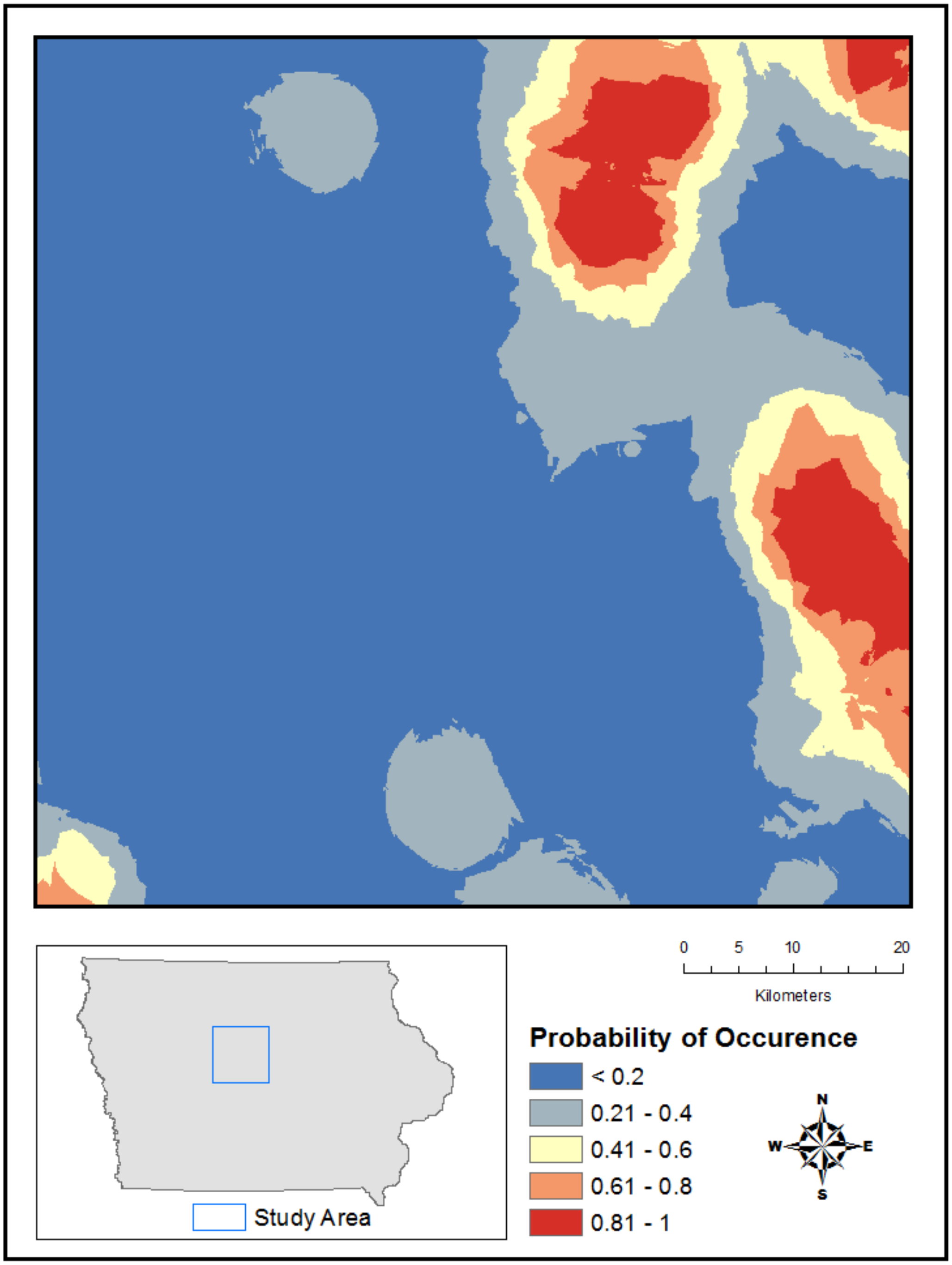

Figure 26: Spatial lag model L4.1 suitable site for turbine occurrence

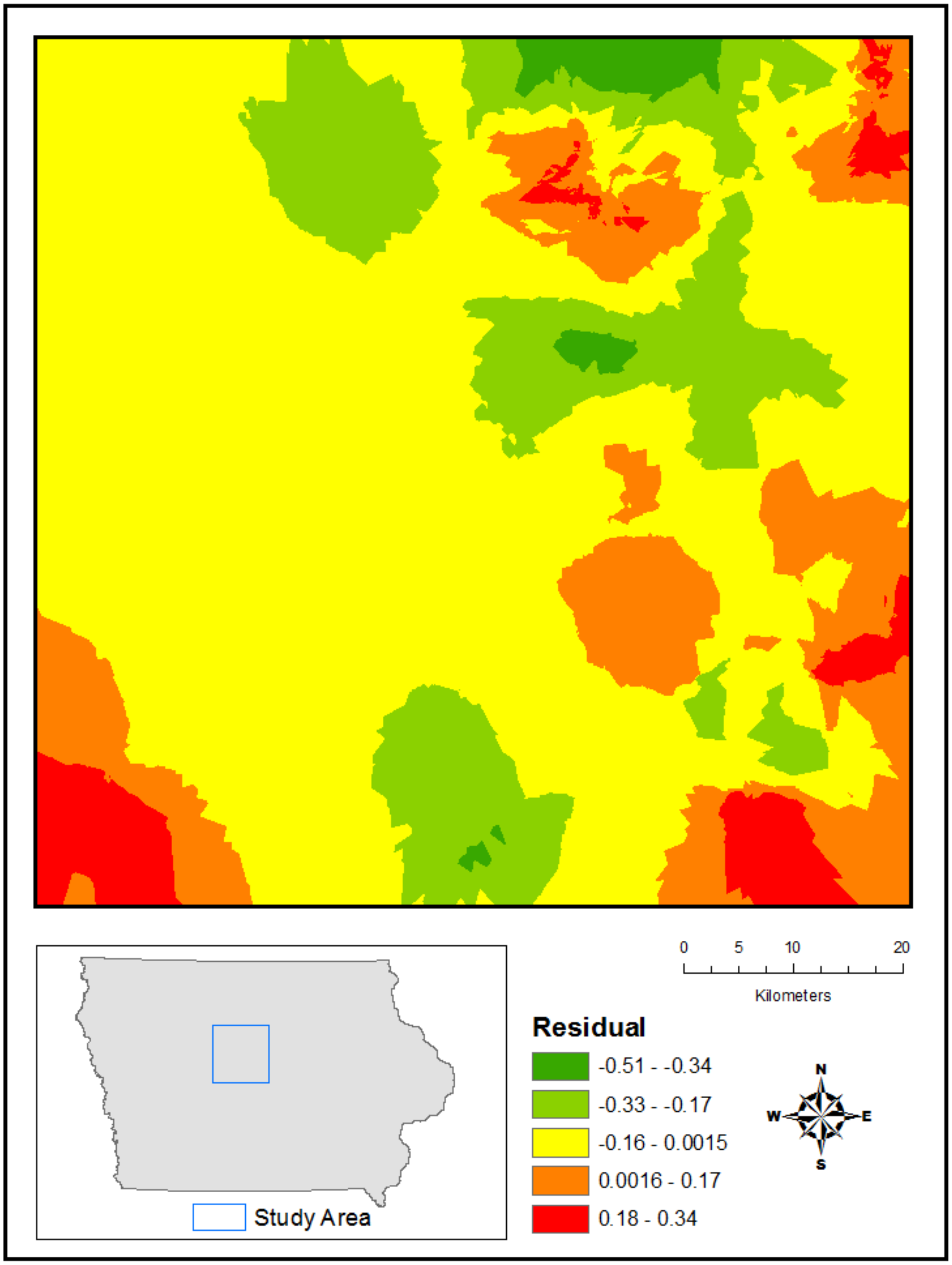

Figure 27: Spatial lag model L4.1 residual diagnostics

4.3 Summary

Spatially lagged models at all scales (Macroscale 1.1, Mesoscale 2.1, 3.1, and Microscale 4.1) performed significantly better than the standard binary logistic regression. Also, all the predictors displayed improved strength (positive or negative) in terms of likelihood of turbine occurrence when the proximity of neighboring turbine variables was included as predictors in the models (Table 10). Wind power class, elevation, distance to city and airport displayed significant contribution to the model L1.1, L2.1, L3.1, and model L4.1 had one predictor (elevation) in common with previous models. Distance to transmission line and highway displayed negative relationship (L1.1, L2.1, and l3.1) but not in model L4.1. The model L3.1 and 4.1 showed river as a positive predictors. Population density only showed significance in model L2.1. Microscale model L4.1 displayed only four predictors (elevation, distance to river, city, and proximity to neighboring turbine) that contributed to the model at this scale. Wind power class, even though the odds ratio was extremely influential, showed sig. > 0.995. At this scale, the distribution of wind power class does not display much variation (uniform wind power class at this scale).

Table 14 displays Nagelkerke R square for each model scale and compares models without spatial lag and with spatial lag factors. At all scales, models with the spatial lag component performed significantly better than standard models based on Nagelkerke R square comparison. Elevation, proximity to neighboring turbine and distance to city were statistically significant at scales. Wind power class and distance to transmission line were significance at the Macroscale, Mesoscale 1, and Mesoscale 2.

Distance to airport and highway appear to be scale dependent since they are significant at Macroscale and Mesoscale 1. Table 15 summarizes spatial lag logistic regression factors that were statistically significant at each scale. The positive or negative indicates the relation to the dependent variable (existence of turbine) based on slope- intercept.

Table 14: *Logistic Regression models comparison*

Scale	Without Spatial Lag	Spatial Lag
	Nagelkerke R Square	Nagelkerke R Square
Macroscale	.581	.861
Mesoscale 1	.308	.783
Mesoscale 2	.503	.794
Microscale	.552	.784

Table 15: *Spatial lag logistic regression factors contribution at different scale*

Scale	Significant Predictor
Macroscale	+WPC, +Elevation, +AP, +city, - TL, - Hwy
Mesoscale	+WPC, +Elevation, +AP, +City, -TL, -Hwy, -Pop Den
Mesoscale	+WPC, +Elevation, +City, +river, -TL
Microscale	+Elevation, +river, +city

4.4 Machine-Learning Algorithm Models

In this study Maxent provided a predictive (suitability) map and percent contribution table to measure the predictor variables influence and gain useful information for all scales (model M1, M2, M3, M4). The environmental variables (factors) used in each model are listed in Table 16, 17, 18, and 19. Also, Figure 28, 30, 31, and 32 illustrate the suitability map for each model.

<u>Maxent Model M1 (Macroscale)</u>

Macroscale model M1 used environmental variables slope, elevation, wind power class, distance to transmission line, highway, airport, river, city, railroad and land cover and population density. Table 16 displays the percent contribution of factors to the model. The most influential variables were elevation with 57.2% contribution and wind power class at a 19.9% contribution to the suitability distribution. On the other hand, the study by Petrov and Wessling (2014) showed wind power class as the most contributing with 51.2% and followed by elevation with 32.6% contribution. The remaining variables contribution had less than 6% each to the model. Furthermore, the jackknifing test showed elevation as the most useful independent environmental variable. Elevation had the highest gain when used in isolation therefore appears to have the most useful information by itself. Elevation decreased the gain the most when it is omitted; hence, it appears to have the most information that isn't present in other variables.

Table 16: *Maxent Macroscale variable percent contribution*

Variable	Percent Contribution
Elevation	57.2
Wind Power Class	19.9
Population Density	5.4
Airport	4.6
River	3.6
Transmission line	2.5
Slope	2.3
Railroad	1.9
City	1.3
Land cover	0.7
Highway	0.5

Figure 28 shows a Macroscale model M1 suitability distribution. Highly suitable areas have probability of 0.81 -1 in red while low suitability areas (blue) have less than 20% probability of turbine occurrence. Suitable conditions are found in the northwest, north central, and south central parts of the state which also align with existing turbines distribution in the state. The ROC curve (AUC) of 0.85 indicates a “good” fit to the model. The areas selected with more than 50% suitability general follow areas of high elevation and high wind power class. These areas are low in population density and very few incorporated cities which means rural areas. The combination of all these factors impacts suitability and existing turbines are in areas of high suitability due to limited constraints from the negative factors.

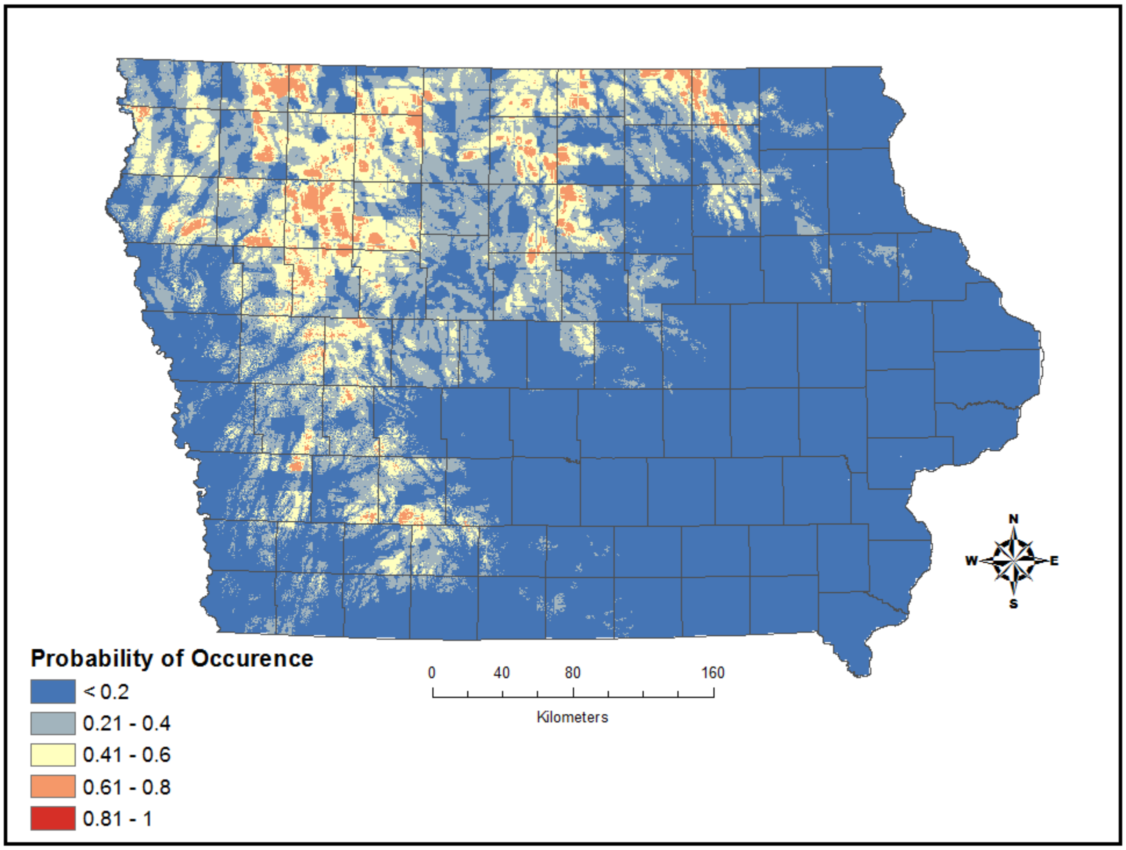

Figure 28: Maxent Macroscale modle M1 suitable site for turbine occurence

Maxent also produced factors response curves which allows us to identify optimal range (Figure 29). Elevation optimal range that is suitable for wind turbine occurrence is between 1050 – 1300 ft (320.04 – 396.24), wind power class (WPC) prefred range is 4 – 6 while slope is optiaml at 3.5% or lower . Cropland and grassland is the prefered land cover type. Similar results were observed for Model M2, M3, and M4 as well.

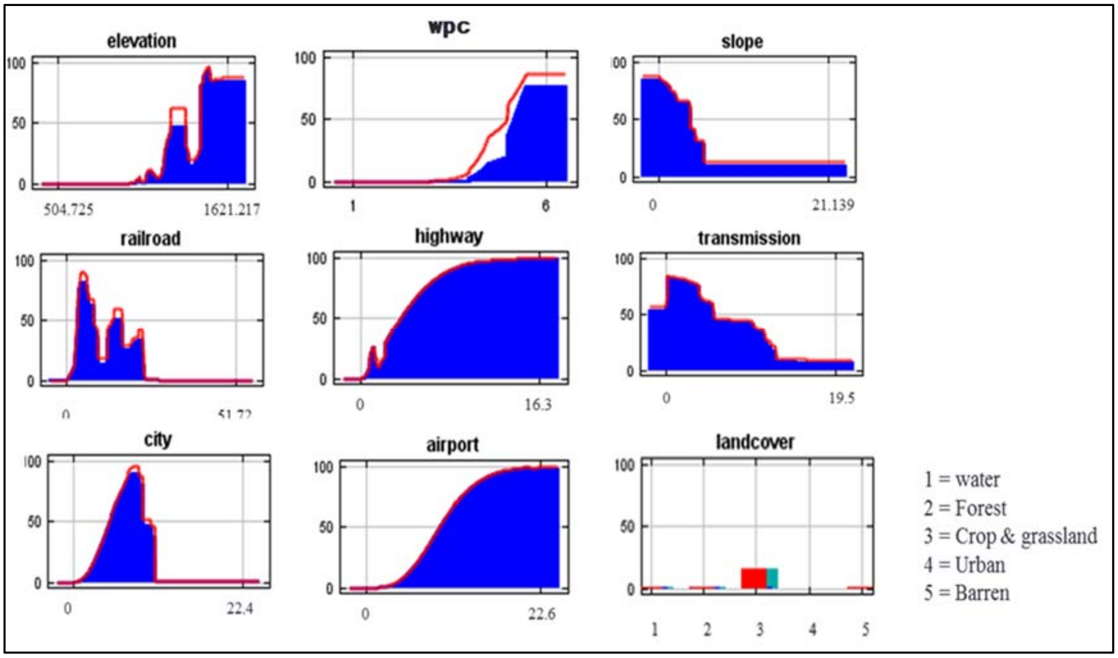

Figure 29: Maxent Macroscale factors response curves

Maxent Model M2 (Mesoscale 1)

Elevation had 50.7% contribution followed by population density at 20.5% contribution indicating a high level of importance for these two factors to the suitability distribution (Table 17). Jackknifing identified elevation as the single most useful environmental variable which provided the most gain when used in isolation indicating to provide the most effective information to predicting the distribution of the turbines occurrence. Elevation also decreases the gain of the model the most when it is omitted therefore appearing to have the most information that isn't present in other variables in the model.

Table 17: *Maxent Mesoscale 1 variable percent contribution*

Variable	Percent Contribution
Elevation	50.7
Population Density	20.5
River	8
Wind Power Class	5.7
City	4.4
Transmission line	3.4
Airport	2.8
Railroad	2.3
Highway	1.6
Land cover	0.4
Slope	0.1

The predictive map with areas of high probability of suitable conditions are in red with blue designating areas having a very low probability (Figure 30). A continuous distribution is shown and the range is 0 to 1. High probability of suitability is found in western and north central regions of the study area. Suitable conditions tend to be closer to existing turbines which agrees with spatial lag regression. The model's goodness-of-test with the ROC curve (AUC) of 0.86 indicates a good fit and a slight improvement from model M1 (0.85).

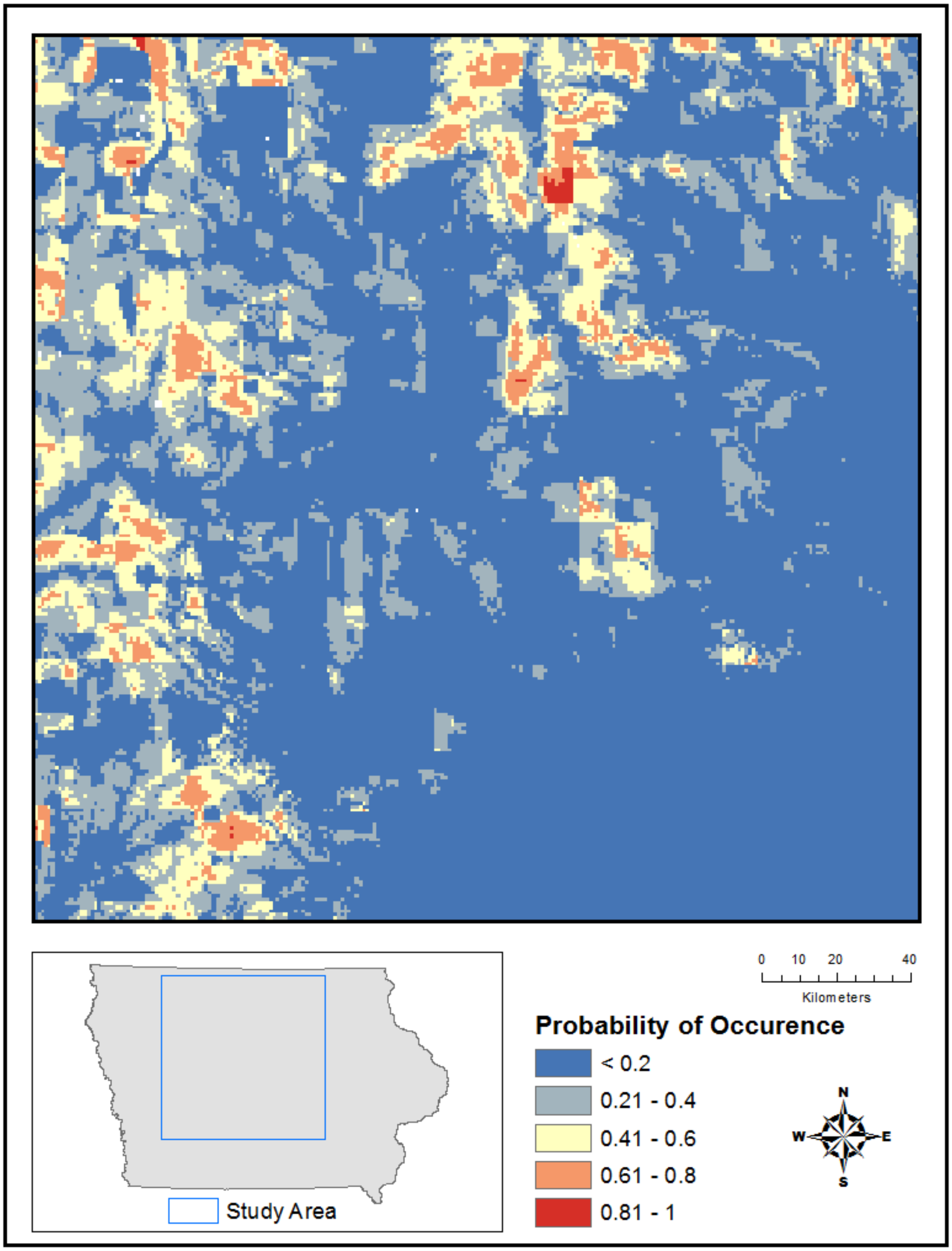

Figure 30: Maxent Mesoscale 1 model M2 suitable site for turbine occurrence

Maxent Model M3 (Mesoscale 2)

In model M3, the elevation had 34.7% contribution followed by population density at 20.2% contribution indicating a high level of importance for these two factors to the suitability distribution (Table 18). In contrast, Petrov and Wessling (2014) Mesoscale (6 - 8 counties region) which is comparable to model M3, displayed elevation with 79.1% contribution and population density of 0.3% contribution. Jackknifing identified population density as the single most useful environmental variable which provided the most gain when used in isolation. Also, population density decreases the gain of the model the most when it is omitted therefore appearing to have the most information that isn't present in other variables in the model.

Table 18: *Maxent Mesoscale 2 variable percent contribution*

Variable	Percent Contribution
Elevation	34.7
Population Density	20.2
Airport	10.4
Wind power class	9.4
City	8.8
Transmission line	7.1
River	5.9
Railroad	1.9
Highway	0.9
Slope	0.7
Land cover	0.2

The suitability map identified areas to the right and center of the image which are within proximity to occurrence data as relatively suitable (Figure 31). ROC curve (AUC) of 0.925 is an excellent model fit and substantial improvement from model M1 and M2.

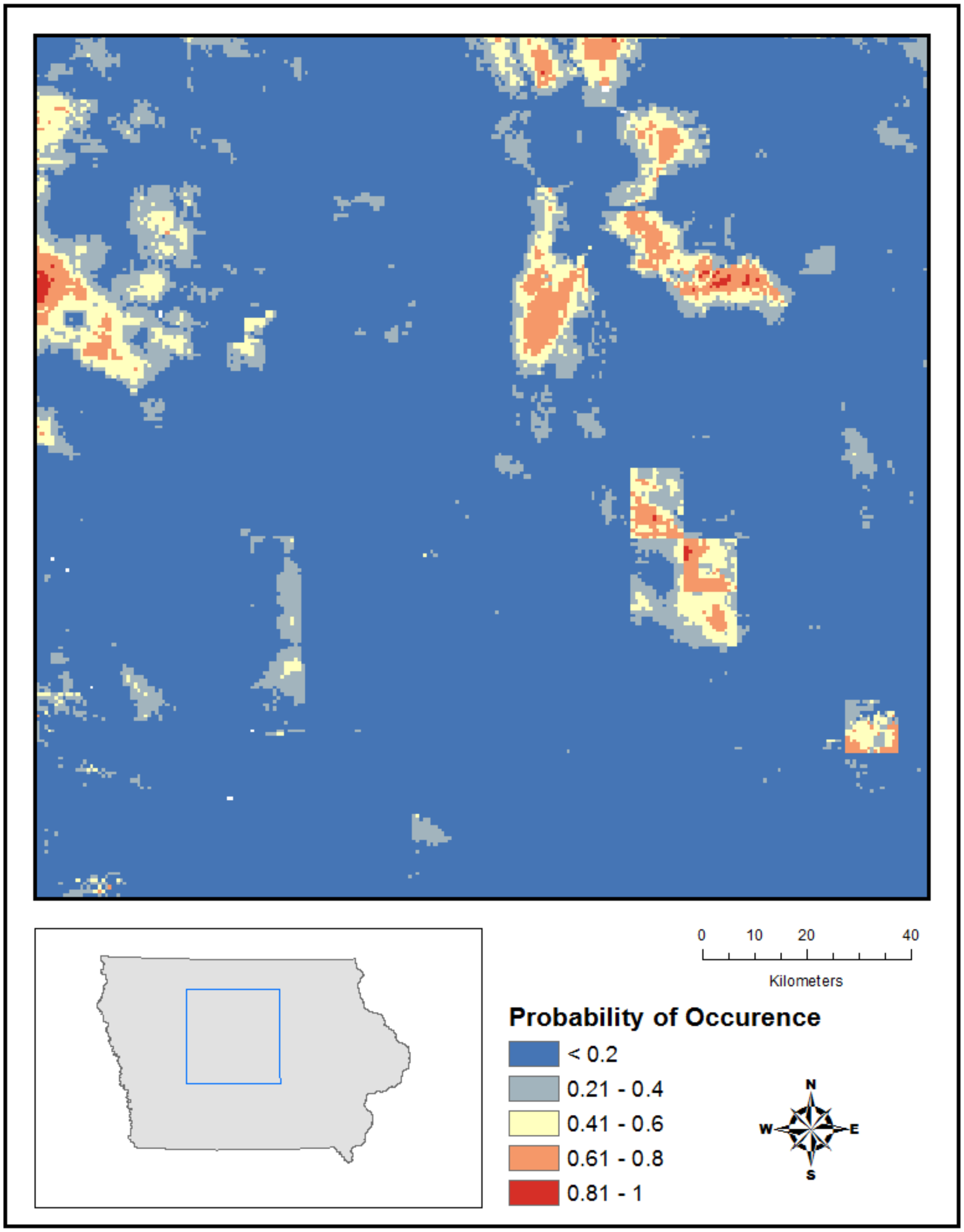

Figure 31: Maxent Mesoscale 2 model M3 suitable site for turbine occurrence

Maxent Model M4 (Microscale)

The Microscale model M4 elevation had 49.2 % contribution followed by a population density at 19.3% contribution (Table 19). Wind power class contributed 15.7% to the model which is the first time any model showed a third contributing environmental variable. Jackknifing identified elevation as the single most useful environmental variable which provided the most gain when used in isolation. But population density environmental variable decreased the gain of the model the most when it was omitted therefore appearing to have the most information that wasn't present in other variables in the model. Figure 32 displays high occurrence within proximity to existing areas. At this scale, suitability estimates are limited to areas near where existing turbines are located.

Table 19: *Maxent Microscale variable percent contribution*

Variable	Percent Contribution
Elevation	49.2
Population Density	19.3
Wind Power Class	15.7
Airport	6
River	3.8
City	1.6
Railroad	1.6
Highway	1.2
Transmission line	0.8
Land cover	0.5
Slope	0.3

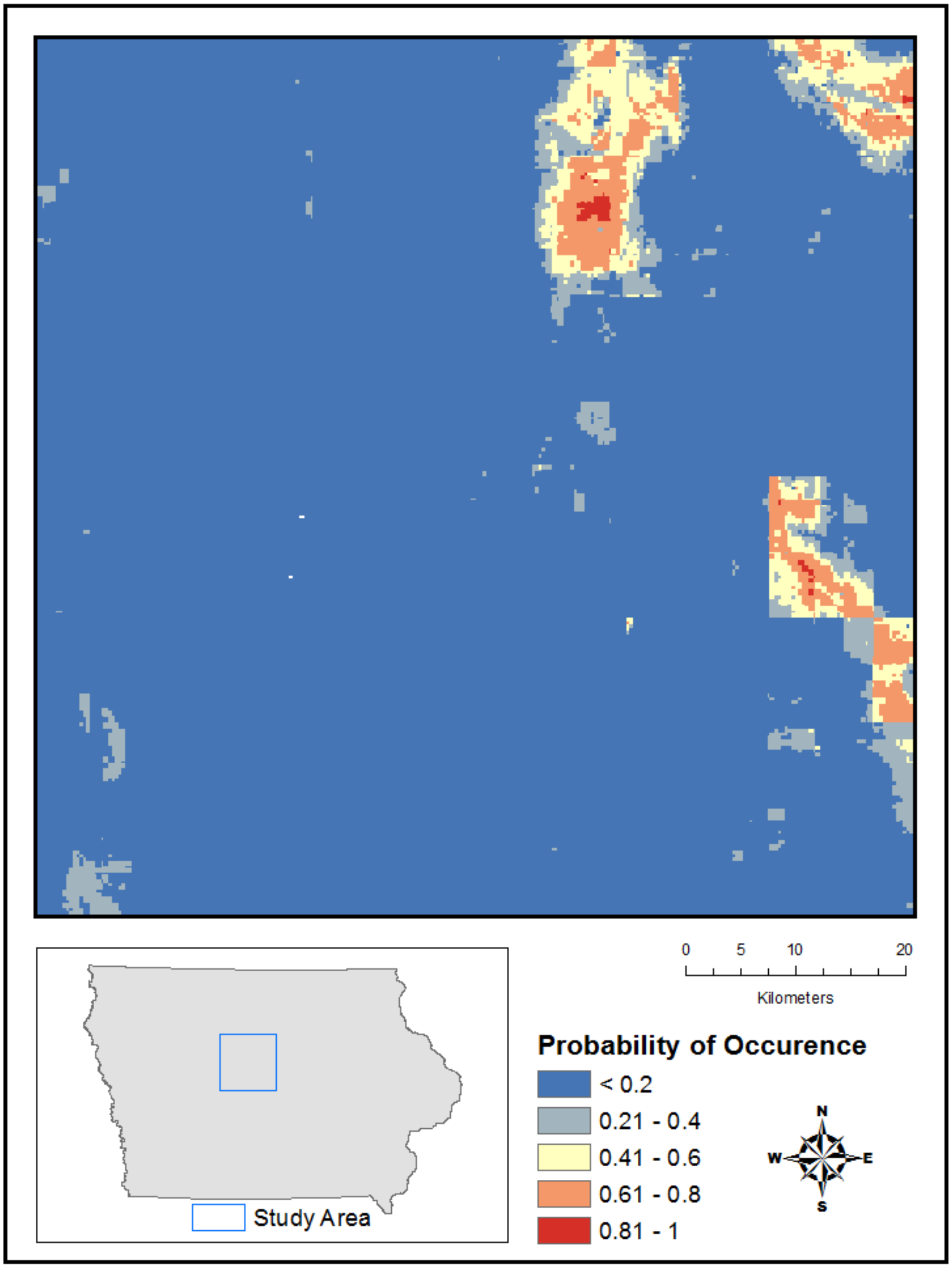

Figure 32: Maxent Microscale model M4 suitable site for turbine occurrence

4.5 Summary

Table 20 displays the machine-learning (Maxent) factors and percent contribution at the different scales. Top three influential predictors are listed since more than two-thirds percent contribution to the fitted models contain the factors listed. In this case, the Maxent component identifies elevation as the most important and influential factors at all scales which collaborates the spatial lag component. Therefore, this indicates that the elevation factor should be included at all scales in any wind energy development. Wind power class on the other hand is the only significant factor at the Macroscale and Microscale (6 county level) which did not correspond to the spatial lag component. The difference might be the result of the machine-learning component only using the locations of existing turbines which have wind power class >= 3 thus causing the model not to identify WPC due to lack of differences in WPC. In the context of Iowa, the majority of the state has WPC >= 3 which is the bottom range WPC ideal for utility scale development. The contribution of WPC is highly dependent scale and therefore should be considered at all scales. Maxent identified population density as a significant contributor at all scales while the spatial lag models significance of population density was minimal.

Examining factors response curve, optimal range for each factor are identified (Table 21). Suitable site generally will have less than 3.5% slope grade with elevation ranging from 320 – 396 m. High elevation is due to exiting turbines being located in the western part of the state where it’s relatively high. In this case, high elevation means high wind power class as well. Optimal range identified by the Maxent model for wind power class ranges from 4 – 6 (7 - 8 m/s wind speed). Now that environment optimal range are identified, distance to infrastructure are also highlighted. In this case, distance from wind farm to transmission line is optimal 2.5 – 5 km range. While cropland, grassland, and barren land are preferred land cover types for placing turbines. However, when erecting turbines, it needs to be cited at least 3 km away from the nearest city and 5 km from the nearest airport. Wind farms should be less than 1.5 km from highway but greater than 2.5 km from railroad tracks.

Table 20: *Maxent models significant predictors and percent contribution*

Scale	Significant Predictor	% Contribution
Macroscale	Elevation, WPC, Pop Density	82.5
Mesoscale	Elevation, Pop Density, and Dist. to river	79.2
Mesoscale	Elevation, Pop Density, and Dist. to airport	65.3
Microscale	Elevation, Pop Density, and WPC	84.2

Table 21: *Maxent optimal site suitability range for turbine placement*

Factor	Optimal Factor Range
Slope	< 3.5%
WPC	4 – 6 (7-8 m/s)
Elevation	320.04 – 396.24 m (1050 – 1300 ft)
Transmission line	< 2.5 – 5 km
Land cover	Cropland, grassland, and barren land
City	>3 km
Highway	< 1.5 km
Airport	>5 km
Railroad	>2.5 km

CHAPTER 5

DISCUSSION

5.1 Methodological Improvements

This research highlights four methodological improvements: (1) Further development of empirical modeling, which is a relatively new approach in site suitability assessment for wind energy development (2) Implementation and testing of spatial lag regression, which accounts for the spatial autocorrelation due to spatial clustering of turbines (3) Explanatory of critical factors using multiple methods and their scale manifestation in the context of Iowa are examined (4) Incorporation of scale to test the impact on factors and site suitability.

First, empirical models based on existing turbines are derived from the spatial lag logistic regression and machine-learning algorithm (Maxent) components. Empirical models showed high accuracy, differentiated factors importance and gained better understanding of the complex factors that generally exist in site suitability assessment. Previous studies employed normative 'expert-based' approaches by combining overlay, buffering, and weighting of criteria to determine suitable site (Grady et al., 2005; Mosetti et al., 1994). Empirical models used in this study identify evidence driven relationship between factors which is a shift from the traditional expert based suitability assessment. Empirical approach is a relatively new in the context of wind energy development and site suitability modeling (Mann et al., 2012; Petrov & Wessling, 2014). As such, spatial lag and machine-learning algorithm methods are high in agreement on site suitability factors, spatial manifestation, and identifying suitable site based on empirically driven approach.

Secondly, spatial lag regression based on existing turbines (3,177 turbine points) were used. This approach accounted for environmental, technical, and social constraints inherently in siting locations. Turbines spatial distribution are highly clustered and spatial autocorrelation must be addressed by introducing the proximity to neighboring turbine (NT) variable in all models. Controlling for spatial autocorrelation, models improved substantially at all scales (Table 14). The main gain is that non-spatial factors were not stretched to account for variance in the models that otherwise could not have been explained.

Thirdly, the lack of understanding of relative importance and scale manifestation of factors are a shortcoming in literature explicitly addressed this study (Table 15). Multiscale empirical approach enabled us to assess scale's impact on factors. Scales' impact on factors is directly from the empirically driven approach. Previous studies examined the importance of factors based on expert assessment methodology (Aydin et al., 2010; Baban & Parry, 2001; Mann et al., 2012). Therefore, empirical approach in this study explicitly addresses this shortcoming by identifying important factors and scale dependency. Elevation, wind power class, proximity to neighboring turbine, and distance to city appear to be the most important factors. However, other factors impact is dependent on scale. At Microscale (six county level), wind power class is statistically insignificant. It might be due the study area containing a single wind power class (no variation in wind speed).

Finally, multiscale analysis highlighted the importance of scale and the impacts it has on factors and site suitability. It also demonstrated the critical need to clearly define scale in order to identify suitable site (Petrov & Wessling, 2014). Furthermore, multiscale analyses fills the knowledge gap on the scale and addresses the MAUP's impact on factors and site suitability. Scale and its impact were illustrated in this study be examining scale at the Macroscale (statewide), Mesoscale 1(regional), Mesoscale 2(sub-regional) and Microscale (6 county level; Table 14).

5.2 Understanding Contributing and Influential Factors

This study shows that important and influential factors are elevation, wind power class, proximity to neighboring turbine and distance to city while the remaining factors were scale dependent. The impact of elevation and wind power class go hand in hand. Majority of existing turbines are located in areas with high elevations located in the Northwest, North central, and south central parts of the state. These areas also contain the highest WPC in the state confirmed by the correlation analysis, higher elevation generally leads to high wind power class. However, this does not mean an increase in elevation always produces an increase in wind power class. The combination of other environmental and climatic factors along with local geographic characteristics are likely to influence wind power (Van Hoesen & Letendre, 2010). For instance, a study by Rodman and Meentemyer (2006) demonstrated how low valleys can serve as a channel for increasing wind power class.

Wind power class is an essential factor in scales except Microscale (6 county level). While Toke et al. (2008) acknowledge that wind resources in a region are not always the driving factor but the consideration of other factors in addition to wind potential is equally important. They noted that one of the most important factors is the level of investment and siting decision are made and who is involved in those decisions. In this study, existing turbines are located in areas with wind power class 3 or greater which means they are in a suitable area just on the basis of wind power class. Petrov and Wessling (2014) noted how wind speed (wind power class) improves site suitability for utility scale wind energy development when wind power class is 3 or greater which also aligns with NREL recommendation.

Spatial lag factor (proximity to neighboring turbine) improved the models performance at all scales and other co-factors coefficient improved as well. Existing turbines are highly clustered which explains why proximity to neighboring turbine is such a strong predicting factor. Spatial autocorrelation of the turbines confirmed proximity to neighboring turbines increases the likelihood of turbine occurrence as well. On the other hand, it's logical to expect if turbines exist (wind farm), it areas near or within proximity to existing turbines, they might have the environmental, technical, social characteristics conducive for development.

Furthermore, factors such as land cover, slope, and distance to highways, and rivers are scale dependent. Land cover as a categorical variable, it displayed significance and importance as well. The empirical approach segmented the classification to determine ideal land cover type for wind energy development. As a result, row crop and grassland displayed the most positive significance while water and urban were not significant. Iowa is mostly agricultural state; as such, it was not surprising since row crop and grassland are the dominant land cover types which are ideal for turbines to be placed. In fact, over 75% of existing turbines occupy row crop and grassland while the remaining 25% located in barren and forest areas. Turbines are placed in the rural areas and most of the rural areas are privately owned farms. This also reveals the willingness of farmers or land owners in Iowa to lend their land for wind energy development (Slattery et al., 2012). Also, Sowers stated (2006) that most Iowan's see turbines as an economic benefits.

Distance to airport, city, and river appear to have a positive relationship. The greater the distance away from these features the more likely the site might be suitable for placing turbine. There are 313 airports of all sizes, and FAA regulations require builders/developers or any persons who want to construct object over 60.96 meters get approval. This regulation is intended to minimize interference with aircrafts and radars. Even if all other criterions are met, site is unsuitable if distance to airport criteria is not met. Further away from city which probably has low population density diminishes the likelihood of turbine occurrence because it won't be suitable. Distance to rivers also appears to be influential but highly depends on the scale. In contrast, further away from transmission lines increases the cost and limits accessibility to the grid to transport wind energy to the market place. The greater the distance from highway makes it difficult to access turbines. Heavy machineries are needed to install and maintain wind turbines so gravel roads are connected to surface roads near the site. Turbines require substantial maintenance due to the wear and tear thus it's essential to have easy access.

5.3 Empirical Modeling and Reduced Regression Comparison

The relative influence and significance of factors are varied between models at the different scales. As illustrated in the results section, the importance of scale in identifying suitable site is critical. Factors importance is scale dependent thus its critical scale should be incorporated in site suitability assessment phase. Macroscale model accuracy displayed a very good fit with Nagelkerke R square 0.861. Maxent model M1 (Macroscale) ROC Curve of 0.85 also indicated a very good fit to the model. Elevation and wind power class are dominant factors in determining suitability from the empirical components. In the Maxent Macroscale model, elevation and wind power class provide over 70% to the model contribution. In addition, slope, distance to airport, city, transmission line, and highway are also significant predictors at the Macroscale level in the spatial lag regression analysis. The difference from the two approaches might be due to Maxent only using 'occurrence' data which tends to over fit. Therefore, factors like population density and land cover impacts is minimal which indicates how sparsely populated the area where turbines are located and the dominance of row crop and grassland land cover in Iowa the statewide scale. This is useful for resource characterization and to paint the larger picture to utilize with maximization. In doing so, wind energy developers have a general understanding of the resources to build the basis for detailed and site specific wind resource study.

Mesoscale 1 (regional level) results showed to the fact that similar factors were significant as in Macroscale. Spatial lag regression model accuracy with Nagelkerke R square of 0.80 is a good fit while Maxent ROC Curve of 0.86 indicating a very good fit and an improvement from Maxent Macroscale. Second, Maxent's use of 'occurrence' only data tends to over fit the model. Elevation and population density provide 70% contribution to Mesoscale 1. Iowa's population density is among the lowest (34^{th}) in the nation and existing turbines located in rural areas indicating suitable sites are in low population areas.

Additional analysis was performed for Macroscale and Mesoscale 1 spatial lag regression models by selecting common factors from both scales to re-run logistic regression. Factors that were statistically significant were selected from Macroscale and Mesoscale 1; the default enter method was applied in SPSS 15.0. This was done to determine whether models with fewer factors performed better based on Nagelkerke R Square compared to the result presented in the section 4.2. Reduced regression analysis for Macroscale and Mesoscale are in Table 22. Macroscale Nagelkerke R square of 0.859 (original Nagelkerke R square 0.861) demonstrates a very good fit and this model with reduced factors performs on par with the model that contains all factors. Mesoscale 1 displayed similar results (original regression) in terms of factors significance and Nagelkerke R square of 0.799 (original Nagelkerke R square 0.8) indicates a good fit. Both scales appear to highlight similar factors being strong predictors with the exception of Mesoscale 1, highway is statistically insignificant.

Both scales (Macroscale and Mesoscale 1) R square indicate little difference between the full and reduced factors regression. However, it's important to use the full regression results for several reasons. First, I would argue using all factors enables us to understand how the factors affect suitability and be able to quantify each factors contribution even if it's minimal. Second, reducing factors for the purpose of improving R square might be irrelevant to developers who are seeking all information to maximize production and reducing cost.

Table 22: *Macroscale and Mesoscale 1 Common Factors Logistic Regression*

	Macroscale*			Mesoscale 1**		
	B	**Sig.**	**Exp(B)**	**B**	**Sig.**	**Exp(B)**
WPC	.765	.000	2.149	1.189	.000	2.111
Elevation	.014	.000	1.014	.020	.000	1.019
LC	n/a	.001	n/a	n/a	.033	n/a
LC(1)	-19.920	.999	.000	-19.754	.999	.000
LC(2)	-1.193	.293	.303	-2.210	.003	.110
LC(3)	-.833	.066	.435	-2.078	.004	.125
LC(4)	-.502	.253	.605	-4.643	.021	.010
LC(5)	-2.434	.000	.088	-4.693	.000	.009
NT	-5.560	.000	.004	-.322	.000	.725
TL	-.202	.000	.817	.317	.000	1.373
City	.154	.000	1.166	.114	.000	1.121
Hwy	-.066	.005	.936	n/a	n/a	n/a
Airport	.107	.000	1.113	.215	.000	1.240
Constant	-5.803	.000	.003	-10.219	.000	.000

* Nagelkerke R square 0.859 ** Nagelkerke R square 0.799

Mesoscale 2 Nagelkerke R square 0.749 indicating good model fit while Maxent ROC curve of 0.925 is excellent model fit. Maxent model M3 displayed significant improvement from model M1 (Macroscale) and M2 (Mesoscale 1) when comparing the ROC curve. Elevation and population density factors provide 50% contribution to Mesoscale 2. Large scale depicts greater detail and the reduction of factors while small scale tends to give broader overview but the core factors don't appear to be significantly impacted.

On the other hand, Microscale (six-county level) Nagelkerke R square 0.784 while Maxent had ROC curve of 0.945 is a good fit. Elevation is the dominate indicator of suitability while wind power class is insignificant due to statewide resolution. At large scale, local terrain characteristics can have a significant effects on the local wind speed variability (Petrov & Wessling, 2014). This indicates small changes in topography are a better predictors of suitability than wind power class. This also indicates the need for localized assessment because localized factors impact suitability.

Common factors that were statistically significant were selected at the Mesoscale 2 and Microscale for additional analysis in all models. Mesoscale 2 Nagelkerke R square 0.749 and Microscale Nagelkerke R square 0.757 are slightly lower than the spatial lag regression from the full model (Table 23). It appears that the regression with common factor performs reasonably well. However, as stated in previous explanation regarding the use of reduced number of factors at different scales, more factors are better equipped to define the relationship between scale and factor. Obtaining the highest R square with fewer variables might not provide all the necessary information. In order to understand scale's impact on factors and how suitability is affected; all factors should be included and base it off the empirical model to identify important factors. Site suitability models that contain all factors are more beneficial to developers, decision makers and stakeholders of wind energy because it provides the most information.

Table 23: *Mesoscale 2 and Microscale Common Factors Logistic Regression*

	Mesoscale 2*			Microscale**		
	B	**Sig.**	**Exp(B)**	**B**	**Sig.**	**Exp(B)**
Elevation	.038	.000	1.039	.069	.000	1.071
NT	-4.488	.000	.011	-4.384	.000	.012
City	.195	.000	1.216	.146	.013	1.158
River	.196	.000	1.216	.279	.000	1.322
Constant	-11.778	.000	.000	-22.963	.000	.000

* Nagelkerke R square 0.749 ** Nagelkerke R square 0.757

5.4 Modifiable Areal Unit Problem in Spatial Data Analysis

To date, multiscale analysis has not been examined in wind energy site suitability assessment. The complexity of natural systems require a multiscale approach to understand the impact of scale on predicting factors examined in this study. This research uses overlapping multiscale analysis; therefore, tackling the MAUP problem when conducting scale dependent spatially explicit data analysis (Openshaw, 1983; 1984). Several studies have confirmed that statistical results vary based on scale which is a cause for concern in geographic research (Dark & Bram, 2007; Flowerdew, 2011). Models are based on spatial datasets that are valid for Macroscale analysis and the use of the same dataset to infer higher-resolution or lower-resolution (WPC data at Macroscale used to assess regional suitability) may produce inaccurate results.

The statistics and model parameters differ between scales. This study has focused on the issue of scale, impact on site suitability and spatial distribution. MAUP impact occurs for both the spatial and temporal scale. Dark and Bram (2007) acknowledge the existence of natural scales at which physical characteristics occur within the landscape. One solution is to use dataset appropriate for each scale because research suggested that scale determines the range of patterns and process that can be detected thus requiring an appropriate level of resolution for the study area (Flowerdew, 2011). Appropriate spatial scale for spatial analysis is an ongoing and unresolved issue within many disciplines of geography. Openshaw (1984) suggest to focus on identifying the appropriate scale and dataset to minimize MAUP impact. This study address this by using multiple scales and comparing results to better understand the MAUP impact on wind farm modeling.

5.5 Developing Spatially Explicit Scale Dependent Modeling Framework

Rapid wind energy development demands a framework that will address the complex, technical, environmental, and social constraints of site suitability assessment for wind energy development. Thus, the methodological approach used in this study incorporates an integrated module that combines the empirical and normative components to create the spatially explicit scale dependent framework (Figure 34).

The empirical module is from the spatial lag logistic regression and machine-learning algorithm (Maxent), while normative module accounts for factors not considered in the empirical module such as regulations and policies. In this particular empirical component uses coefficients derived from the spatial lag where factors were statistically significant. The logistic regression equation predicts the probability of Y taking a specific value. In the application case study presented here, the empirical module is based on computation of wind turbine probabilities and according to the following formula derived from logistic regression. The logistic regression formula is as follows:

$$P(Y) = \frac{e^{b_0+b_1x_1+\cdots+b_nx_n}}{1+e^{b_0+b_1x_1+\cdots+b_nx_n}}$$

P: probability of Y occuring

e: natural logarithm base

b_0: interception at y-axis

b_1: line gradient

b_n: regression coefficent of X_n

X_1: preictor variable

The normative module contained characterization based on known regulations (1), suitable or (0) not suitable. The ArcGIS Map Algebra function was used to combine the empirical and normative modules in ArcGIS ModelBuilder (Figure 33). Application case study at the Mesoscale was conducted. Final suitability map was produced and validation was conducted using the pixel value (probability of occurrence) for each turbine and using a cut-off value of 0.5. The cutoff value was determined based on the probability of turbine occurrence or non-occurrence at 50%. This value also aligns with existing turbines existing turbines occupying areas with probability 0.5 because conditions in these areas meet the minimum requirements of wind power class greater than 3 and other environmental factors also appear to be significant. Furthermore, total number of turbines with probability greater than 0.5 were divided by the total number of turbines to get the accuracy percentage.

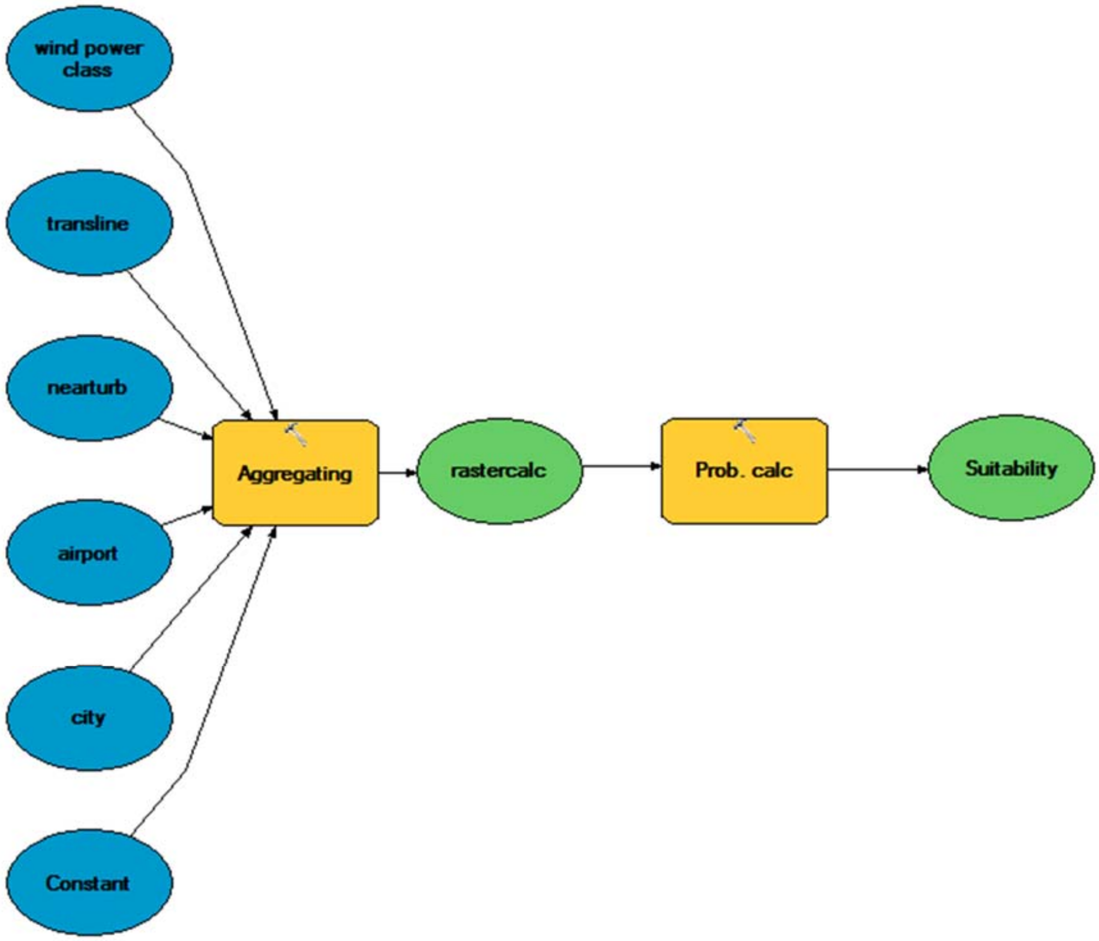

Figure 33: ModelBuilder site suitability using Map Algebra function

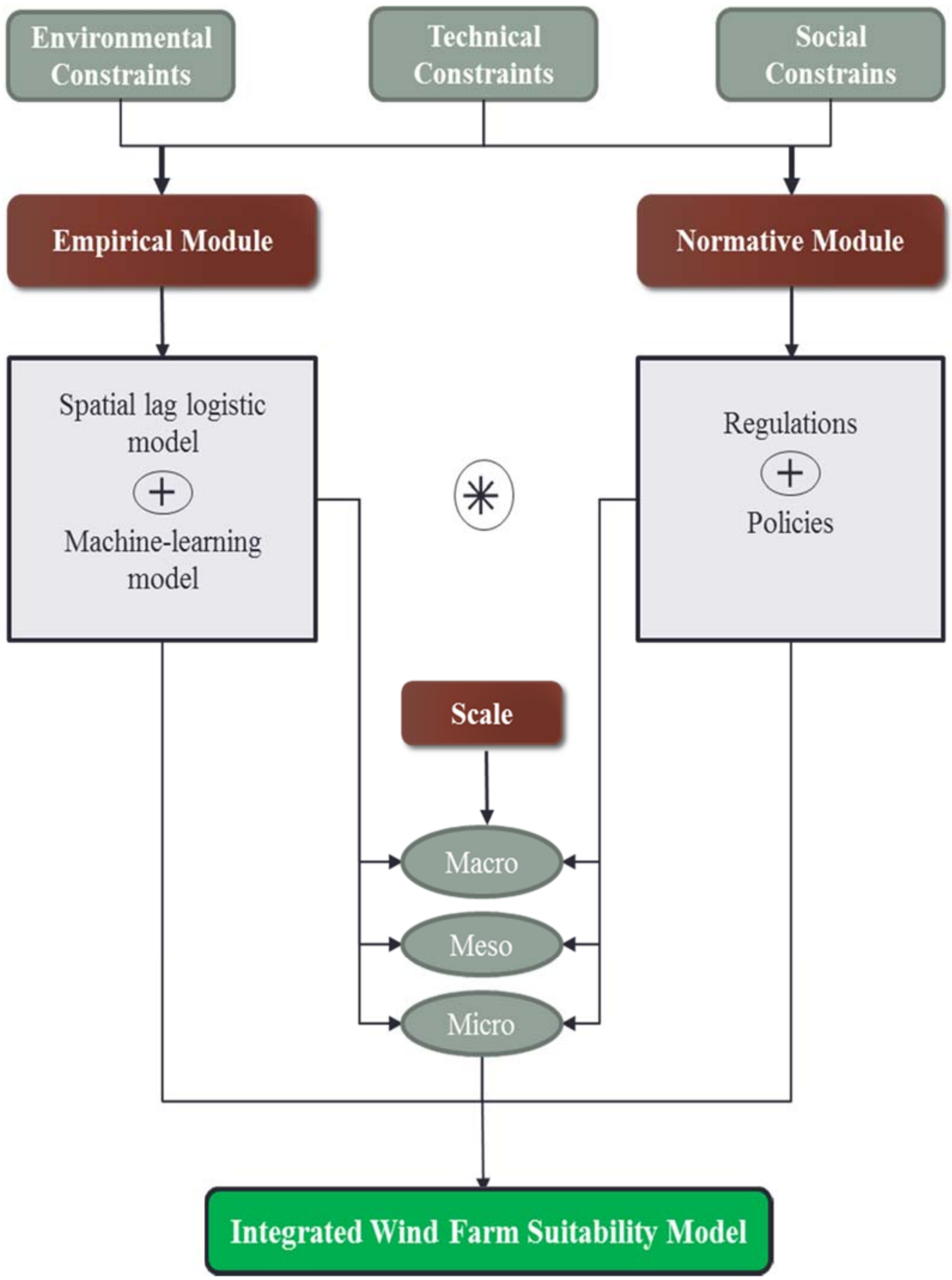

Figure 34: Spatially Explicit Scale Dependent Suitability Assessment Framework

Mesoscale Framework Application

Mesoscale suitability map without spatial lag component has 85% model accuracy (Figure 35). While Mesoscale suitability map with spatial lag component derived from the framework showed 90.5% mode accuracy (Figure 36). Both maps display the same core area with high suitability (red), limited medium suitability (burnt orange), and even less area with low suitability (yellow) are highlighted.

High suitable areas (red) have 0.81-1 probability of turbine occurrence. While areas 0.61-0.80 probability of occurrence are suitable even though very limited in terms of available land area especially in the suitability map with spatial lag component. In contrast, areas with probability 0.21 – 0.4 range might be suitable; however, more disadvantages, i.e. higher costs and technical difficulties might exist.

Comparing the two maps illustrates Mesoscale without spatial lag appears to be to highlight even areas that don't meet the suitability criteria. On the other hand, the map with spatial lag component appear to be inclusive and selecting substantial areas of suitability. Spatial lag model contains red dots which represent existing turbines. It appears there are existing turbines that are not in suitable areas; therefore, they might not be as productive. At this scale, suitability maps illustrates the variation in spatial variation over short distance. Site suitability are impacted by local factors thus local site assessment and factors should be thoroughly investigate. Local environmental and physical characteristics appear to impact wind power class especially at the Microscale (Petrov & Wessling, 2014; Toke et al., 2008).

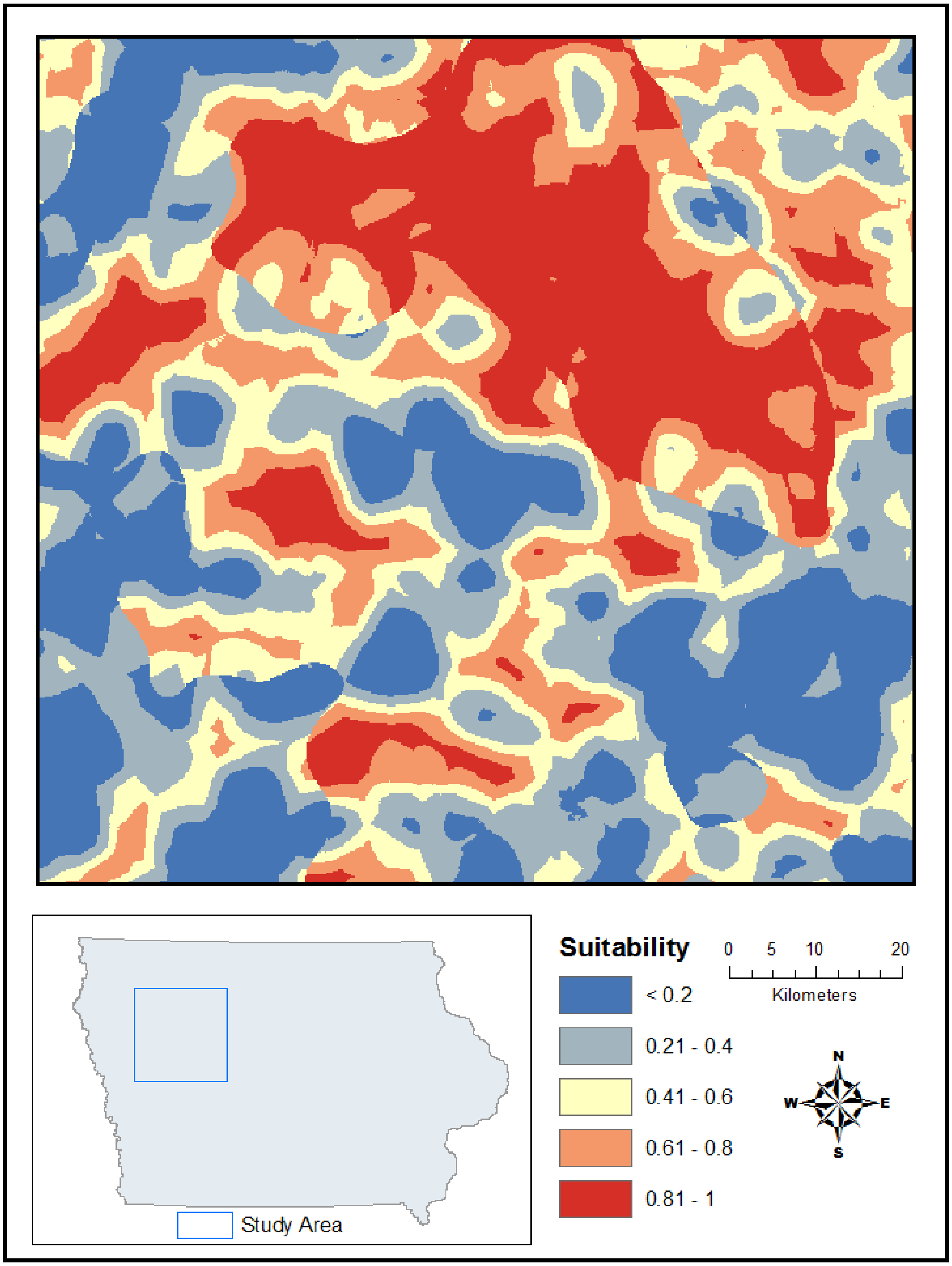

Figure 35: Case study site suitability without spatial lag component

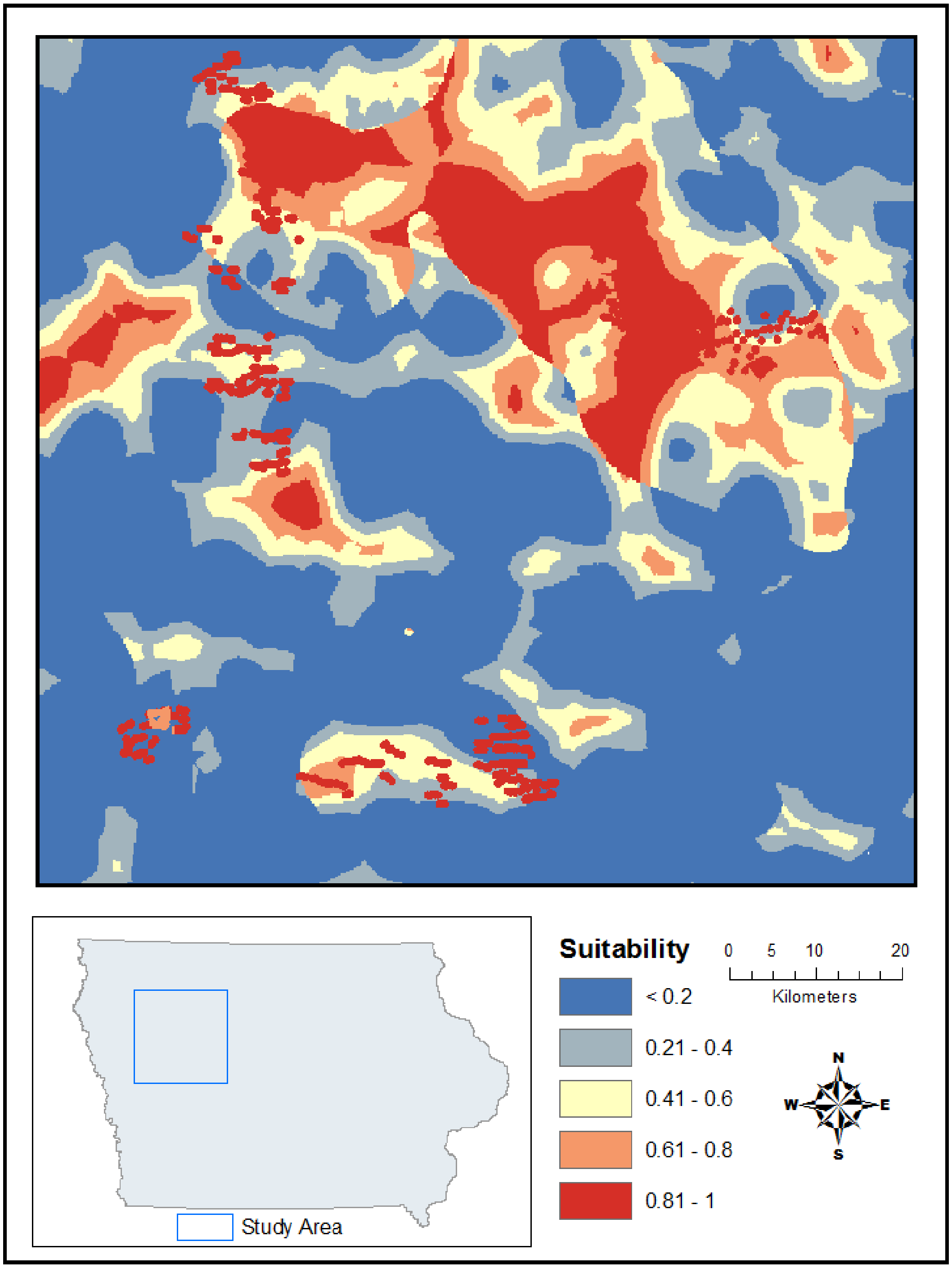

Figure 36: Case study site suitability with spatial lag component

CHAPTER 6

CONCLUSIONS

6.1 Conclusions

The limitations of traditional studies on site suitability has prompted the search for advanced empirically driven modeling framework for wind farm site suitability assessment. This research examined the effects of scale on suitability factors, and develops a spatially explicit scale dependent modeling framework for placing wind farms in Iowa. This study developed and examined an empirical driven approach and coupled it with normative components to assess wind farm suitability. Existing models are insufficient as they are based on incomplete data and a limited number of considered spatial factors and turbines are spatially clustered which requires autoregressive control. This study addressed these shortcomings by implementing spatially explicit scale dependent modeling framework derived from spatial lag logistic regression and machine-learning algorithms coupled with normative factors to account for the technical, environmental, and social factors for site suitability assessment in Iowa.

The three methodologies examined in this research prove components to modeling wind farm site suitability. (1) The empirical model derived from spatial lag logistic regression and machine-learning algorithm (Maxent) were used to examine suitability distribution for wind energy development in Iowa by analyzing the locations of existing turbines and identifying factors importance. Factors identified from extensive literature review and applicable in the context of Iowa were used. As a result, empirically driven models were developed which identified statistically significant factors and incorporated multiple scales.

(2) Spatial lag logistic regression significantly improved modeling accounting for spatial autocorrelation. It also allowed to analyze empirically driven measurements of factors impact at different scales. This approach was appropriate since turbines are spatially clustered. Accounting for spatial autocorrelation, variables are forced by the model to account for variances in the models otherwise would not be explainable.

(3) The impact of scale on factors importance on the models are illustrated in this research. While previous studies have focused on a single scale (i.e. wind farm, regional or state) this research fills the knowledge gap by implementing multiscale analyses. As a result, our evidence suggests that scale does affect factors contribution to models and site suitability. The Macroscale model with Nagelkerke R square of 0.861 identified proximity to neighboring turbine, wind power class, elevation, slope and distance to city, airport, transmission line, and highway as significant factors that contribute at Macroscale level. These factors are indicative of ideal suitability at the Macroscale thus should be thoroughly assessed and incorporated. Mesoscale 1 model (regional level) with Nagelkerke R square of 0.801 identified wind power class, elevation, and distance to airport, city, transmission line, highway, and population density as significant factors for site suitability. Mesoscale 2 (micro-regional) with Nagelkerke R square of 0.794 identified wind power class, elevation, distance to city, river, and transmission line as predictors for site suitability. While, Microscale with Nagelkerke R square of 0.784

identified elevation, distance to river and city as significant factors for predicting suitable site at this scale. As scale changes, factors importance and significance also changes. Overall, elevation, proximity to neighboring turbine, and distance to city are factors that do not appear to be impacted by scale. In contrast, other suitability factors importance or significance appears to be scale dependent. This indicates localized nature of factors so it's ideal to conduct local measurements and assessment to determine site suitability.

The goal of this research was to develop a spatially explicit scale dependent modeling framework for wind farm site suitability assessment. The framework developed, incorporates the technical, environmental and social constraints for siting wind energy development. The framework is based on multiscale empirical module derived from spatial lag regression and machine-learning algorithm coupled with normative component (regulations and policies). Based on the framework, application case study was conducted at a Mesoscale. The model accuracy Nagelkerke R square of 0.88 indicating a good fit. The framework accounts for the complex technical, environmental, and social constraints to identify suitable sites in Iowa with high accuracy. Even though the framework is developed in the context of Iowa, it can be modified to other geographic locations.

6.2 Limitations

There are several limitations of this study: (1) The models developed and presented in this thesis are based on Iowa context and may not be readily suitable for different geographic locations wind farm development based. (2) The scales were determined based on geographical mean of existing turbine distribution thus one scale encompasses another which can lead to the coverage of the same area at different scales. Results may change if others scale selection principle is adopted. (3) Some of the dataset had limited resolution and maybe not suitable for Mesoscale and Microscale analysis which lead to an emergence of Modifiable Areal Unit Problem (MAUP). (4) The assumption of empirical models are based on existing turbines are located on suitable site; therefore, future development might not accurately project new areas if this assumption is violated. (5) Literature indicates economic factors are critical components but due to lack of data, they are not considered in the models or the framework developed.

6.3 Future Directions

Future work in this research could consist of the following: (1) Incorporate production data to make the empirical modules more robust. In doing so, site suitability assessment can be improved because production output of each turbine and with the environmental, technical, and social constraints can give us a better understanding on what makes a particular site suitable. (2) Develop a real-time predicative web based scale dependent spatial decisions support system (SDSS) based on proposed framework to provide developers, policy makers, and the public with sound assessment and relevant information instantly. The tool should be user friendly so that the seasoned or the average persons can equally operate. (3) Expand the framework to encompass multistate scale to gain insight, plan and develop wind energy projects to meet the 20% wind energy by 2030 initiative set by the DOE. Multistate level framework will enable us to meet the high demands of metropolitan areas, regions and states.

REFERENCES

Acker, T. L., Williams, S. K., Duque, E., Brummels, G., & Buechler, J. (2007). Wind resource assessment in the state of Arizona: Inventory, capacity factor, and cost. *Renewable Energy, 32* (9), 1453-1466.

American Wind Energy Association. (2012a). *Iowa State Fact Sheets*. Retrieved from http://awea.files.cms-plus.com/FileDownloads/pdfs/iowa.pdf

American Wind Energy Association. (2012b). *U.S. Wind Industry Annual Market Report*. Retrieved from http://awea.files.cmsplus.com/images/AWEA_USWindIndustryAnnualMarketReport2012_ExecutiveSummary (2).pdf

AWS Scientific, Inc. (1997). *Wind resource assessment handbook: Fundamentals for conducting a successful monitoring program*. Golden, CO. Retrieved from http://www.nrel.gov/wind/pdfs/22223.pdf

Aydin, N. Y., Kentel, E., & Duzgun, S. (2010). GIS-based environmental assessment of wind energy systems for spatial planning: A case study from Western Turkey. *Renewable and Sustainable Energy Reviews, 14* (1), 364-373.

Baban, S. M. J., & Parry, T. (2001). Developing and applying a GIS-assisted approach to locating wind farms in the UK. *Renewable Energy, 24* (1), 59-71.

Baldwin, R.A. (2009). Use of Maximum Entropy Modeling in Wildlife Research. *Entropy, 11* (4), 854-866. doi: 10.3390/e11040854.

Baller, R., Anselin, L., Messner, S., Deane, G., & Hawkins, D. (2001). Structural covariates of U.S. County homicide rates: incorporating spatial effects. *Criminology, 39* (3), 561-588.

Bennui, A., Rattanamanee, P., Puetpaiboon, U., Phukpattaranont, P., & Chetpattananondh, K. (2007). Site Selection for Large Wind Turbine Using GIS. *PSU-UNS International Conference on Engineering and Environment.*

Blair, N., Hand, M., Short, W., & Sullivan, P. (2008). *Modeling sensitivities to the 20% wind scenario report with the WinDS model*. Paper presented at the American Energy Association (AWEA) WindPower conference, Houston, TX. Retrieved from http://www.nrel.gov/docs/fy08osti/43511.pdf

Dark, S. J., & Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography, 31* (5), 471-480.

Elith, J., Phillips, S.J., Hastie, T., Dudik, M., Chee, Y.E., & Yates, C.J. (2011). A statistical explanation of maxent for ecologists. *Diversity and Distribution, 17*, 43-57. Retrieved March 16, 2014, from http://www.cs.princeton.edu/~schapire/maxent/

EnerNex Corporation and WindLogics Inc. (2004). *Characterization of the Wind Resource in the Upper Midwest. Wind Integration Study - Task 1*. Excel Energy and the Minnesota Department of Commerce. Retrieved from http://www.uwig.org/XcelMNDOCwindcharacterization.pdf

Federal Aviation Administration (2013). Retrieved from https://oeaaa.faa.gov/oeaaa/external/public/publicAction.jsp?action=showCaseDownloadForm

Flowerdew, R. (2011). How serious is the Modifiable Areal Unit Problem for analysis of English census data? *Population Trends, 145,* 106-118.

Freedman, J. (2010). *Iowa, past, and present* [electronic resource] / Jeri Freedman. New York, NY: Rosen Central, 2010.

Garland, J. H. (1955). *The North American Midwest: A regional geography*. New York, NY: Wiley.

Gehlke, C. E., & Biehl, K. (1934). Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material. *Journal of the American Statistical Association, 29,* 169-170.

Geiszler, G., Koppel, J., & Gunther, P. (2013). Wind energy and environmental assessments - A hard look at two forerunners' approaches: Germany and the United States. *Renewable Energy, 51,* 71-78.

Grady, S. A., Hussaini, M. Y., & Abdullah, M. M. (2005). Placement of wind turbines using genetic algorithms. *Renewable Energy, 30* (2), 259-270.

Grassi, S., Chokani, N., & Abhari, R. S. (2012). Large scale technical and economical assessment of wind energy potential with a GIS tool: Case study Iowa. *Energy Policy, 45,* 73-85.

Griffiths, J. C., & Dushenko, W. T. (2011). Effectiveness of GIS suitability mapping in predicting ecological impacts of proposed wind farm development on Aristazabal Island, BC. *Environment, Development and Sustainability, 13* (6), 957-991.

Halvatzis, S., & Keyser, D. (2013). *Estimate Economic Impacts of Utility Scale Wind Power in Iowa*. Retrieved from http://www.nrel.gov/docs/fy14osti/53187.pdf

Hand, M., Blair, N., Bolinger, M., Wiser, R., O'Connell, R., Hern, T., & Miller, B. (2008). *Power system modeling of 20% wind-generated electricity by 2030*. Paper presented at IEEE Power and Energy Society, Pittsburgh, PA.

Hofstrand, D. (2010, June). *Understanding Iowa Corn Suitability Ratings (CSR)* [File C2-86]. Ames: Iowa State University Extension. Retrieved from Ag Decision Maker website: www.extension.iastate.edu/agdm

Houston, D. (2014). Implications of the modifiable areal unit problem for assessing built environment correlates of moderate and vigorous physical activity. *Applied Geography, 50,* 40-47.

Hurlbut, D., & National Renewable Energy Laboratory (U.S.). (2008). *State clean energy practices: Renewable portfolio standards*. Golden, CO: National Renewable Energy Laboratory. Retrieved from http://www.nrel.gov/docs/fy08osti/43512.pdf

Iowa Department of Natural Resources Geographic Information Systems Library (2012). Retrieved from http://www.igsb.uiowa.edu/nrgislibx/

Iowa Energy Center. (2012). *Iowa 50 meters wind maps [Data CD]*. Available from Iowa Energy Center, www.iowaenergycenter.org

Johnson, G.L. (1985). *Wind energy systems*. Retrieved from http://www.fastonline.org/CD3WD_40/JF/JF_OTHER/BIG/Wind%20Energy%20-%20Johnson.pdf

Josimovic, B., & Pucar, M. (2010). The strategic environmental impact assessment of electric wind energy plants: Case study 'Bavanište' (Serbia). *Renewable Energy, 35* (7), 1509-1519.

Kissling, W. D., & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography, 17*(1), 59-71.

Leung, D. Y. C., & Yang, Y. (2011). Wind energy development and its environmental impact: A review. *Renewable and Sustainable Energy Reviews, 15* (8), 1031-1039.

Mann, D., Lant, C., & Schoof, J. (2012). Using map algebra to explain and project spatial patterns of wind energy development in Iowa. *Applied Geography, 34,* 219-229.

Marmidis, G., Lazarou, S., & Pyrgioti, E. (2008). Optimal placement of wind turbines in a wind park using Monte Carlo simulation. *Renewable Energy, 33* (7), 1455-1460.

Menz, F. C., & Vachon, S. (2006). The effectiveness of different policy regimes for promoting wind power: Experiences from the states. *Energy Policy, 34* (14), 1786-1796.

Miller, G. A., & Iowa State University. (2012). *Corn suitability ratings: Index to soil productivity*. Ames, Iowa: Cooperative Extension Service, Iowa State University. Retrieved from https://store.extension.iastate.edu/Product/Corn-Suitability-Ratings-An-Index-to-Soil-Productivity

Mohandes, M., Rehman, S., & Rehman, S.M. (2011). Estimation of wind speed profile using adaptive neuro-fuzzy inference systems (ANFIS). *Applied Energy, 88*(2011), 4024-4032. Retrieved from http://dx.doi.org/10/1016/j.apenergy.2011.04.015

Moller, B. (2006). Changing wind-power landscapes: regional assessment of visual impact on land use and population in Northern Jutland, Denmark. *Applied Energy, 83* (5), 477-494.

Mosetti, G., Poloni, C., & Diviacco, B. (1994). Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm. *Journal of Wind Engineering and Industrial Aerodynamics, 51* (1), 105-116. http://dx.doi.org/10.1016/0167-6105(94)90080-9

Nadai, A. (2007). "Planning", "siting" and the local acceptance of wind power: Some lessons from the French case. *Energy Policy, 35* (5), 2715-2726. doi: 10.1016/j.enpol.2006.12.003

Nelson, H. L. (1967). *A geography of Iowa*. Lincoln: University of Nebraska Press.

New York State Energy Research and Development Authority. (2010). *Wind Resource Assessment Handbook*. Retrieved from http://www.inscc.utah.edu/~krueger/5270/wind_resource_handbook.pdf

Nguyen, K. Q. (2007). Wind energy in Vietnam: Resource assessment, development status and future implications. *Energy Policy, 35* (2), 1405-1413.

O'Brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity, 41,* 5, 673-690.

Openshaw, S. (1983). *The modifiable areal unit problem*. Norwich, UK: [Norfolk]: Geo Books.

Openshaw S. (1984). *The Modifiable Areal Unit Problem*. Norwich, UK: GeoBooks.

Openshaw, S., & Taylor P (1979). A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In Wrigley N (Ed.) *Statistical applications in the spatial sciences* (pp.127-144). London, UK: Pion.

Parisien, M.A., & Moritz, M. A. (2009). Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecological Monographs, 79 (*1), 127-154.

Petrov, A. N., & Wessling, J. M. (2014). Utilization of machine-learning algorithms for wind turbine site suitability modeling in Iowa, USA. *Wind Energy, 1*. doi: 10.1002/we.1723

Phillips, S. J., Anderson, R.P., & Schapire, R.E. (2006). Maximum entropy modelling of species geographic distribution. *Ecological Modelling, 190* (3-4), 231-259.

Phillips, S.J., & Dudik, M. (2008). Modeling of species distribution with maxent: new extensions and a comprehensive evaluation. *Echography, 31* (2), 161-175.

Phillips, S.J., Dudik, M., & Schapire, R.E., (2004). A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on Machine Learning* (pp. 655-662). New York, NY: ACM Press.

Portman, M. E., Duff, J. A., Koppel, J., Reisert, J., & Higgins, M. E. (2009). Offshore wind energy development in the exclusive economic zone: Legal and policy supports and impediments in Germany and the US. *Energy Policy, 37* (9), 3596-3607.

Ramirez-Rosado, I. J., Garcia-Garrido, E., Fernandez-Jimenez, L. A., Zorzano-Santamaria, P. J., Monteiro, C., & Miranda, V. (2008). Promotion of new wind farms based on a decision support system. *Renewable Energy, 33* (4), 558-566.

Rodman, L. C., & Meentemeyer, R. K. (2006). A geographic analysis of wind turbine placement in Northern California. *Energy Policy, 34* (15), 2137-2149.

Sathyajith, M. & Philip, G.S. (2011). *Advances in wind energy and conversion technology*. Berlin, Germany: Springer.

Slattery, M. C., Johnson, B. L., Swofford, J. A., & Pasqualetti, M. J. (2012). The predominance of economic development in the support for large-scale wind farms in the U.S. Great Plains. *Renewable and Sustainable Energy Reviews, 16* (6), 3690-3701.

Sowers, J. (2006). Fields of opportunity: wind machines return to the plains. *Great Plains Quarterly, 26 (*2), 99-112.

Swofford, J., & Slattery, M. (2010). Public attitudes of wind energy in Texas: Local communities in close proximity to wind farms and their effect on decision-making. *Energy Policy, 38* (5), 2508-2519.

Tegou, L. I., Polatidis, H., & Haralambopoulos, D. A. (2010). Environmental management framework for wind farm siting: Methodology and case study. *Journal of Environmental Management, 91* (11), 2134-2147.

Toke, D., Breukers, S., & Bolsink, M. (2008). Wind power deployment outcomes: how can we account for the differences? *Renewable and sustainable energy reviews, 12* (4), 1129-1147.

U.S. Census Bureau. (2010). *Iowa census tract 2010*. Retrieved from ftp://ftp2.census.gov/geo/pvs/tiger2010st/19_Iowa/

U.S. Department of Energy (DOE), Energy Efficiency and Renewable Energy. (2008). *20% wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply*. Retrieved from http://www.nrel.gov/docs/fy08osti/41869.pdf

U.S. Department of Energy, (DOE). Energy Efficiency and Renewable Energy. (2011). *2012 wind technologies and market report*. Retrieved from http://www1.eere.energy.gov/wind/pdfs/2011_wind_technologies_market_report.pdf

U.S. Department of Energy, (DOE). Energy Efficiency and Renewable Energy. (2012). *2012 wind technologies and market report*. Retrieved from https://www1.eere.energy.gov/wind/pdfs/2012_wind_technologies_market_report.pdf

Van Hoesen, J., & Letendre, S. (2010). Evaluating potential renewable energy resources in Poultney, Vermont: A GIS-based approach to supporting rural community energy planning. *Renewable Energy, 35* (9), 2114-2122.

Wiser, R., Namovicz, C., Gielecki, M., & Smieth, R. (2007). The experience with renewable portfolio standards in the United States. *Electricity Journal, 20* (4), 8-20.