UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

PROJECTING FUTURE LOCATIONS FOR COMMERCIAL WIND ENERGY DEVELOPMENT IN THE CONTERMINOUS UNITED STATES USING A LOGISTIC REGRESSION-CELLULAR AUTOMATA MODEL

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

JOSHUA JON WIMHURST
Norman, Oklahoma
2023

PROJECTING FUTURE LOCATIONS FOR COMMERCIAL WIND ENERGY DEVELOPMENT IN THE CONTERMINOUS UNITED STATES USING A LOGISTIC REGRESSION-CELLULAR AUTOMATA MODEL

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL SUSTAINABILITY

BY THE COMMITTEE CONSISTING OF

Dr. John Scott Greene, Chair

Dr. Elinor Martin

Dr. Jennifer Koch

Dr. Xin (Selena) Feng

Dr. Anthony Levenda

Acknowledgements

In some ways it is hard to believe I am even writing what I am about to write. When I arrived at the University of Oklahoma six years ago to start my Master's degree in Environmental Sustainability, the eventual completion of my doctorate felt so far away that it would seemingly never come. I also did not have the faintest idea of what my time as a graduate student would have in store; the research direction I would take; who I would meet; the opportunities that would come my way. The completion of my Master's thesis brought about an interest in building my own model pertaining to wind energy development, rather than using wind speed data that other institutes' climate models had been used to generate. This dissertation is the outcome of that initial idea that I had during my thesis defense four years ago. What followed were months of collecting and processing secondary datasets, several more months of writing Python code to process the data and construct usable outputs from it, a year of writing publications, and now a dissertation that summarizes all of my hard work. While my initial ideas evolved over time, my original aspiration to build a model that could be used to determine where commercial wind farms should be built across the United States never changed.

Here I now stand as a doctoral candidate, on the precipice of receiving my PhD in Geography and Environmental Sustainability. I have walked across the graduation stage, I have started a postdoctoral program at the South Central Climate Adaptation Science Center, and the three manuscripts that went into this dissertation are already in the publication process, one of which has already been published at the time of writing this dissertation. I have a long history of impostor syndrome and self-deprecation about my own abilities, but for the first time in my life I feel confident enough to say that I am proud of what I have accomplished, and that I am a capable scientist.

Obviously, I could not have achieved any of this without the love and support of so many people, particularly during the darkest moments that I faced throughout my time as a graduate student. After attempting suicide in August 2021, I thought for a while that the dream I had had since being 14 years old of completing a PhD was over, but it was because of the incredible people I get to call friends, family, confidantes, and soulmates that I was able to carry on after giving myself time to recover. To the people in the Oklahoma drag community, who I leant on heavily in those last few months of 2021 (particularly Wyatt "Ouch" Smith, Kenna Green, Luis Barajas, Sage Runsabove, and Rolando Hernandez), thank you for being there for me to turn to when I needed you. To Claire Marlatt: thank you for giving me a roof over my head in the final weeks of 2021. You are an unfailingly kind person, and without that kindness I don't know that I would have remained well-adjusted into the years that followed. To Trey Lee: we have supported each other in our research endeavors and personal lives since I first met you. Having you as a close friend within the Geography department has been invaluable. To Cass Wantland, Matthew Pearson, and Stephen Whitmore: you took me in last year when I didn't know I needed a group of soulmates to give me space to understand myself and rediscover an authenticity that I didn't realize I had suppressed for so long. The three of you mean more to me than I know how to put into words, and you have been there through all the tears and all the stress as I worked to complete this dissertation; I love you all. To Dedrick Perkins: you are a new friend in my life but I feel like we already know so much about each other. Having you by my side in the homestretch toward my dissertation's end has been more than I knew I needed, especially as I was coming to terms with things in my personal life, and for that I thank you. To my advisor, Dr. Scott Greene: since I met you seven years ago, you have become as much a friend as you have been my mentor. Thank you for being unwaveringly supportive of all of my research ideas, thank you for

allowing me to be my authentic self, and thank you for never leaving my side when my needs for your counsel extended beyond academia. To Shannon Hanchett: my dearly departed sweet friend. I will never forget you or everything we accomplished in Norman together. I know you are smiling and cheering me on as my dissertation journey draws to a close. And finally, to my wonderful parents, Lisa and Guy Wimhurst: you are the most important people in the world to me. Six years ago, you let me travel halfway across the planet to pursue an academic passion, and not once have you ever stopped believing in me. You are my biggest cheerleaders, and the unconditional love you have shown me throughout my entire life is something for which I can never thank you enough. I hope you are as proud of me for completing my PhD as I am.

Table of Contents

List of Tables

List of Figures

Abstract

Pressures to decarbonize the United States' electricity production, reduce dependence on foreign energy imports, and the declining levelized cost of renewable electricity is making wind energy an increasingly appealing means of meeting electricity demand in the United States. However, the installation of new commercial wind farms to meet this demand requires knowledge of the most suitable locations for their installation, which depends on a combination of environmental, technical, economic, political, and social characteristics. Wind Farm Site Suitability (WiFSS) models are frequently enlisted to assist in this decision-making process in countries around the world for both onshore and offshore wind farm siting decisions. However, existing WiFSS models serve to assess present-day wind farm siting potential, rather than project specific locations for future wind energy development. Taking cues from Socio-Environmental Systems (SES) models of urban growth, this dissertation presents a Logistic Regression-Cellular Automata (LRCA) model, henceforth referred to as WiFSS-LRCA, conceived to produce maps that identify scenarios of potential future locations and timing of future commercial wind farms across the Conterminous United States (CONUS) between now and the year 2050.

Following a review of existing WiFSS modeling approaches, and of common practices by which WiFSS modeling studies select and represent their predictors, the niche that WiFSS-LRCA serves to fill was consequently identified. The majority of WiFSS studies take a Geographic Information Systems-based Multi-Criteria Decision Analysis (GIS-MCDA) approach that combines spatial data layers corresponding to selected predictors to construct a composite suitability surface. Other common approaches include Non-GIS-MCDA models that rank discrete potential wind farm sites to prioritize their order of development, Bayesian Network (BN) models that construct and convey probabilistic relationships between predictors, and

Logistic Regression (LR) models that perform either spatial or non-spatial assessment of a wind farm's suitability of presence based on the log-odds of a linear combination of predictors. The common limitation of these modeling approaches is their lack of a temporal component, meaning that they can assess WiFSS only at a single point in time. WiFSS-LRCA fills this niche by combining an LR equation with the decision rules of Cellular Automata (CA) to iteratively advance the computed probabilities of each grid cell, based on areas constrained from development and neighboring grid cells that already contain wind farms.

WiFSS-LRCA enlists a large set of predictors ranging from wind speed to legislation in effect in order for the model to represent the influence that environmental, technical, economic, political, and social predictors have on wind farm siting decisions. Data were aggregated at 20 different grid cell resolutions, collated in four different predictor configurations, and adjustments to the model's constraint, neighborhood effect, and equation-based scenario transition rules were incorporated into the model's construction, facilitating WiFSS-LRCA's sensitivity and scenario analysis of model outputs by end-users. WiFSS-LRCA incorporates both calibration of its LR equation's predictors and validation of the model's performance to determine its ability to correctly identify the observed locations of present-day wind farms. Subsequently, the model constructs a WiFSS map whose interpretation and predictive accuracy are informed by the calibration and validation process. Construction of scenarios that modify WiFSS-LRCA's predictors allow for the model to consider the impacts of changes in these predictors on the locations of future wind energy development (e.g., new transmission line construction, opinions of wind energy improving with time, increasing temperatures due to climate change).

The ability of WiFSS-LRCA to produce suitability surfaces with verifiable accuracy is greatest under the following conditions: when running the model over an individual U.S. state rather than

the CONUS, when using a smaller grid cell size, when using a more complete (*Full* configuration) or more refined (*Reduced* configuration) set of predictors, and when the selected study area contains a larger number of present-day commercial wind farms. Across most study areas, however, WiFSS-LRCA is typically able to correctly identify 75-85% of grid cells that do and do not contain commercial wind farms, with these classifications most often associated with high wind speed, proximity to transmission lines, legislation that supports wind energy development, and large tracts of undeveloped land. CONUS-level model runs indicate five regions as being the most suitable for present wind energy development: Southern California, the Pacific Northwest, the Central Plains, the Great Lakes, and the Northeastern United States. CONUS-level model runs have a tendency to over(under)-estimate grid cell probabilities within (outside) the Central Plains and Great Lakes, which makes state-level model runs useful for revealing smaller-scale differences in the probabilities computed within these five broad regions.

Subsequent iterations of WiFSS-LRCA out to the year 2050 show projected wind energy development to remain concentrated within these same regions. Many of the grid cells initially classified as false positive in the model's first iteration are those that gain wind farms in subsequent iterations, particularly false positive grid cells that were part of high-probability hotspots identified by Getis-Ord statistics. Running WiFSS-LRCA over states outside of these five regions projects wind energy development potential in low-probability areas (as shown in this dissertation for Florida and Kentucky) with projected wind farms in these states concentrated closer to existing infrastructure and away from protected natural areas. The Odds Ratios (ORs) computed during WiFSS-LRCA's initial calibration provide geographical insight into its projections, with grid cells characterized by high wind speed, undeveloped land, and ambitious Renewable Portfolio Standards (RPS) being the most likely to gain wind farms in future decades.

The model's projections are, however, shown to be sensitive to end-user definitions of parameters, with neighborhood effect and constraint definitions greatly affecting the location and timing of projected wind farm locations. The scenario setup, by contrast, is shown to mostly influence the timing of these projections, with grid cell size moderately affecting both.

Multiple limitations exist in the application and interpretation of WiFSS-LRCA. Firstly, the lack of existing LRCA approaches to assessing wind farm siting potential meant few standards existed to guide this model's development, such as the setting of default constraints and establishing cutoff statistics for refining the model's enlisted predictors. Secondly, the use of an LR equation to construct suitability surfaces in the model's first iteration means that both classes of the dependent variable must be filled, requiring a study area to contain at least two commercial wind farms, compromising the model's reliability in runs over the Southeastern United States. Finally, the lack of spatial stratification during WiFSS-LRCA's calibration and validation means that the model is trained to recognize predictors associated with wind energy development in regions where many wind farms exist, namely the Central Plains and Great Lakes, hence the greater number of Type 2 errors in CONUS-level model runs outside of these regions. Selecting stratified samples of grid cells that contain wind farms from different parts of the CONUS could be incorporated into WiFSS-LRCA to address this bias. Other directions for future work with WiFSS-LRCA include the following: optimization to assess offshore wind energy development potential by training the model with proposed offshore wind farm sites surrounding the CONUS; adapting WiFSS-LRCA to run over multiple states simultaneously to identify predictors that influence wind farm siting decisions at regional spatial scales; and performing projections of other types decentralized land-use change, such as solar energy development given similarities in the required model predictors.

Chapter 1: Introduction

1.1. Scope of Dissertation

Global wind energy capacity expanded greatly in the last decade, increasing from 300 Gigawatts (GW) installed in 2013 to 899 GW in 2022 [1]. This expansion has occurred in nations and trade blocs across the planet, with 289 GW of wind energy capacity installed in China by 2020 [2], the European Union receiving 16.4% of its consumed electricity from wind energy in 2021 [3], and the United States similarly reaching 10.2% wind-derived electricity as of 2023 [4]. Future plans for continued adoption of renewable energy exist at multiple scales and take multiple forms, such as the African Renewable Energy Initiative's continent-wide 300 GW by 2030 renewable energy capacity target [5], New Zealand's 50% by 2035 renewable energy consumption target [6], and corporations like BP and Repsol setting "net-zero" targets by 2050 that explicitly require renewable energy development to achieve [7,8].

Among other processes, pressure to decarbonize electricity generation is arguably the major driver of the wind energy sector's current growth [9]. The lifecycle greenhouse gas emissions of wind-derived electricity are lower than those produced by coal or natural gas [10], with Wang et al. [11] noting a three-times lower emission intensity for an onshore wind farm compared to a coal-fired power plant of equivalent capacity. Wind farms also have lower water demands than power plants, since the former do not require water for condensation or fuel processing [12], hence wind energy development is often recommended both for sustainability and climate change mitigation measures [13,14]. Decision-makers commonly select wind energy when seeking to improve energy security, such as small islands that lack a domestic fossil fuel supply [15], regions facing geopolitical tension that may necessitate foreign energy imports [16], and nations ensuring that energy demands are consistently met [17]. Evidence of a declining

Levelized Cost of Electricity (LCOE) for wind energy in many countries [18] suggests onshore wind energy is now cost competitive, or even less expensive, than other electricity sources. Offshore wind energy is more expensive due to sea depth and transmission line challenges [19], though its capital costs continue to fall as larger numbers of high-capacity turbines continue to be built [20]. Finally, gradual improvement over time in social and political perceptions of wind energy among governments [21], the private sector [22], and the general public [23] reduce the likelihood of resistance to future wind energy development.

These drivers of wind energy development raise the question of where new commercial wind farms would be best installed to meet capacity targets and future electricity demands. According to Latinopoulos and Kechagia [24], efforts to meet renewable energy capacity targets require identifying sites for land development that maximize energy production and minimize land-use conflicts [25]. Suitability analysis assists this identification process, defined as: “a process of systematically [spatially] identifying or rating potential locations with respect to a particular use” [26]. Suitability analysis performed using environmental models simplifies the interconnecting processes that influence site identification, thus benefiting understanding of system behavior [27]. Kelly et al. [28] note that improved system understanding provided by environmental models benefits the decision-making process, in this case decisions of where to build commercial wind farms. Wind Farm Site Suitability (WiFSS) models specifically enlist datasets for a collection of predictors, with each predictor exerting an influence on the wind farm siting process. Besides the most obvious predictor of whether sustained wind speeds are in the optimal range for wind energy production [29], other key predictors may include distance to transmission lines [30] and levels of public support [31], each of which pose a potential barrier to a location’s

wind energy development. All WiFSS models have a goal of combining predictor datasets to conduct assessments of where wind farms are best installed.

Using WiFSS models to inform wind farm siting decisions has advantages beyond strategizing pathways toward wind energy capacity targets. Firstly, including predictors such as bird habitat ranges and public opinion allows WiFSS models to recommend development in areas less likely to face public backlash, which is important given the incidents of high bird mortality rates [32,33] and forced cancelations [34,35] of some past wind farm projects. Secondly, due to the risk involved in investing into wind energy over established technologies [36], WiFSS models are valuable to economists interested in valuing a potential wind farm site's return on investment. WiFSS models incorporate economic concerns in different ways, with some studies adjoining equation-based analysis to assess profitability of development sites [37,38] and others instead including predictors that represent project costs [39]. Finally, WiFSS models frequently enlist Geographic Information Science (GIS) to present their outputs in map form [40], thus providing communicable results for informing policymaking efforts [41], as well as educational tools for academics and non-experts [42]. Continued growth of the wind energy sector increases pressure to construct wind farms in areas that are profitable, socially acceptable, and low in environmental impact, needs that WiFSS models are built to address.

The most common WiFSS models are those that take a Multi-Criteria Decision Analysis (MCDA) approach [43], which in a GIS context means combining geospatial datasets representing different predictors to construct a composite suitability surface [44]. GIS-MCDA approaches to WiFSS have been conducted on every continent except Antarctica [45-50]. Non-spatial approaches to MCDA (i.e., approaches that do not use GIS) typically collect expert opinions about each predictor to rank a small set of discrete wind farm sites [51,52]. However,

both GIS and Non-GIS MCDA approaches may (or may not) apply weighting schemes (such as the Analytic Hierarchy Process (AHP) [53,54], Best-Worst Method (BWM) [55] or Ordered Weighted Averaging (OWA) [56]) to account for each predictor's relative importance to wind farm siting decisions. Machine learning algorithms that are trained and tested to predict potential wind farm locations are an alternative to MCDA [57], particularly those based on logistic regression equations [58,59]. Logistic regression's ability to combine discrete and continuous data with limited error [60] means easier incorporation of implicitly spatial predictors [61], such as demographics and election results, hence their inclusion by Harper et al. [62] and Roddis et al. [63] for spatially assessing public acceptance of wind farm projects. Bayesian networks, though less commonly used in WiFSS modeling studies, construct joint probability distributions between the possible states of each predictor to assess both onshore [64] and offshore [65] WiFSS. Although each has a different theoretical basis and results presentation, each WiFSS modeling approach specifically assesses current wind farm installation potential. Development of a logistic regression-based WiFSS model will be the focus of this dissertation.

While all WiFSS models present deterministic and/or probabilistic assessments of locations for building new wind farms, existing models do not include a temporal component that allows the locations most suitable for development first to be dynamically identified. Such capability would allow model users to prioritize tracts of land for wind energy development at specific future times, rather than selecting among many high-probability locations. A temporally (as well as spatially) explicit component would also allow WiFSS models to incorporate scenario-based modeling, such as varying a model's predictors and parameters to account for a range of possible future land-use changes [66]. One means of adding a temporal component to land-use change models is through iterative updates to the modeled surface, such as by using Cellular Automata

(CA) to simulate dynamic spatial interactions by applying transition rules to a model's gridded surface [67]. White and Engelen [68] noted CA to be effective at simulating city- to region-level land-use change, specifically urban sprawl, of which multiple published examples exist [69-71]. The "neighborhood effect" transition rule of CA allows them to capture cluster-like patterns of land-use change over time, a characteristic feature of urban sprawl [72] and one that commercial wind farm installations in the Conterminous United States (CONUS) also exhibit [73]. Incorporation of CA into a WiFSS model should thus allow specific areas for future wind energy development to be prioritized temporally, with this work using the CONUS as a case study.

1.2. Justification for this Work.

This dissertation's objective is to present a model that combines a Logistic Regression (LR) equation and CA decision rules to project suitable locations for commercial wind energy development across the CONUS out to the year 2050. The LR equation provides the CA's "equation-based" transition rule, which governs the state of all relevant predictors (e.g., wind speed, distance to transmission lines, vulnerable species, etc.) within each individual grid cell of the CA. Since most published examples of Logistic Regression-Cellular Automata (LRCA) modeling are in urban sprawl contexts [74-76], application to nationwide commercial wind farm projects represents a much larger spatial scale than existing LRCA applications. The CONUS was selected as this work's case study for three reasons:

1. *The predictors relevant to WiFSS are easily representable thanks to publicly available datasets.* Plassin et al. [77, p.1] comment that datasets prepared within different disciplines and geographical contexts result in inconsistent data formatting and content, thus "limit[ing] the study of transboundary socio-environmental systems". The United States' federal agencies maintain online access to datasets for many of the predictors that

affect the wind farm siting process, such as the Energy Information Administration's (EIA) state-averaged electricity costs [78], the United States Geological Survey's (USGS) digital land elevation [79] and land cover type rasters [80], and the United States Census Bureau's county-averaged demographics [81,82]. Consistently formatted and easily accessible datasets across such a large spatial domain circumvent the transboundary concern, while also enabling WiFSS analysis across the CONUS from local to national spatial scales.

2. *Predictors can take wide-ranging values across the CONUS.* The CONUS' size means that predictors are seldom constant across the entire study area. For instance, average wind speeds tend to be greatest over the Central Plains [83,84], wind energy markets are more mature in states such as California and Texas [85,86], and public support for wind energy is typically higher in Democrat-leaning states [87,88]. This variability should allow a WiFSS model to capture spatial differences in each predictor's influence on wind farm siting decisions.
3. *The United States' federal government is currently advocating for renewable energy development.* Recent federal initiatives seek to expand the renewable energy sector. Examples include newly enacted capacity targets for offshore wind energy development [89], a 30% investment tax credit scheme for commercial [90] and residential [91] renewable energy projects, and the Department of Energy's *Wind Vision* report detailing a strategy to meet 35% of nationwide electricity demand with wind energy by 2050 [92]. As such, a WiFSS model that projects suitable wind energy development locations is beneficial for strategizing how to realize these targets and to understand the required land demands for this development.

The first purpose of this work is *to demonstrate an LRCA model's ability to project a range of wind farm siting futures that can be explained geographically*. LR equations in this context yield the probability of dichotomous land-use change taking place, based on a trained and tested relationship with the equation's predictors [93]. Existing WiFSS studies express these relationships using Odds Ratios (ORs) [62,94], such that each predictor's OR communicates the association between the predictor and whether a location is projected to gain a commercial wind farm. For instance, an OR for wind speed is likely to be greater (i.e., more positively associated) when running this LRCA model (henceforth referred to as WiFSS-LRCA) over a state that possesses optimal wind speed for wind energy development, suggesting wind speed's importance for wind farm siting decisions. Such interpretation of ORs should be done knowing the limitations of weak associations [95] and confounding predictors [96]. The "range" of siting futures comes from constructing scenarios by modifying the LR equation's trained and tested relationships and observing the effects on WiFSS-LRCA's projections. Sohl et al. [97, p.108] note that scenario-based modeling should "develop a meaningful understanding of how various driving forces that caused landscape change interact with other factors to determine a regional outcome". WiFSS-LRCA simulates said interactions by creating future scenarios (e.g., infrastructure development, climate change, changing demographics) that modify the relationships of groups of predictors in a similar vein to Yang et al. [98], allowing wind farm locations to be projected under multiple potential circumstances.

This work has a second purpose *of showing that LRCA modeling can represent spatial scales and processes larger than its more common application to localized land-use change.* WiFSS-LRCA takes cues from the various urban sprawl examples in the literature, particularly Shu et al.'s [76] simulation of observed sprawl in China's Xuzhou Prefecture, and Mustafa et al. [71] capturing

recent urban densification across Southern Belgium. These studies selected an LRCA approach specifically for the combination of repeated calibration and simplified spatial dynamics offered by the LR and CA components, respectively, advantages similarly noted by Shahbazian et al. [99] in their projections of future land cover change across Iran's Ilam Province. These examples, however, enlist grid cells that cover small land tracts or groups of buildings (< 100 meters across), with grid lattices not exceeding county-level in scale. By contrast, wind turbines within the same wind farm can be situated up to kilometers apart due to micrositing limitations [100], necessitating larger grid cells and thus "neighborhoods" that cover greater distances than in existing LRCA applications. Liao et al. [101] found that the decay effect that neighborhood effects represent in CA-based models depends partly on the physical distance between developed and undeveloped land units. As such, the appropriateness of applying neighborhood effects to wind farm siting on a CONUS-wide scale is addressed in WiFSS-LRCA's presentation.

The final purpose of this dissertation is *to present a WiFSS model that is ready for immediate use by a wide range of audiences, including political decision-makers and non-experts.* Increasing public skepticism about the production of science makes its communication in accessible terms important [102]. Indeed, Grimm et al. [103, p.129] emphasize the value of "linking modelers and model users, for example stakeholders, decision makers, and developers of policies" through documentation of the model development process. To that end, WiFSS-LRCA's development takes several steps to ensure readiness for end-use, such as aggregation of its predictor datasets at multiple grid cell resolutions, construction of multiple scenarios, and documentation that instructs users on how to run WiFSS-LRCA for themselves. Uploading model resources to online repositories facilitates a culture of no-cost model amendment and application by other scientists and stakeholders for their specific purposes [104,105] (see Appendix A1 for GitHub link). As

such, WiFSS-LRCA aids decision-making efforts for end-users wishing to perform their own experiments on wind farm siting futures across the CONUS, with WiFSS-LRCA's robustness for this purpose being explored throughout this dissertation.

The following three research questions are posed based on the case study context and the purposes of this work, along with the Chapters in which each question is addressed:

1. **Where is Logistic Regression-Cellular Automata situated within the broader scope of Wind Farm Site Suitability modeling approaches?** Chapter 2 details the outcome of a systematic literature review of existing WiFSS studies, highlighting the most common modeling approaches and the lack of standardization in predictor selection and representation.
2. **What are currently the most suitable locations for present wind energy development across the CONUS?** After first detailing WiFSS-LRCA's data selection, training, and testing techniques in Chapter 3, Chapter 4 presents outputs from running the model's LR equation across the CONUS and a collection of states, ultimately attesting to LR's ability to capture present commercial wind farm locations.
3. **Which regions of the CONUS (at nationwide and state-level scales) are projected to acquire wind farms out to the year 2050, and what geographical features may explain these projections?** Combination of the LR equation with Cellular Automata to illustrate common patterns in projected locations, along with a sensitivity analysis of the projections to changes in WiFSS-LRCA's parameters (grid cell size, neighborhood size, constraints, scenario setup) are presented in Chapter 5. Chapter 6 concludes the dissertation by discussing implications of WiFSS-LRCA's outputs, the limitations of the overall modeling process, and possible directions for future work.

Chapter 2: Background – Summarizing Existing WiFSS Studies.

2.1. Common Approaches to WiFSS Modeling.

2.1.1. GIS-Based MCDA.

Combining GIS and MCDA can be defined as "a process that transforms and combines geographical data and value judgments to obtain information for decision making" [61, p.703]. According to Cowen [106], GIS explicitly supports problem solving through the integration of data with a common spatial component, which MCDA aids by providing the decision making framework for solving said problems. Such a broad scope means that GIS-MCDA has been used to perform suitability analysis across multiple disciplines and contexts, from bicycle facility planning in Wisconsin [107] to identifying high flood risk zones in Iran [108]. Eastman [109] notes that the spatial datasets GIS-MCDA combines to construct a suitability surface typically take one of two forms: Booleans that constrain locations' development potential by classifying them as either suitable or unsuitable, and predictors whose suitability is evaluated continuously in space and often weighted to account for relative importance compared to other predictors. In the context of performing WiFSS analysis using GIS-MCDA, example Booleans might be limiting land slope to 10% or less due to the relative ease of building wind turbines on flat terrain [110], or not allowing wind energy development within 500 meters of cultural monuments [111]. Examples of continuous evaluation in space include suitability increasing with proximity to transmission lines [112] and in areas with greater average wind speed [49]. Different GIS-MCDA approaches to WiFSS often use different predictors, and studies that do use the same predictors may enlist them as Boolean constraints and/or for continuous evaluation [40]. However, all GIS-MCDA approaches have a common objective of layering spatial datasets to construct composite suitability surfaces that identify suitable locations for wind energy development.

Maps that illustrate likelihoods or scores for wind energy development based on combining dataset layers are the primary output of a GIS-MCDA approach to WiFSS. Figure 1 shows an example of the maps typically produced by GIS-MCDA, taken from Villacreses et al.'s [47] study of WiFSS in Ecuador. These maps show the areas unsuitable for development based on Boolean constraints (Figure 1a), continuous evaluation in space of nine core predictors (Figure 1b), and the final suitability surface that combines the weighted predictors and constraints (Figure 1c). Since GIS-MCDA is the most common approach for performing WiFSS assessment [40], multiple variants have been developed on the overall modeling technique in terms of predictor selection, criteria weighting, and results validation, as recent reviews by Rediske et al. [113] and Shao et al. [114] have covered. Older approaches generally enlist a smaller set of predictors with a more simplistic implementation. One such example is Baban and Parry's [115] 2001 study of WiFSS in the United Kingdom, in which all predictors are enlisted as Boolean constraints and each predictor's dataset is weighted based on perceived importance by the model developers. By contrast, Elkadeem et al.'s [48] 2021 study of WiFSS in Kenya exhibits techniques of more contemporary GIS-MCDA approaches, such as citing previous studies to justify Boolean constraints, employing additional techniques to weight predictors for their importance (Best Worst Method; BWM [116]), and ranking map locations with the greatest wind energy development potential (Technique for Order Preference by Similarity to an Ideal Solution; TOPSIS [117]). Techniques like BWM and TOPSIS, and especially the Analytic Hierarchy Process (AHP) [118], represent attempts at quantifying aspects of GIS-MCDA that involve human decision making, such as using AHP to numerically derive criteria weights for all predictors [44] and incorporating fuzzy logic to account for the subjectivity of expert opinions when deriving said weights [119].

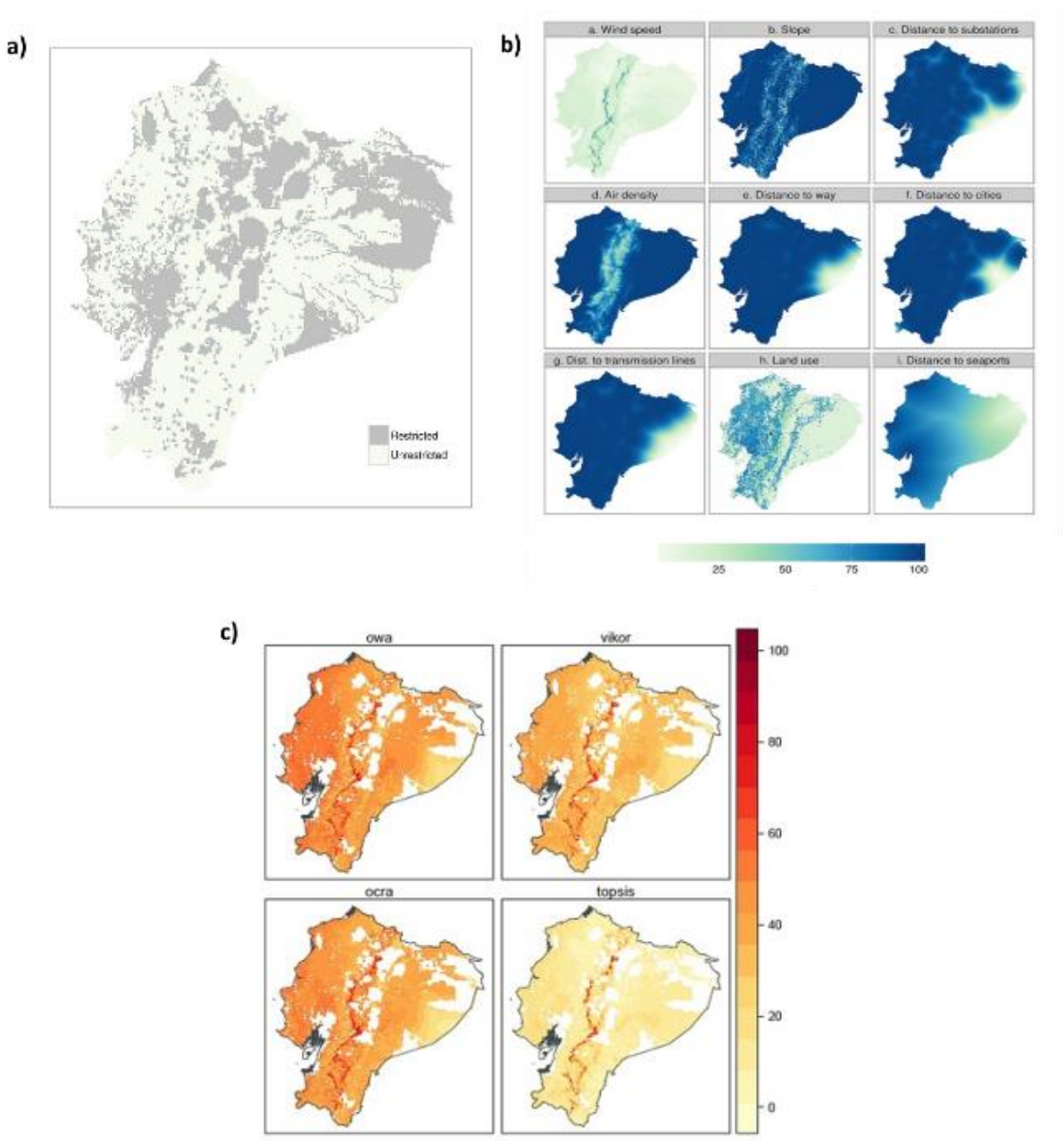


Figure 1: Example maps produced by a GIS-MCDA approach for assessing wind farm site suitability. Figure 1a shows locations in Ecuador identified as (un)suitable for wind energy development based on Boolean constraints. Figure 1b shows the continuous evaluation of each individual predictor. Figure 1c shows the result of applying four different weighting schemes when merging Figure 1b, with the white zones indicating the constrained locations from Figure 1a. Source: Villacreses et al. [47].

Gonzalez and Enríquez-de-Salamanca [120] identify several advantages of GIS-MCDA in their review of its applications to environmental contexts. These advantages include making suitability analysis open to exploration and scenario building, ease of combination of heterogeneous datasets, and applicability across multiple geographical contexts and spatial scales. Another cited advantage of GIS-MCDA is its readiness for non-expert participation, as done by Mekonnen and Gorsevski [42] in their creation of a web-based GIS tool that allows communities to inform offshore wind farm siting plans in Lake Erie. Ease of access to publicly available spatial datasets means that predictors relevant to WiFSS can be freely depicted in GIS-MCDA models [40]. Indeed, González-Ramiro et al. [121] noted the value of free access to GIS-MCDA models based on public datasets for decision makers with limited resources. The increasing role that GIS plays in informing government policy [122] is well-suited to WiFSS assessment since many GIS-MCDA studies of renewable energy potential already use current political context to inform their research methods. For example, Giamalaki and Tsoutsos [123] use Greece's sustainable development legislation to constrain areas of Crete prohibited from solar energy development. Refinement across many existing studies [114], such as expansion of enlisted predictors and development of weighting and ranking techniques, has sufficiently prepared GIS-MCDA approaches to WiFSS for making a tangible impact on present wind farm siting decisions.

Several limitations exist in GIS-MCDA's application, as noted by Gonzalez and Enríquez-de-Salamanca [120]. These limitations include missing data necessary for key constraints resulting in false representations of suitable locations and the loss of information from combining data layers to create the composite suitability surface. Another limitation is sensitivity of composite surfaces to the weighting and ranking criteria used to construct them, as illustrated by Figure 1c [47]. This sensitivity also occurs when comparing separate GIS-MCDA studies of WiFSS over

the same study area, such as Li et al. [119] and Xu et al.'s [124] studies of WiFSS in China's Liaodong Peninsula. The former's combination of fuzzy logic and AHP to derive weights for each predictor resulted in a larger region being classified as unsuitable for wind energy development. While these weighting and ranking methods are more sophisticated than assuming equal importance [115] or assigning arbitrary numbers to each predictor [125], newer methods are not immune to subjectivity, since expert opinions are still required for setting predictor weights/ranks [126]. Furthermore, discrete predictors like political affiliation and health conditions are difficult to include in GIS-MCDA models without explicit spatialization [61] or being proxied by a continuous predictor; hence, few examples exist of their inclusion in a WiFSS context [58,127]. WiFSS-LRCA provides a solution to the following limitations: (1) the calibration of its LR equation's removes human decision-making from weighting importance of predictors [128]; (2) LR equations can combine discrete and continuous predictors with limited error by constructing a relationship between the logit of wind farm occurrence and the linear combination of predictors [60,93].

2.1.2. Non-GIS Based MCDA.

Performing environmental assessment (WiFSS or otherwise) using MCDA does not require combination with GIS. According to Huang et al. [129, p.3579], MCDA is used "to discover and quantify decision maker and stakeholder considerations about various non-monetary factors in order to compare alternative courses of action," a definition that does not necessitate spatial analysis. The objective of applying Non-GIS-MCDA to WiFSS nevertheless remains the same, to select suitable locations for wind energy development based on predictors relevant to the wind farm siting process [130]. Non-GIS-MCDA and GIS-MCDA indeed share similarities in their WiFSS evaluation approaches, such as enlisting a comprehensive predictor set including wind

speed, environmental impact, terrain, social acceptability, transportation infrastructure, and many others [131]. Non-GIS-MCDA is also compatible with the aforementioned weighting and ranking techniques. Examples include Rouyendegh et al. [132] combining fuzzy logic and TOPSIS to rank potential wind farm sites in Turkey from most to least suitable, and Wu et al. [133] using AHP to obtain criteria weights for predictors in their assessment of the East China Sea's offshore wind potential. There are, however, two key differences when implementing Non-GIS-MCDA to assess WiFSS, the first being that this approach does not construct composite suitability surfaces, instead presenting results as tables that compare discrete wind farm sites based on chosen predictors and implemented weighting/ranking schemes [134]. The second difference is Non-GIS-MCDA more frequently relies on primary data to inform its predictors, with common practice being to enlist academic or industrial expert opinions to identify and rate the importance of predictors at each discrete wind farm site [51,135].

Figure 2 presents Toklu and Uygun's [136] general application of Non-GIS-MCDA to WiFSS, which uses criteria weights derived from applying AHP (Figure 2a), and data for five predictors collected at three hypothetical wind farm sites (Figure 2b), to determine the best site for wind energy development. Their approach normalizes the predictor data to produce "information content" indices, such that the site possessing the lowest total index is the best for development, identified as Loc-B in Figures 2c and 2d. Combining criteria weights and predictor data to rank wind farm locations from best to worst has long been the baseline for those that use Non-GIS-MCDA's approach to WiFSS assessment. Early examples include Aras et al.'s [137] AHP assessment for placing a wind turbine near Osmangazi University's campus (Turkey), and Gamboa and Munda's [138] evaluation of social predictors and actors that limit wind energy development in Catalonia (Spain). As with GIS-MCDA, application of Non-GIS-MCDA has

a)

Evaluation Criteria	Criteria Weights
C1: Wind speed (WS)	0.362
C2: Wind power density (WPD)	0.29
C3: Capacity factor (CF)	0.245
C4: Distance to power grid (DPG)	0.093
C5: Land roughness value (LRV)	0.01

b)

	Alternative Locations		
	Loc-A	Loc-B	Loc-C
Wind speed (WS)	4-6 m/s	5-7 m/s	3-6 m/s
Wind power density (WPD)	300-500 W/m2	300-600 W/m2	250-450 W/m2
Capacity factor (CF)	% 25-30	%20-40	%15-30
Distance to power grid (DPG)	Very Good	Very Good	Fair
Land roughness value (LRV)	Semi-Rough	Smooth	Semi-Rough

c)

Criteria Weights	0.362	0.29	0.245	0.093	0.01	
Alternatives	I_{WS}	I_{WPD}	I_{CF}	I_{DPG}	I_{LRV}	$\sum I$
Loc-A	1	1	0	0.263	3	5.263
Loc-B	0	0.585	0.415	0.263	1	2.263*
Loc-C	1.584	2	0	1.585	3	8.169

d)

Alternatives	IWS	IWPD	ICF	IDPG	ILRV	$\sum I$
Loc-A	0.362374	0.289863	0	0.000001	1.010739	1.662977
Loc-B	0	0.157259	0.027660	0.000001	0.009723	0.194642*
Loc-C	1.181371	2.000000	0	1.043713	1.010739	5.235822

Figure 2: Example results from using a Non-GIS-MCDA modeling approach to assess three discrete wind farm sites. From top to bottom are the criteria weights obtained from applying AHP (Figure 2a), the predictor data at the three discrete sites (Figure 2b), the transformation of predictor data to produce "information content" indices (Figure 2c), and the product of multiplying these indices by the criteria weights and the summed information content indices for each discrete site (Figure 2d). Source: Toklu and Uygun [136].

been refined over time, with the greatest difference in contemporary methods being the technique used to obtain the final wind farm site rankings. For instance, Wu et al. [139] use the PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations) method to rank four potential offshore wind farm sites in China's Guangdong Province, with this method selected for its modification to the initial rank order based on decision-makers' attitudes toward risk factors, such as extreme weather and profitability. By contrast, Wang et al. [52] take the TOPSIS approach of computing each predictor's "distance" from ideal conditions and thus determining which of seven candidate locations in Vietnam is closest to these conditions. While fewer examples exist of Non-GIS-MCDA's application to WiFSS compared to GIS-MCDA [40], the maturation of applied techniques is similar.

Since Non-GIS-MCDA requires similar decision-making techniques to GIS-MCDA, it shares some of its advantages. For instance, the range of weighting and ranking methods used across Non-GIS-MCDA studies enables exploration of which discrete wind farm sites are the most suitable under different circumstances, and the geographical study context and scale are only limited by where primary data can be collected [120]. Reliance on primary data is an advantage of Non-GIS-MCDA. Whereas GIS-MCDA approaches compile secondary spatial datasets to construct a suitability surface [140,141], Non-GIS-MCDA approaches typically solicit expert opinions from questionnaire responses [130], interviews [138], and focus groups [132] about each predictor's relative importance at each discrete wind farm site. Since "data quality standards are [usually] developed from the perspective of data producers" [142, p.4], Non-GIS-MCDA is less susceptible to poor data quality than WiFSS modeling approaches that rely on secondary data. Furthermore, representing predictors as expert opinions allows Non-GIS-MCDA to incorporate predictors that are often elusive to GIS-MCDA approaches to WiFSS. For instance,

Fetanat and Khorasaninejad's [143] study of offshore wind farm site potential in Iran sought opinions about commercial feasibility, importance of regulations, and aesthetic pollution, predictors that are not explicitly spatial [61] and are thus harder to include in GIS-MCDA models. A final advantage of Non-GIS-MCDA is that access to, or a working understanding of, GIS software is not required. Education in the use of GIS may be prohibited by high costs and lack of software availability [144], particularly in less industrialized nations [145]. A lack of prohibitive technologies and costs makes Non-GIS-MCDA a more internationally accessible means of performing WiFSS assessment than GIS-based options.

Non-GIS-MCDA also shares some of GIS-MCDA's limitations, in particular the sensitivity to the weighting and ranking methods selected by model developers, and the loss of information due to combining and aggregating data for selected predictors [120]. Without the unifying presentation of GIS, Non-GIS-MCDA approaches can vary widely in how they select and represent predictors relevant to wind farm siting. Jun et al. [146] use 13 predictors ranging from wind power density to traffic conditions to rank seven potential wind/solar hybrid power station sites in China, compared to Deveci et al.'s [134] larger set of 23 predictors that explicitly uses existing literature to justify predictor selections. Furthermore, the latter study enlists expert opinions to assist in setting criteria weights for each predictor, whereas the former does not. The many variants of WiFSS assessment using Non-GIS-MCDA, particularly in how predictors are selected and represented, makes comparisons between studies difficult [40]. Moreover, as Figure 2 illustrates, expressing WiFSS across multiple tables is not as easily interpretable as a suitability map. The core result of most GIS-MCDA approaches is a composite suitability surface that illustrates the best and worst places for wind energy development [44,49,110]. This synthesis of data into a single map facilitates communication of interdisciplinary research findings [147] and gives

direction to policymaking efforts [148], primarily due to the increased problem-solving time required to interpret several tables [149]. To that end, much like other Logistic Regression-Cellular Automata modeling studies [76,99], WiFSS-LRCA's primary output is a map showing locations across the CONUS projected to acquire wind farms in the following decades, serving as an easily communicable result for decision-makers.

2.1.3. Bayesian Networks.

According to Ben-Gal [150, p.1], a Bayesian Network (BN) is a "graphical structure used to represent knowledge about an uncertain domain. In particular, each node in the graph represents a random variable, while the edges between nodes represent probabilistic dependencies among the corresponding random variables." BNs thus compute and construct statistical relationships between variables to predict system behavior based on probabilistic connections between said behavior and the system's individual parts. Aguilera et al. [151] note several reasons for a BN's usefulness for environmental modeling, such as its ability to process uncertainty, its accommodation of complex systems, and its openness to participatory modeling that incorporates the priorities of experts and stakeholders. Example applications of BNs as an environmental model range from predicting chemical toxicity to juvenile fish [152] to determining consistent causes of Turkish forest fires [153]. In the context of WiFSS, each "parent" node of a BN represents a single predictor, with the "child" node being the binary outcome of whether a location is suitable for wind energy development. Figure 3 presents the BN produced by Li et al. [65] in their study of offshore wind farm siting potential in the North Sea, with predictors such as wind speed, wave height, and distance to roads and electricity grids among the parent nodes. As in other BN applications to WiFSS, the probability distribution constructed for each parent node conveys the likeliest state(s) of each predictor; data comprising each distribution may be

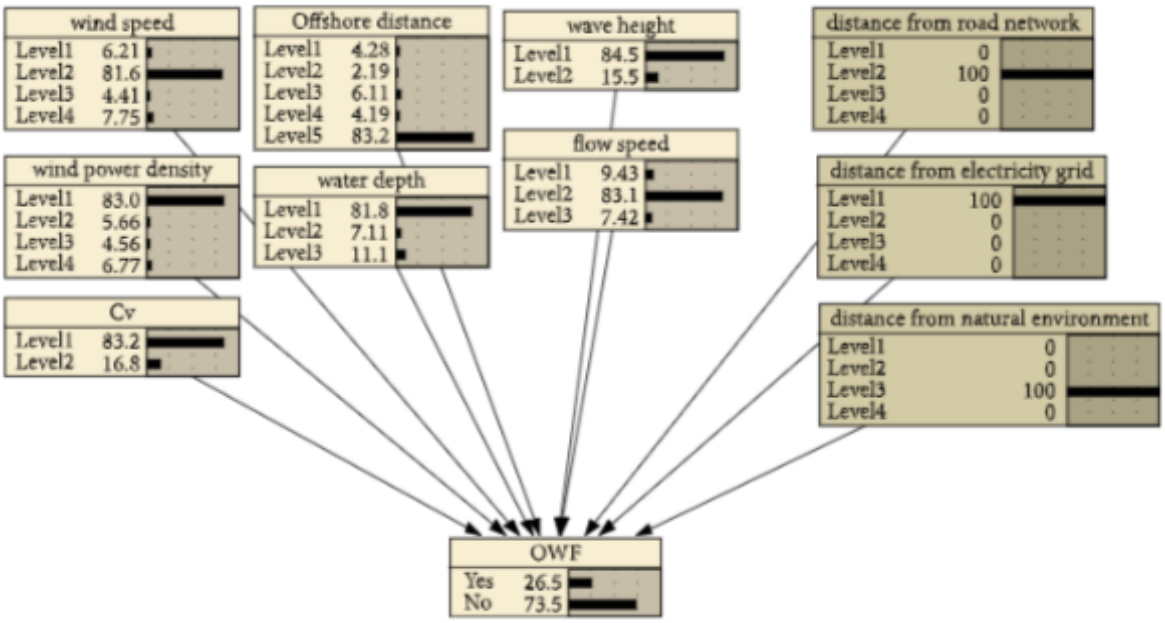


Figure 3: An example Bayesian Network for offshore wind farm siting in the North Sea. Probability distributions were constructed for each predictor (or "parent" node), with all arrows pointing to the likelihood of offshore wind farm site suitability (the "child" node). The quantitative data for each parent node were transformed into discrete levels. Source: Li et al. [65].

generated using long-term observational records [154] or Monte Carlo simulations [155]. BNs accommodate wind farm siting scenarios by modifying parent nodes and observing effects on the child node, such as Pınarbaşı et al. [156] observing an increased available area for European offshore wind energy development if likelihood of underwater cable construction is increased. Despite a BN's provided insight into how predictors influence each other, they are less commonly applied to WiFSS than (Non-)GIS-MCDA approaches.

Compared to other WiFSS modeling approaches, BNs have several unique advantages. Firstly, constructing probabilistic connections between predictors allows BNs to explicitly estimate risk in system behavior [157]. Indeed, BNs are frequently employed for environmental risk assessment, with Kaikkonen et al. [158, p.63] commenting that the probability distributions of each parent node "explicitly address uncertainty in different parts of the analyzed system",

therefore representing risk. In a WiFSS context, BNs may be used to compute risks due to interactions between environmental, social, and political predictors [159], such as potential backlash due to avian mortalities caused by a new wind farm project. BNs share Non-GIS-MCDA's advantage of using expert knowledge to construct predictor datasets, while simultaneously handling predictors based on quantitative data [160]. The intent of BNs is to depict each parent node as discrete possible states to be taken (as in Figure 3), hence data for each predictor are discretized to produce probabilities of each state's occurrence [161]. Datatypes for each predictor, whether expert knowledge or numerical observations, are thus treated equally and allow for a wide range of predictors to be included in BN-based WiFSS models. A final important advantage is BNs' ability to illustrate the causal structure between predictors [162]. More complex BNs consist of multiple levels of parent nodes, for instance Borunda et al.'s [64] BN for onshore wind energy development across Mexico uses "City" as its highest parent node, with wind speed and energy consumption nodes at lower levels. Causal structures give BNs the "capability to represent the conditional dependence between events" [163, p. 758], allowing the binary outcome of wind energy development to be explained based on probabilistic states of individual and groups of predictors.

Among BNs' limitations is the inconsistent method for discretizing the data of each parent node's probability distribution [157]. Continuous data may be discretized for inclusion in BNs by splitting data up into equal-sized quantiles or equal-length intervals [164]. Studies such as Flores et al. [165] and Beuzen et al. [166] found that the selected discretization method impacts a BN's predicted outcomes, because of how the method changes the shape of a node's probability distribution. Another limitation of BNs is that the assumed probability distribution for any node might not always be appropriate, again compromising the model's predicted outcomes. A BN's

developer must choose a distribution shape for each node (normal, uniform, multimodal, etc.) that reflects each predictor's true variability [167], a decision made more difficult with less prior knowledge (e.g., few wind speed observations) [168]. An inaccurate probability distribution means a greater risk of incorrect WiFSS assessment when running scenarios through the network. A BN with many nodes and connections can make quantitative validation methods difficult. According to Sperotto et al. [169], the various dataset sources and types used to create BNs reduces validation to single components of the network, or to qualitative approaches such as comparison against BNs in existing literature [170]. A final important limitation is the acyclic nature of BNs resulting in a lack of temporal and spatial explicitness [157]. Stritih et al. [171] state that even repeatedly running a BN at different points in space and time cannot account for interactions between processes at different scales. Integrating BNs with a GIS and producing a cyclic network can grant them spatiotemporal explicitness [172], with examples of this in existing WiFSS assessments [156]. By contrast, any model using Cellular Automata is inherently spatiotemporally explicit [173], meaning that WiFSS-LRCA can capture iterative updates to future commercial wind farm locations without additional model components.

2.1.4. Logistic Regression.

Logistic Regression models are useful for analyzing events that are "naturally or necessarily represented by binary variables" [174, p.67]. A common application of LR models is analyzing the probability of dichotomous land-use change, such as residential development [175], urban sprawl [176,177], and wind farm installation [58,59]. LR assumes that the logit (or log-odds) of a binary dependent variable is a function of the linear combination of a set of independent variables [93] (i.e., predictors of WiFSS). Using the logit means that discrete and continuous independent variables can be combined with limited error [60], hence published LR applications

to WiFSS do enlist discrete variables, such as political affiliation [62] and reported health conditions [178]. In the context of environmental modeling, LR is similar to Bayesian Networks in that both approaches probabilistically assess the occurrence of an event, such as land-use change. However, LR's applications also resemble GIS-MCDA in their frequent usage to construct a suitability surface at a single point in time, as shown in Figure 4, a raster constructed by Mann et al. [58] in their LR-based WiFSS assessment of Iowa. Existing LR-based WiFSS models often enlist relatively small predictor sets. Stevens et al. [179] assessed the impacts of a wind farm project on migratory birds in Texas based on bird populations and wind turbine proximity. Foley [59] applied a larger predictor set that enlisted wind power class, household income, elevation, and population density in a WiFSS study over Maine. Outside of the United States, LR approaches have focused on analyses of social acceptance of individual wind energy projects. Examples include surveyed opinions about existing wind farms in Italy's Apulia Region [180] and stakeholder opinions of potential offshore wind farms in Taiwan [181]. Regardless of application context, the overall LR approach in environmental contexts is to use a maximum likelihood method to estimate the coefficients of each independent variable [182], to then use this fitted equation to predict and/or estimate the environmental feature of interest, and then finally to validate the equation's performance based on prediction error rate [183].

One of LR's greatest advantages is its ability to create spatially explicit model outputs. Figure 4 shows one of many examples of LR being applied to gridded surfaces to construct suitability maps, others being Raja et al.'s [184] study of landslide susceptibility in Turkey and Abdel-Kader's [185] construction of soil maps in coastal Egypt. This easy translation into a spatially explicit context is a key advantage of LR over BN approaches to environmental modeling, with predictions by LR models also less likely to be compromised by relying on expert opinions to fill

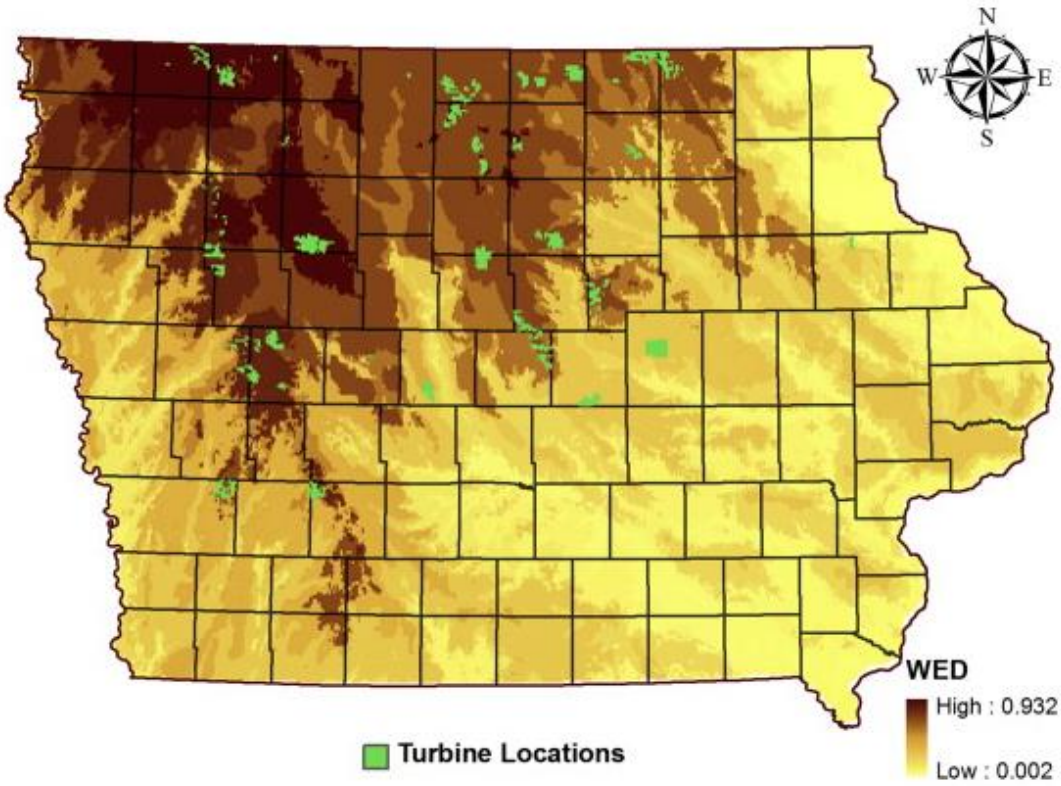


Figure 4: Example of a suitability surface constructed by applying a logistic regression equation to a raster surface over the State of Iowa. Each cell is colored based on the probability of commercial wind farm existence, and the locations of existing wind farms (as of 2012) are shown in green. Source: Mann et al. [58].

gaps in datasets [186]. However, much like BN-based models, LR can handle predictors based on discrete data [187], because the use of a link function (i.e., a logit) allows variables in an LR equation to be discrete or continuous and not require normal distributions [188]. Consequently, discrete data are often used to represent predictors in LR-based WiFSS models, such as Mueller and Brooks [189] including median age among predictors of wind energy's distributional justice in the United States, and Harper et al. [62] using elected officials to represent political climate

during times of wind farm construction across the United Kingdom. While GIS-MCDA can handle discrete datasets with a spatial component, such as transmission lines and major roads [45], applying non-spatial datasets often results in all pixels of a suitability surface being assigned the same value [190]. As such, unless proxying discrete data with a continuous variable, e.g., representing aesthetic pollution of offshore wind farms with distance to the shoreline [191], GIS-MCDA approaches to WiFSS assessment are less adept at depicting social and political predictors than LR equations. LR's compatibility with constructing suitability surfaces and its ability to handle discrete and continuous data were important reasons for the decision to construct an LRCA-based WiFSS model for this dissertation.

An LR equation that incorporates many predictors is vulnerable to overfitting. According to Merckx et al. [192, p.590], a model with too many degrees of freedom "will generally have poor predictive performance, as it can exaggerate minor fluctuations in the data." In the case of LR, too many predictors increase the number of possible interactions between them, making effects of these individual interactions on the dependent variable difficult to distinguish [193], particularly when the number of predictors exceeds the number of data points [194]. Another limitation of LR-based modeling approaches is that the basic equation structure cannot account for the scale of influence of each individual predictor. This limitation is sometimes circumvented with a multi-level LR structure that accounts for the dependence of a cluster of predictors on a single, higher-level predictor [93]. For example, Khan and Shaw [195] use a multi-level LR equation to predict rates of contraceptive use across Bangladesh by assessing survey responses at three levels of increasing spatial scale: individuals, villages, and administrative areas. A third limitation is that, like BN and MCDA approaches, LR equations are not temporally explicit, due to predictor data representing a single point in time. Temporal explicitness can be realized by

rerunning the LR equation given new data in subsequent timeframes, as in Kong et al.'s [196] study of fire detection using surveillance videos, though this requires continued refitting of the equation's coefficients. Another option is to integrate Cellular Automata that update grid cell states iteratively based on the trained relationships between the LR equation's predictors and each cell's transition suitability, a process used by Liao et al. [74] and in various LRCA simulations of land-use change [71,76,99]. Temporal explicitness by adding CA is what allows WiFSS-LRCA to project the specific future times at which grid cells across the CONUS gain commercial wind farms.

2.2. A Systematic Review of Predictor Standardization in WiFSS Modeling Studies.

Despite the common objectives of the above modeling approaches, i.e., to improve system understanding and to inform the decision-making process for siting wind farms [28], these approaches vary in their enlisted predictors. Rediske et al. [113] summarized that certain predictors are frequently enlisted in such studies, many of which describe physical features (e.g., wind speed, distance to roads/transmission lines/urban areas, land type, slope, etc.), though these predictors are not all applied to every study. Several reasons explain the exclusion of certain predictors, such as their effects on siting decisions being perceived as lower [197], covariance with other predictors [198], or if they are simply irrelevant (e.g., ocean depth does not impact onshore wind farm siting decisions). Moreover, non-physical predictors (such as project cost [199], government policies in effect [200], and demographics [201]) are harder to include in GIS-based approaches to WiFSS because of their need to be spatialized [61] for evaluation on a continuous domain. However, the use of expert opinions to rank candidate wind farm sites makes inclusion of non-physical predictors easier in Non-GIS approaches [51,52]. Predictor selection is also contextual; some WiFSS studies, particularly those using GIS-MCDA, are performed to

assess wind farm siting potential based only on features of the land itself [47,125]. Conversely, other studies situate their analysis within a broader social or economic context, such as wind farm project acceptability [202,203] or total project costs [30,204], thereby altering the enlisted predictors. In short, although many predictors can be considered relevant to wind farm siting decisions, some are incorporated more frequently into WiFSS models than others, and the same predictors may be represented in different ways by different studies.

The growth of WiFSS modeling, evidenced by recent review papers [40,43,113,114], highlights the need for model developers to prioritize communication and knowledge sharing and thus to ensure the continued refinement and policymaking benefits of these models [41]. Modeling collectives from other disciplines, such as the Coupled Model Intercomparison Project [205], encourage their participants to utilize a common set of variables and experiments in order to standardize climate model performance and facilitate comparisons of different models' outputs. Jakeman et al. [27] note that model building and usage are inherently subjective and benefit from standardization. There is consequent value in recommending predictor standards for current and future WiFSS model development, especially since predictor selection and representation impacts a model's suitability analysis (e.g., use of predictors as Boolean constraints or continuous evaluation criteria., see Section 2.1.1). Published review articles have covered the wind farm site selection process [113], application of GIS-MCDA (and Non-GIS-MCDA) to siting renewable energy projects [114], and wind energy development's social [206,207] and environmental [12,208] impacts. However, a systematic review of how existing WiFSS studies have selected and represented their predictors did not exist prior to this dissertation. The remainder of Section 2.2 and Section 2.3 are taken from Wimhurst et al.'s [40] review of this very topic, with the following paragraphs detailing the article selection and data collection

approaches. Sections 2.2.1 to 2.2.5 present results of a thematic synthesis of these articles. Section 2.3 provides recommendations for standardizing predictor selection and representation in future WiFSS studies and how this review informed the construction of WiFSS-LRCA.

This systematic literature review took a thematic synthesis approach that grouped findings into five themes [209], allowing for methodical article selection and thus reduction of author bias in the review process [210]. Articles were sought about WiFSS models that detail the selection and representation of their predictors, from which inconsistencies in how different studies selected and represented them could be deduced. Snowballing and a database search via Web of Science/Scopus were done to identify articles for this review over the period from March 2022 to May 2022. Articles published between January 2000 and May 2022 were sought for the database search, ensuring that this review consisted of contemporary articles up to the time of the search process. Existing reviews on related topics [43,113,114] took a similar approach when identifying potential articles. A modeling approach was not specified for the dataset search (e.g., MCDA, BN, LR), because contrasting predictor representation across modeling approaches was of interest to this review. Snowballing added to the database search by using the reference lists of existing review articles, thus expanding the list of obtained articles for this review [211]. Specifically, a "backward snowballing" approach [212] for articles that mentioned WiFSS modeling in their titles and/or abstracts was enlisted. The reference lists of two specific review articles were used: Rediske et al.'s [113] review of the wind farm site selection process and Shao et al.'s [114] review of MCDA applied to renewable energy site selection, both of which are relevant to the current topic and were published in high-impact journals.

A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Flow Diagram [213] illustrating the refinement of the collected articles is shown in Figure 5. The article search identified 206 articles in total, 109 of which came from the database search and 97 more from snowballing (54 articles from Rediske et al. [113] and 43 from Shao et al. [114]), 27 of which were duplicates identified by both approaches. The title, abstract, and keywords of the 179 non-duplicate articles were screened for references to wind energy or suitability analysis. This screening removed 39 articles that focused on other sources of energy (e.g., solar, tidal, geothermal) or on assessing the suitability of non-energy systems. A full-text assessment of the remaining 140 articles sought details about predictors for a suitability analysis and/or techniques associated with WiFSS modeling (e.g., MCDA). This full-text assessment removed 20 articles that did not specify their predictors, three articles lacking a full-text version (despite requests from their authors), and one article not written in English. The PRISMA approach left 116 articles eligible for inclusion in the systematic review.

A spreadsheet was created using the articles identified by this systematic review (see Appendix A2 for link to Supplementary Material), which compiled each article's predictors along with other relevant information under the column headings described below:

1) *Year of Publication*, for ensuring that methods of predictor selection and representation summarized in this review are contemporary. Of the 116 included articles, 66 of them (41 from the database search, 25 from snowballing) were published from 2018 to 2022.

2) *Country of Origin*, for illustrating case study contexts of these articles. These contexts are illustrated in Figure 6, showing a large number of studies from China (16), Turkey (15), Iran (13), Greece (9), the United States (7), Spain (5), and Saudi Arabia (4). Studies were

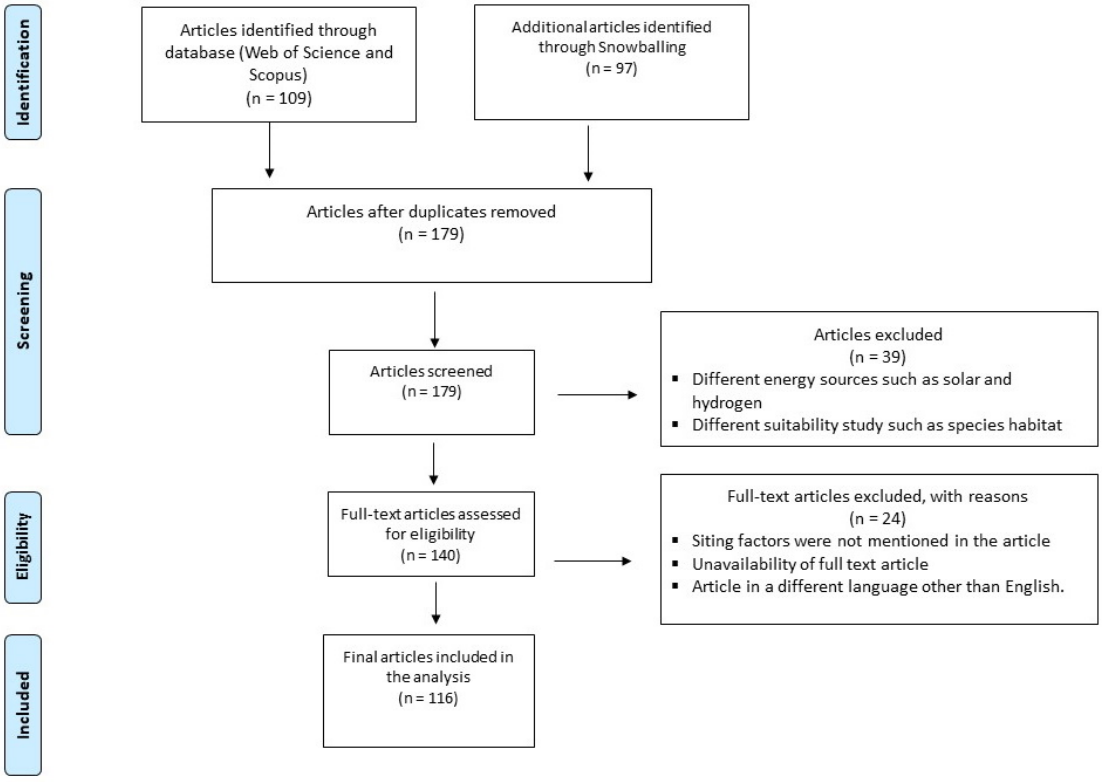


Figure 5: A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Flow Diagram that illustrates the method by which articles for this systematic review were identified, screened, and finalized. Source: Wimhurst et al. [40].

also conducted on multiple spatial scales, whether regional (city/county/state), national, or international; country and spatial scale could impact the treatment of predictors.

3) *Onshore or Offshore,* for documenting whether articles assessed onshore or offshore WiFSS. Offshore WiFSS studies accounted for 21 (18%) of the 116 articles included in this review (offshore study locations in Figure 6 (yellow points) are based on their approximate centroids). The number of offshore WiFSS studies is sufficiently large to allow for the assessment of onshore versus offshore predictor differences.

4) *Modeling Approach*, for summarizing the type of model and predictor ranking/weighting methods enlisted by each article. Table 1a shows that 98 (85%) of the 116 reviewed studies enlisted a (Non-)GIS-MCDA approach for assessing WiFSS, most of which were performed in a GIS environment (81 out of 98 studies) that utilized secondary datasets (e.g., digital elevation models, land cover rasters, census statistics, etc.), and 51 of which used AHP to construct a weighting scheme. Conversely, 18 (16%) studies did not use a (Non-)GIS-MCDA approach (Table 1b), with Data Envelopment Analysis (DEA) and other GIS-based models being frequent choices. The types of data collected for WiFSS models are often connected to the modeling approach, e.g., Non-GIS-MCDA relying on primary data collection to rank candidate wind farm sites (see Section 2.1.2). As such, modeling approach can affect the selection and representation of predictors.

5) *Basis for Predictor Selection*, for documenting how each reviewed article decided upon its predictors, with all studies relying on at least one of four methods. Eighty (69%) of the 116 studies used previous literature to justify their predictor choices, 38 (33%) studies relied on their authors' opinions about which predictors to include, 24 studies (21%) enlisted external expert opinions to inform these decisions, and 20 studies (17%) stated that predictor selection was influenced by knowledge of local geography and/or legislation.

Because this review was performed to examine predictor differences across WiFSS studies, several properties of all predictors enlisted in each reviewed article were documented in the Supplementary Material. These properties consisted of "File Type" (Vector, Raster, Point Observations, Unspecified), "Predictor Type" (Constraint, Evaluation, Unspecified), "Constraint Nature; Logic" (if the predictor was implemented as a constraint, how was it implemented, and

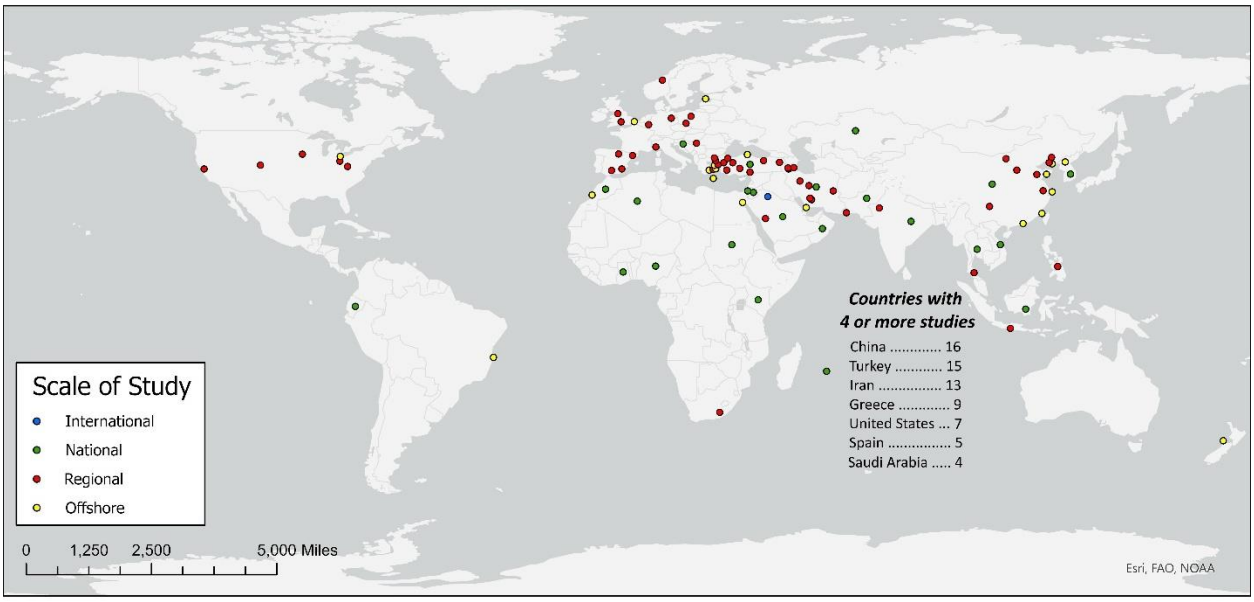


Figure 6: Study locations of the articles included in the systematic review. Points are colored by the spatial scale of the study performed. Note that some points overlap due to studies being performed over the same spatial domain. Basemap from Esri [214]. Source: Wimhurst et al. [40].

by what logic), "Data Source" (the primary or secondary data that supplied the predictor, e.g., expert opinions, a website source, legislation), "Classification" (Economic, Environmental, Social, Technical, etc., otherwise Not Classified), and "Combined Predictors" (describes grouped sub-predictors, and/or predictor names that represent the same concept across different studies). Each of these six properties underpins the thematic synthesis presented in the systematic review, and the discussion of each of the five themes is accompanied by figures and tables that provide bibliometrics for each of these properties [120]. For instance, the "Classification" property is critical to Section 2.2.2, with bibliometrics computed that quantify the commonest language (e.g., technical, economic, environmental, etc.) used to classify predictors across all reviewed articles. Absent information about predictor properties was documented in the Supplementary Material as "N/A", for example, an unspecified data source or file type for a particular predictor.

MCDA Approach	Study Context	Data Type (Secondary/ Primary)	Weighting/Ranking Method	References
GIS-based	Onshore	Secondary	AHP	[29,37,44,46,47,49,54,110,197,199,215,216, 217,218,219,220,221,222,223,224,225,226, 227,228,229,230,231,232,233,234, 235,236,237,238,239,240,241]
			ANP	[56,242,243]
			TOPSIS	[46,47,48,230,238]
			VIKOR	[47,48]
			OWA	[37,48,56,227,228,244,245]
			BWM	[47,55,246]
			Prescribed Weights	[24,45,247,248,249,250]
			No Scheme/Equal Weights	[38,115,251,252,253,254,255,256,257]
			Other	[111,139,258,259]
		Primary and Secondary	AHP-VIKOR	[119,124]
			TOPSIS	[260]
	Offshore	Secondary	AHP	[28,112,191,261,262,263,264,265,266]
			TOPSIS	[263]
			Prescribed	[42,267]
			No Scheme/Equal Weights	[268]
Non-GIS-based	Onshore	Secondary	AHP	[136]
		Primary	AHP	[51,130,137]
			VIKOR	[51]
			PROMETHEE	[131]
			ELECTRE	[135,146]
			TOPSIS	[52,132]
			Intuitionistic Fuzzy Logic	[269]
		Primary and Secondary	No Scheme/Equal Weights	[138]
	Offshore	Secondary	AHP	[133]
			TODIM	[39]
			No Scheme/Equal Weights	[270]
		Primary	BWM	[134]
			PROMETHEE	[139]
			ANP	[139,143]

Table 1a: Articles included in the systematic literature review, for MCDA approaches only. Acronyms for the listed MCDA methods: Analytic Hierarchy Process (AHP), Analytic Network Process (ANP), Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), Multicriteria Optimization and Compromise Solution (VIKOR, in Bosnian), Ordered Weighted Averaging (OWA), Best-Worst Method (BWM), ELimination Et Choice Translating REality (ELECTRE, in French), Preference Ranking Organization Method of Enrichment Evaluation (PROMETHEE), Multicriteria Interactive Decision Making (TODIM, in Portuguese). Source: Wimhurst et al. [40].

Model Approach	Study Context	Data Type (Secondary/Primary)	References
Artificial Neural Network	Onshore	Secondary	[271]
Benefit-Cost Analysis	Offshore	Secondary	[272]
Data Envelopment Analysis	Onshore	Secondary	[273,274,275,276]
		Primary and Secondary	[277,278]
GIS - Boolean Logic	Onshore	Secondary	[279]
GIS - Correlation Analysis	Onshore	Secondary	[280]
GIS - Least Cost Distance	Onshore	Secondary	[281]
Ideal Matter-Element Model	Onshore	Primary	[282]
Logistic Regression	Onshore	Secondary	[58]
Machine Learning	Onshore	Secondary	[57]
Maximum Entropy Model	Onshore	Secondary	[283]
Mixed Integer Linear Programming	Onshore	Secondary	[284]
Picture Fuzzy Modeling and TOPSIS	Offshore	Primary	[285]
Wind Atlas Analysis and Application Program	Onshore	Secondary	[286]

Table 1b: Same as Table 1a but for approaches that do not use MCDA only. Source: Wimhurst et al. [40].

2.2.1. Theme 1 – Deciding upon Predictors.

Predictor decisions are often motivated by the predictors used in prior WiFSS studies. Doing so ensures that current modeling studies do not exclude important predictors, while also facilitating model output comparisons for the same spatial contexts. Figure 7 shows that, among the 116 reviewed articles (95 onshore WiFSS studies in Figure 7a, 21 offshore WiFSS studies in Figure 7b), predictors that describe wind resources (e.g., *Wind Speed*, *Wind Power Density*; note that italicized predictors designate those identified by the systematic review), natural limitations (e.g., *Slope*, *Elevation*, *Ocean Depth*), and distance to land features (e.g., *Distance to Roads/Transmission Lines/Protected Areas/etc*.) are often selected, suggesting that WiFSS studies value consistent predictor choices. This consistency continues in how WiFSS studies detail their selected predictors. Eighty-five (73%) of the 116 reviewed articles summarized their

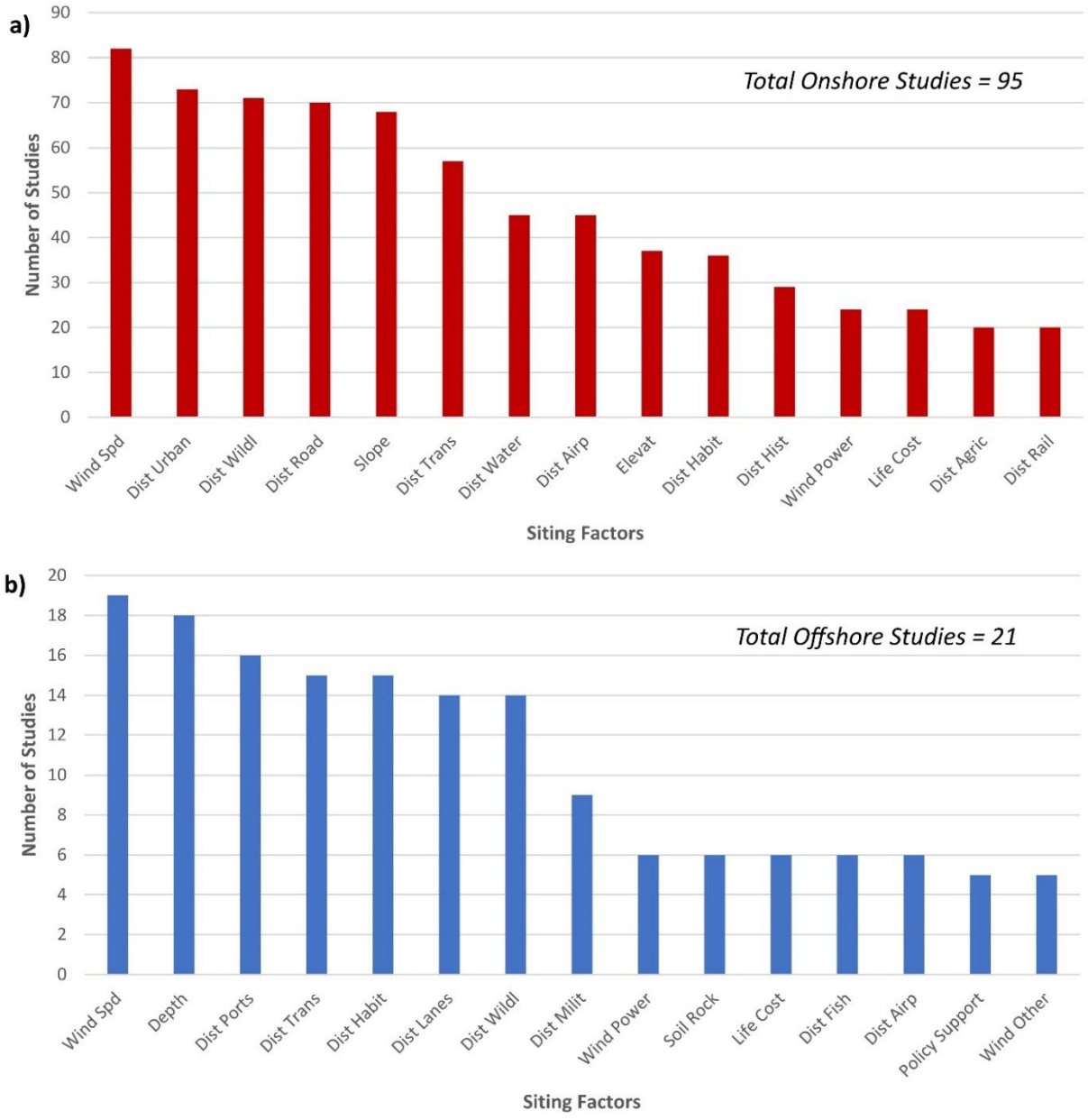


Figure 7: The 15 most commonly enlisted predictors in the onshore (7a, top) and offshore (7b, bottom) WiFSS (Wind Farm Site Suitability) studies included in this systematic review. Refer to Table 1 for full predictor names. Source: Wimhurst et al. [40].

chosen predictors in table form (see Supplementary Material for details), with table columns typically detailing each predictor's description [55,272], dataset source [191,253], citations [230,243], and implementation for constraint or evaluation [47,54]. Older WiFSS studies, such as Baban and Parry [115] and Rodman and Meentemeyer [125], along with recent, high-impact studies [24,202], are often cited to justify predictor choices, establishing a frequently emulated style of predictor selection and tabular presentation.

Figure 7 also demonstrates how some predictors are enlisted more consistently for WiFSS studies than others. Apart from *Life Cycle Cost* (24 onshore studies, six offshore studies) and *Policy Support* (five offshore studies), non-physical predictors are absent among those most frequently enlisted. Malczewski [61] refers to non-physical predictors as being "implicitly spatial," meaning they have potential to be expressed in a spatial context, which Non-GIS-MCDA models often realize by having experts rate the importance of predictors, and then re-expressing these ratings numerically [260]. However, this rating and ranking of candidate wind farm sites lacks common predictors across studies. Presumably important predictors such as *Elevation* [130,132] and *Distance to Airports* [135,270] are enlisted by few of the Non-GIS-MCDA studies in Table 1a, and some of these studies excluded *Wind Speed* [52,138,146]. Such studies sometimes establish a context of ranking wind farm sites based on social acceptability [138] or economic viability [39,130], hence Non-GIS-MCDA studies often select non-physical predictors relating to electricity demand [132,139]**,** local attitudes to wind farms [131,143], noise pollution [52,134], and many others. Some other Non-GIS-MCDA studies propose a general model framework without a case study [136,270], allowing selected predictors to remain generalized for other modelers to implement. Ultimately, the predictors used in Non-GIS-MCDA approaches to WiFSS depend on the study context.

Context dependence of predictor choices also applies to GIS-MCDA WiFSS studies. This context may simply be a study's objective, such as Gkeka-Serpetsidaki and Tsoutsos [191] examining social acceptance of offshore wind farm sites in Crete by using *Distance to Coastlines* as a proxy for noise pollution and visual disturbance. Similarly, Vinhoza and Shaeffer [265] assessed offshore wind's economic attractiveness in Brazil by using *Distance to Shipping Ports* to represent development costs. Predictor inclusion may also be justified by geographical context, such as Díaz-Cuevas [226] including proximity to tourist facilities due to South Spain's large tourism sector, or Ouammi et al. [252] not considering *Distance to Water Bodies* because those in Italy's Savona Province do not obstruct eligible wind farm sites. Pamučar et al. [55] and Sánchez-Lozano et al. [259] both noted that predictor choices depend on the geographical area in question. Study context also explains why *Wind Speed* was not included in some of the reviewed articles (101 out of 116, Figure 7). *Wind Speed* may be deemed invariable across a small study domain [251], a study's authors may instead use wind power density to represent the wind resource [244,276] or including *Wind Speed* may not assist the study's objective [114,138]. Predictor decisions therefore also depend on decisions made by a study's authors [197,217], not just geographical context and the chosen modeling approach.

Figure 7b shows that, of the 21 offshore studies in this review, many enlisted *Wind Speed* (19) *Ocean Depth* (18), *Distance to Shipping Ports* (16), *Distance to Transmission Lines* (15), and *Distance to Animal Habitats* (15) as predictors, suggesting some level of consensus about important predictors for offshore WiFSS studies. However, predictors that have documented relevance to offshore wind energy, such as *Natural Disaster Risk* [287,288] and *Distance to Commercial Fishing Areas* [289,290], were not as frequently used in these 21 studies, featuring in four [134,135,139,272] and six [42,133,262,264,265,268] studies, respectively. The exclusion

of these predictors could be due to a lack of relevant datasets or the relatively small amount of offshore WiFSS literature; the oldest offshore WiFSS study in this review was published in 2013 [272], compared to 2001 for onshore WiFSS studies [115]. The importance of prior WiFSS studies for deciding upon predictors will likely increase, as will the incorporation of underused predictors as the demand for offshore wind energy research grows.

2.2.2. Theme 2 – Classifying Data and Predictor Terminology.

WiFSS modeling studies often classify their predictors by grouping them under terms such as environmental [38,247,258], economic [37,261,283], social [219,224,282], and technical [49,110,218]**.** Classification allows for vocabulary control when describing predictors with similar effects on WiFSS. For example, the distance to the nearest city, transmission line, or road all present a common technical limitation to wind farm siting [110,226]. Additionally, environmental limitations are posed by the noise pollution and visual impact associated with wind turbines [217,244]; hence, these predictors are similarly classified together. A second benefit of classification in WiFSS studies is organizing one's analysis. Predictors classified as environmental or technical often serve as constraints, such as limited development in protected wildlife areas [202,254], land that is too elevated or steep [49,235], or areas with insufficient wind speeds [233,264]. Similarly, economic predictors like land leasing and maintenance costs are often incorporated into WiFSS models within a subset of equations that calculate cost competitiveness of candidate wind farm sites [263,271,284], hence their common classification.

Table 2 presents the classification terms most utilized across this systematic review. For both onshore (Table 2a) and offshore (Table 2b) studies, *Distance to Animal Habitats or Migration Routes* was most frequently classified as an environmental predictor [140,255,267]**,** as were other predictors that relate to natural land features, such as *Distance to Protected or Wildlife*

Predictors	Common Classifications (Frequency)	Studies Without Classification
Wind Speed	Technical (15); Economic (14); Wind/Weather (12); Climate (7); Environmental (5)	30
Distance to Protected or Wildlife Areas	Environmental (37); Social (3); Protective (2); Location (2)	28
Slope	Economic (10); Technical (8); Topography (7); Geographical (5); Environmental (3)	28
Distance to Urban Centers	Environmental (14); Social (14); Planning (6); Economic (5); Technical (3)	26
Distance to Roads	Economic (26); Environmental (7); Technical (7); Location (3); Planning (3)	24
Distance to Transmission Lines or Substations	Economic (17); Technical (5); Environmental (3); Infrastructural (3); Location (2)	24
Distance to Airports	Environmental (9); Economic (5); Protective (3); Location (2); Political (2)	20
Distance to Water Bodies	Environmental (23); Social (3); Economic (2); Location (2); Protective (2)	15
Elevation	Technical (6); Environmental (5); Topography (5); Economic (3); Geographic (2)	15
Distance to Animal Habitats or Migration Routes	Environmental (21); Social (3); Protective (2)	15
Distance to Agricultural Areas	Economic (2); Technical (2)	14
Distance to Historic Places	Environmental (10); Social (4); Cultural (2)	11
Distance to Railroads	Environmental (4); Economic (2)	11
Wind Power Density	Wind/Weather (8); Technical (4); Climate (2)	10
Life Cycle Costs	Economic (13); Technical (3); Social (2)	8

Table 2a: Language used to classify the 15 most common predictors in the onshore WiFSS studies included in this systematic review. The number of studies including each predictor that did not use classification is also given. See the "Classification" columns of the Supplementary Material (link in Appendix A2).
Source: Wimhurst et al. [40].

Areas [115,226,242] and *Distance to Water Bodies* [57,243,254]**.** This consistency was lacking for other predictors, with different studies classifying *Wind Speed* as a technical [110], economic [257], environmental [235], or climate [215] predictor. Similar inconsistency exists for predictors relating to distance from infrastructure. Tables 2a and 2b show studies commonly classifying

Predictors	Common Classifications (Frequency)	Studies Without Classification
Distance to Shipping Ports or Coastlines	Economic (5); Local Conditions (3); Technical (2); Region Characteristics (1); Social (1)	6
Distance to Shipping Lanes	Political (2); Protective (2); Social (2); Technical (2); Environmental (1)	6
Distance to Animal Habitats or Migration Routes	Environmental (10); Protective (1)	5
Distance to Protected or Wildlife Areas	Environmental (8); Protective (2)	5
Ocean Bathymetry	Economic (5); Technical (5); Construction (2); Sea State (2); Region Characteristics (2)	4
Distance to Transmission Lines or Substations	Local Conditions (4); Economic (3); Technical (2); Region Characteristics (1); Safety (1)	4
Distance to Military Zones	Protective (2); Environmental (1); Political (1); Region Characteristics (1); Safety (1)	4
Distance to Commercial Fishing Areas	Social (2)	4
Wind Speed	Technical (9); Wind/Weather (6); Economic (4); Geographical (1)	2
Wind Power Density	Wind/Weather (4)	2
Life Cycle Costs	Economic (4)	2
Distance to Airports	Local Conditions (1); Environmental (1); Safety (1); Technical (1)	2
Soil or Rock Type	Construction (1); Environmental (1); Region Characteristics (1); Sea State (1); Technical (1)	1
Other Wind Properties (Turbulence, Effective Wind Hours, Direction)	Wind/Weather (3); Environmental (1)	1
Policy Support	Social (3); Cultural (1); Economic (1); Safety (1)	0

Table 2b: Same as Table 2a but for the offshore WiFSS studies included in this systematic review. Source: Wimhurst et al. [40].

Distance to Airports [247] *Distance to Urban Centers* [254] and *Distance to Railroads* [226] as environmental predictors, whereas *Distance to Roads* [217], *Distance to Agricultural Areas* [235], and *Distance to Transmission Lines and Substations* [244] were often classified as economic predictors. Inconsistent classification is further complicated by the prescribed influence of predictors. For example, some studies [235,258] prescribe proximity to population centers as a social and economic asset (lower construction costs; closer to demand areas), but other studies [217,244] prescribe this proximity as a social and economic detriment due to

increased noise pollution and visual impact from new wind farms. The classification terms adopted for the predictors of WiFSS models therefore depend to an extent on the subjective decisions of model developers.

Another example of this subjectivity is the decision not to use a classification scheme. Table 2b shows that, of the 21 offshore WiFSS studies, six (29%) studies that included *Distance to Shipping Lanes* did not classify their predictors [42,135,262,264,267,285]. Of these six studies, Wu et al. [135] and Zhang et al. [285] do not use a GIS-based approach (of the 32 Non-GIS studies in this review, eight studies did not classify their predictors [58,136,275-278,284,286], and six others classified only some of them [51,133,134,143,270,274])**.** The other four of these six studies that included *Distance to Shipping Lanes* are GIS-based, despite Tercan et al. [264] stressing the importance of having technical, economic, environmental, and social criteria for evaluating potential offshore wind farm sites. An absent classification scheme sometimes appeared concurrently with other modeling decisions. Of the 32 studies in this review that did not classify *Wind Speed* (Table 2a and 2b), GIS-based studies that included an equation-based economic/technical analysis of potential wind farm locations often lacked a classification scheme [55,56,239], as did studies with a small number (fewer than seven) of predictors [220,250,279]. The application of classification schemes across WiFSS studies is inconsistent and not well-defined, thus making the intended role of predictors when comparing WiFSS study approaches potentially unclear.

This lack of clarity also comes from WiFSS studies using different terminology to describe the same predictors. *Wind Speed* was referred to by several different terms throughout the systematic review, such as “Wind Potential” [45], “Wind Sources” [39], “Average Wind Blow” [273], and “Efficiency” [51]. It can be implied that these terms describe the same predictor, but not for

certain, and this uncertainty increases if the means of data collection for each WiFSS study is different, e.g., a downloadable wind speed dataset versus expert opinion about wind speed's importance. Conversely, some studies used the same terminology to describe different predictors; the term "Protected Areas" was used to describe forests [57,225], bird habitats [29,272], marine habitats [28,267], or combinations of these features. Common language for both describing and classifying predictors is essential for any modeling discipline, such as climate modeling [205], especially given the number of recently published WiFSS studies (see Supplementary Material). WiFSS modeling would benefit from nomenclature for predictor terms and their classification, thereby assisting communication when using prior literature to inform studies and establishing a standard language for model developers to adopt [41].

2.2.3. Theme 3 – Implementing Predictors for Constraint or Evaluation.

The implicitly spatial nature of non-physical predictors [61] means that their depiction as continuous in space requires joining them to a gridded dataset (see Mann et al.'s [58] approach to census demographics, or Brewer et al.'s [127] to social attitude surveys), or being proxied with a physical predictor, such as inferring noise pollution or visual impact based on distance from infrastructure [248,281]. Either of these approaches allow non-physical predictors to be used alongside datasets that commonly represent physical predictors in GIS-based WiFSS models, such as line shapefiles of powerlines for transmission line proximity [231,237] or rasters of wind speed for assessing the resource itself [47,225]. The common function of the datasets for each predictor is to inform a GIS-based model's assessment of wind farm potential across a continuous spatial domain. Where GIS-based WiFSS studies frequently differ is in their implementation of the same predictors as constraints and/or evaluation criteria.

In this review's context, constraints are Boolean restrictions that eliminate potential wind farm locations based on a minimum standard [24], such as land being too steep [247,251], being too close to historic landmarks [191,230], among many others. These constraints are typically either a buffer distance around land features (e.g., no wind farms within 500 meters of a river [235]), or prohibition within an area of conflicting land-use (e.g., no wind farms in a designated vulnerable bird habitat [46]). Evaluation criteria assess WiFSS outside the constrained zones, either as an ordinal [229,250] or a quantitative [112,279] value. The two commonest types of evaluation criteria are those that assess suitability with distance from physical features (e.g., WiFSS being greater closer to roads [44]), and those based on magnitude at a singular point in space (e.g., WiFSS being greater in high-altitude locations up to 2000 meters [215]). Table 3 summarizes onshore (Table 3a) and offshore (Table 3b) physical predictors that were frequently enlisted as constraints and/or evaluation criteria by the studies in this review, with predictors representing distance to land features (e.g., *Distance to Urban Centers, Distance to Shipping Ports or Coastlines, Distance to Protected or Wildlife Areas*), *Slope*, *Wind Speed*, and *Ocean Depth* being especially common as constraints. Some of these studies used predictors to both constrain and evaluate WiFSS, such as Ajanaku et al. [110] excluding areas of West Virginia more than 10 kilometers from transmission lines (constraint), and having suitability increase with proximity to transmission lines outside of constrained zones (evaluation). Many predictors were used to both constrain and evaluate WiFSS across the studies in this review, particularly *Wind Speed* [220,236], *Distance to Roads* [216,231], and *Ocean Depth* [249,267], hence the high counts in both columns of Table 3 for these predictors.

Predictor	Frequency as Constraints (Unspecified Logic)	Frequency as Evaluation Criteria
Distance to Urban Centers	57 (15)	38
Distance to Protected or Wildlife Areas	52 (10)	23
Slope	51 (14)	42
Distance to Roads	49 (24)	51
Wind Speed	47 (18)	70
Distance to Airports	38 (14)	13
Distance to Water Bodies	37 (9)	12
Distance to Transmission Lines or Substations	30 (13)	42
Distance to Animal Habitats or Migration Zones	29 (9)	13
Distance to Historic Places	24 (7)	7
Elevation	21 (8)	20
Distance to Railroads	16 (4)	4
Distance to Agricultural Areas	9 (4)	14
Wind Power Density	7 (4)	18

Table 3a: Number of studies that employed predictors as constraints and/or evaluation criteria, among the most common predictors used in the onshore WiFSS studies included in this systematic review. The frequency of unspecified logic for constraint criteria for each predictor is also given in parentheses. See the "Predictor Type" and "Constraint Nature; Logic" columns of the Supplementary Material (link in Appendix A2). Source: Wimhurst et al. [40].

Despite these similarities in the implementation of constraints and evaluation criteria across WiFSS studies, this review highlighted some important differences:

1. *Specifying logic for the selection and setting of constraints.* Most studies in this review utilized existing legislation (e.g., laws prohibiting wind energy development in protected areas [262]), previous WiFSS modeling studies (e.g., setting a maximum land slope based on a prior study [246]), and/or a chain of reasoning in the main text (e.g., setting a buffer around airports to mitigate radar signal interference [111]) to justify the selection and setting of constraints. Some studies, however, did not provide logic for their models'

Predictor	Frequency as Constraints (Unspecified Logic)	Frequency as Evaluation Criteria
Ocean Depth	14 (1)	13
Distance to Shipping Ports or Coastlines	12 (1)	10
Wind Speed	11 (0)	13
Distance to Animal Habitats or Migration Routes	11 (0)	2
Distance to Protected or Wildlife Areas	11 (1)	2
Distance to Shipping Lanes	9 (0)	5
Distance to Military Zones	9 (0)	1
Distance to Transmission Lines or Substations	9 (2)	7
Distance to Airports	5 (0)	2
Distance to Commercial Fishing Areas	5 (0)	1
Soil or Rock Type	4 (1)	3
Wind Power Density	2 (2)	5
Other Wind Properties (Turbulence, Effective Wind Hours, Direction)	1 (1)	4

Table 3b: Same as Table 3a but for the offshore WiFSS studies included in this systematic review. Source: Wimhurst et al. [40].

constraints. Table 3a shows that, of the 49 studies that implemented *Distance to Roads* as a constraint, 24 (49%) did not justify this constraint in any of the manners mentioned previously. Unspecified logic was, conversely, significantly less common for offshore predictors (Table 3b). Not justifying constraints leaves readers to guess whether a constraint is appropriate, and furthermore whether the constraint is transferable to other contexts. For instance, a 300-meter buffer distance around railroads might be acceptable for a WiFSS study in Northwest Iran [254], but whether 300 meters would be acceptable for studies in other locations is uncertain due to absent logic.

2. *Inconsistent implementation of constraints.* Table 3 suggests a common set of constraints employed by WiFSS studies, such as, for example, a minimum wind speed [256,268], limiting wind farm construction in protected areas [54,125], and a minimum distance from urban centers [251,283]. However, the magnitude of these common constraints varies widely. Pamučar et al. [55] enlisted a maximum land slope of 7%, in contrast to Tegou et al.'s [218] constraint of 30% (some studies instead constrained land slope with a degree angle [197,215], adding further inconsistency). Additionally, for distance-based constraints, some studies enlisted a prohibition rather than a buffer. Whereas Cradden et al. [267] only prohibited wind energy development in Europe's Important Bird Areas (IBAs), Ayodele et al. [216] also included a 300-meter buffer around Nigeria's IBAs, despite both studies using the same dataset [291]. A third facet of inconsistent implementation is the exclusion of important constraints. Compared to predictors listed in Table 3, few studies in this review enlisted constraints for *Distance to Mines or Pits* [24,237], or *Distance to Fault Lines* [38,257], despite the known risks of building wind farms in earthquake-prone areas [292,293] and over mines [294]. The decision to implement specific constraints in WiFSS models is sometimes context-dependent (e.g., there is no need to include fault line proximity if the study area does not experience earthquakes), but constraints having a consistent magnitude, units, and nature (prohibition or buffer distance) across studies is nevertheless important.

3. *Enlisting predictors as constraints or evaluation criteria in different studies.* While some studies enlist predictors to both constrain and evaluate WiFSS, as previously discussed, other studies may enlist a predictor only for constraint or only for evaluation. For instance, Mekonnen and Gorsevski [42] and Kazak et al. [248] set WiFSS to increase

with distance away from bird habitats, with no specified minimum distance or similar constraint. Conversely, Değirmenci et al. [225], Genç et al. [268] and Ouammi et al. [252] implemented proximity to bird habitats and migration routes strictly as a constraint. An increasing suitability with distance from a bird habitat (i.e., evaluation) would produce a different model output than just buffering the same habitat (i.e., a constraint), resulting in two different WiFSS outputs for the same spatial context. This difference presents a planning risk, knowing the negative impacts of improper wind farm siting on avian species [32,33]. The decision to implement predictors for either constraint or evaluation may be motivated by usage of these predictors in prior studies, and/or the modelers' objectives, for example, the evaluation of the wind resource in remaining locations after applying all other predictors as constraints [38,256,284].

The implementation of predictors as constraints or evaluation criteria can be subjective, again depending on a WiFSS study's context and individual modeler preference. Addressing this subjectivity could benefit the consistency of GIS-based WiFSS studies by normalizing the use of literature and legislation to inform the magnitude of constraints, thus explicitly justifying the use of specific predictors for constraint and/or evaluation. This implementation can also depend on regulations observed in a particular country or region. For example, some counties in the United States enforce setback distance constraints on wind energy development, other counties do not, and constraints in some cases are enforced at higher levels of government [295].

2.2.4. Theme 4 – Utilizing Primary and Secondary Data.

Tables 1a and 1b show that most studies in this review utilized secondary datasets, particularly studies that built GIS-MCDA models, usually in the form of downloaded geospatial data [219,238] and previously recorded observations [240,277]. Some secondary datasets were

enlisted by multiple studies, such as road information obtained from OpenStreetMap [225,231,237], Digital Elevation Models from the United States Geological Survey [125,216,228], and wind speed information from the Global Wind Atlas [49,236,241]. The use of such datasets for WiFSS studies exemplifies the value of free resource access for public sector model development [296], because developers are thereby encouraged to use a common set of predictors, facilitating standard language and comparisons of modeling approaches that are less biased by predictor choices. Differences in model outputs for the same geographical contexts could indeed be partially attributed to their enlisted secondary datasets. For instance, WiFSS studies from Turkey have represented protected areas with secondary datasets enlisted data either from the state government [229,232] or from larger organizations such as the European Environment Agency [283] and the United Nations Educational, Scientific and Cultural Organization [225]. Each data source has its own unique definition of protected areas, which combined with modeler preferences results in quite different depictions of protected areas for the same country. Selected secondary datasets therefore have important consequences for the consistency and comparability of model outputs across (GIS-based) WiFSS studies.

By contrast, primary data were enlisted almost entirely by the Non-GIS-MCDA studies in this review (Table 1a). These primary data are usually the collected opinions of academic or industrial experts, whether from questionnaire responses [137,269], conducted interviews [52,138], or focus groups [39,132], with the objective of assessing discrete wind farm sites based on the rated importance of a set of predictors. These ratings are usually ordinal and employ either a linguistic scale to express each predictor's individual importance [51,135,143] or a ranking of predictors relative to each other [131,134,231]. There are multiple examples in this review of WiFSS studies that relied solely on primary data for both physical and non-physical predictors

(e.g., studies that enlisted both *Wind Speed/Power Density* and *Life Cycle Cost* [134,143,269])**,** because they are both collected using opinion-based methods. However, the predictors incorporated into studies that rely on primary data are not consistent. Aras et al. [137] and Gamboa and Munda [138] excluded *Wind Speed* due to their interest in the technical and social feasibility of wind farm sites, respectively, and important predictors like *Distance to Military Zones* [134,135] and *Distance to Water Bodies* [132] were rarely included. Differences in predictor choices could result in inconsistent wind farm site characterization, which can be a problem when comparing Non-GIS-MCDA studies with the same geographical contexts, such as those from China [39,139,146] or Turkey [51,132,134]. The collection and application of primary data in these studies varied in other important ways, such as some studies using outside expert opinions, rather than those of the authors alone, to help decide upon predictors [130,134,269]. Additionally, most studies applied fuzzy logic to both quantify expert opinions and address the uncertainty inherent to linguistic decision-making [132], though some studies did not [131,137,146]. Much like with secondary data, there are several conventions in the use of primary data to represent predictors in WiFSS models, such as the use of linguistic scales and enlisting expert opinions, though these conventions are not universal.

A small number of studies in this review presented models that combined primary and secondary data in their assessment of WiFSS [119,124,138,260,277,278], eliminating the need for proxies or dataset transformations. For instance, Rezaei-Shouroki et al. [277] combined secondary wind speed observations and primary opinions about land price into the same set of Data Envelopment Analysis equations. Studies that enlist primary and secondary data also often take different approaches to WiFSS modeling, such as Xu et al. [124], which used datasets of bird migration routes and power plant locations to constrain suitable wind farm sites, and subsequently

evaluated a grid of remaining potential wind farm sites using expert opinions about a host of predictors. Li et al. [119] took a similar GIS-based approach but also assessed future wind resources under climate change. These efforts represent possible new directions for assessing WiFSS, but the lack of a standard approach makes predictor selection and representation highly variable. For instance, Gamboa and Munda [138] accounted for candidate wind farm sites' visual pollution with secondary simulations of viewshed (in square kilometers), in contrast to the more common method of assessing visual impact using the primary opinions of experts [51,52,143]. Another example is Pambudi and Nananukul's [278] decision to collect wind speed data using questionnaire responses. Studies that combine data collection methods show that WiFSS is not confined to representing predictors with only primary or only secondary data, but also that a common standard for their combination would benefit comparisons between study approaches.

2.2.5. Theme 5 – Data Source and Accessibility.

Beyond specifying data types (i.e., primary or secondary), it is also important for WiFSS studies to specify data sources and to ensure their accessibility [113,114]. Specifying data sources is important for several reasons, firstly doing so enables the replication of similar modeling studies and findings in the same or different spatial contexts [217,221]. Replication allows scientists and other modelers seeking to conduct similar research to better understand how and where to source candidate data sources for their respective studies, thus facilitating practices of knowledge transfer and data sharing for the public [41]. Secondly, action based on the results of WiFSS modeling studies requires their acceptance by scientists and decision makers. This acceptance is more likely if data sources and details regarding their accessibility are specified, thereby creating transparency in the research process and allowing other modelers and the public to trust a study's findings more readily [138]. Lastly, it is important to give credit to the producers or hosts of all

enlisted data sources. Citation is the primary means of demonstrating how credit should be given to existing studies and their data sources, while also helping other scientists and modelers locate data sources for their own research [52,114].

This review found that although many studies indeed specified data sources for their enlisted predictors [48,232,233], there were also studies that did not [38,273,274]. Tables 4a and 4b show the number of onshore and offshore WiFSS studies, respectively, that did and did not specify data sources for common predictors for the countries with the most studies in this review (see Figure 6). The tables suggest that not specifying data sources was more common for onshore WiFSS studies, particularly those from Greece, Iran, and Turkey, with datasets for *Distance to Airports/Protected or Wildlife Areas/Roads/Urban Centers* being the most frequently unspecified. There exist a few reasons why WiFSS studies may not specify their data sources, first among which is that the datasets used in these studies may be proprietary, as was often the case for studies that incorporated military zones [221,285], bird migration habitats [244,246] and protected areas [138,254] into their predictor choices. This non-disclosure of data sources can also be for legal and/or regulatory reasons [233]. Secondly, discussions about the importance of sharing datasets for environmental model development started relatively recently; the practice of dataset sharing is crucial for asserting any modeling practice as its own discipline [41]. The importance of providing citations for secondary datasets is particularly salient for predictors enlisted as constraints (see Section 2.2.3). Specifying these datasets means that studies performed in the same geographical context could brand the same locations as being (un)suitable for wind energy development, allowing for focused refinement of the predictors used for evaluation and of the models themselves [219]. Specifying dataset sources for predictors is also important for studies that employ primary datasets. Some studies in this review did not clarify

	Countries (Number of Onshore Studies)						
Predictor	**China (10)**	**Turkey (13)**	**Iran (12)**	**Greece (6)**	**United States (6)**	**Spain (5)**	**Saudi Arabia (4)**
Distance to Agricultural Areas	1(0)	1(0)	2(0)	3(1)	2(0)	1(1)	-
Distance to Airports	1(0)	2(5)	2(2)	2(4)	1(2)	2(2)	1(2)
Distance to Animal Habitats or Migration Zones	3(1)	3(4)	-	3(2)	2(0)	1(2)	1(1)
Distance to Historic Places	0(1)	1(2)	1(2)	3(3)	-	2(1)	1(0)
Distance to Protected or Wildlife Areas	3(1)	6(5)	5(4)	4(2)	4(1)	2(3)	1(2)
Distance to Railroads	1(0)	1(2)	0(2)	-	2(0)	2(1)	-
Distance to Roads	3(2)	4(4)	4(3)	3(3)	2(3)	2(1)	3(1)
Distance to Transmission Lines or Substations	4(1)	5(4)	3(4)	2(/2)	2(1)	2(1)	2(0)
Distance to Urban Centers	4(2)	5(3)	6(4)	3(3)	4(2)	2(2)	1(2)
Distance to Water Bodies	2(1)	4(3)	4(2)	1(2)	3(1)	2(1)	0(2)
Elevation	3(1)	5(1)	5(2)	-	1(1)	1(0)	-
Life Cycle Cost	4(0)	4(1)	2(4)	0(1)	-	1(1)	1(0)
Slope	5(1)	6(3)	7(4)	4(1)	2(1)	3(1)	0(2)
Wind Power Density	4(1)	2(5)	4(0)	-	1(0)	-	1(0)
Wind Speed	6(1)	8(2)	7(4)	4(2)	5(1)	2(2)	4(0)

Table 4a: Countries with four or more studies (see Figure 6) that specified data sources for the 15 most common predictors in this systematic review (see Figure 7) for onshore WiFSS studies. Each cell contains the number of studies that did and did not specify their data sources, the latter in parentheses. See the "Data Source" columns of the Supplementary Material (link in Appendix A2).
Source: Wimhurst et al. [40].

how they obtained expert opinions about predictors, nor did they provide the questions that were posed in the questionnaires or interviews conducted with them [119,278]. Although these data would be useful, their non-disclosure could be for ethical reasons such as protection governed by institutional review boards [266,267].

	Countries (Number of Offshore Studies)						
Predictor	**China (6)**	**Turkey (2)**	**Iran (1)**	**Greece (3)**	**United States (1)**	**Spain (0)**	**Saudi Arabia (0)**
Distance to Airports	1(1)	1(0)	-	1(2)	-	-	-
Distance to Animal Habitats or Migration Routes	3(1)	2(0)	1(0)	2(0)	1(0)	-	-
Distance to Commercial Fishing Areas	-	1(0)	-	-	1(0)	-	-
Distance to Military Zones	-	2(0)	-	3(0)	-	-	-
Distance to Protected or Wildlife Areas	3(0)	1(0)	1(0)	3(0)	-	-	-
Distance to Shipping Lanes	2(1)	2(0)	1(0)	3(0)	1(0)	-	-
Distance to Shipping Ports or Coastlines	2(1)	1(0)	1(0)	3(0)	1(0)	-	-
Distance to Transmission Lines or Substations	1(2)	2(0)	1(0)	3(0)	1(0)	-	-
Life Cycle Cost	2(1)	1(0)	1(0)	1(0)	-	-	-
Ocean Depth	3(1)	2(0)	1(0)	3(0)	-	-	-
Other Wind Properties	-	-	-	-	-	-	-
Policy Support	0(2)	2(0)	1(0)	-	-	-	-
Soil or Rock Type	1(0)	1(0)	-	1(0)	-	-	-
Wind Power Density	1(2)	-	-	-	-	-	-
Wind Speed	2(4)	2(0)	1(0)	3(0)	-	-	-

Table 4b: Same as Table 4a but for the offshore WiFSS studies included in this systematic review. Source: Wimhurst et al. [40].

Beyond specifying data sources, there is also the issue of data accessibility for predictors enlisted in WiFSS models. Although most studies in this review cited their datasets [216,236,286], others only listed the names of the institutes who provided their datasets in the Methods section [215,218], and some did not list their datasets at all [38,248]. By not providing full citations with functional links to enlisted datasets, the nature of the data that informed each predictor becomes difficult to ascertain. There was also some inconsistency in how these studies reported their dataset sources. While some studies provided data source details in the main text [233,248], other

studies provided details in an Acknowledgements section [282], and others only in their lists of references [48,265]. Without a recognizable, consistent way of identifying dataset sources, it becomes harder for readers to identify what datasets informed each predictor, as well as the data preparation that would have been necessary to incorporate them into a WiFSS model. This review also highlighted an issue with incomplete citation for selected datasets. Even among studies that did provide citations for their dataset sources, these citations were sometimes only the name of the government website, piece of legislation, or research institute that provided the data, without specific details regarding the dataset source [222,228]. When links to data sources were included, they were sometimes inaccessible, even in recent WiFSS studies [56,229]. Greater emphasis should be placed on ensuring that data sources for predictors are fully sourced in a consistent manner across WiFSS studies, in order to encourage the use of publicly available, high-quality datasets and to standardize their citation and presentation in published work.

2.3. Recommendations Based on the Systematic Review.

By performing a systematic literature review on 116 identified articles, it was found that predictor choices in WiFSS models and the reasons for adopting common standards for them were best discussed under the headings of five themes (see Sections 2.2.1 to 2.2.5). The summary of each theme below includes recommendations for standardizing predictor selection and representation in future work. This section concludes with remarks about how WiFSS-LRCA's construction was informed by this systematic review:

1. *Deciding upon Predictors*. WiFSS studies frequently justify predictors to be included in their models based on those used in prior work. Their usage in previous studies, and reasoning by model developers, has resulted in some predictors, such as *Wind Speed, Distance to Roads*, and *Elevation*, being used more than others. The reasoning employed by modelers to select

certain predictors over others may be motivated by overarching study context (e.g., assessing social acceptance of wind energy development), the geographical context, the perceived importance of predictors, or whether the study is one of onshore or offshore WiFSS. The decision to include or exclude predictors in a WiFSS model is sometimes obvious, such as only certain study areas being vulnerable to earthquakes [272], but not explaining predictor selections can be detrimental when comparing different studies with similar geographical contexts. For instance, Deveci et al. [297] included water depth in their study of offshore WiFSS in New Jersey to represent the required foundation structures and costs, but Mekonnen and Gorsevski's [42] participatory GIS model of offshore WiFSS in Lake Erie excluded water depth with no justification. Given the United States Federal government's interest in expanding offshore wind energy in the coming years [298,299], and the limitations water depth places on offshore wind farms [300], model developers could increase their impact on decision-making by using consistent predictor sets that raise fewer questions for non-modelers. If a decision is made to exclude predictors frequently used in prior work, then standard practice should be for WiFSS model developers to justify that decision explicitly.

2. *Classifying Data and Predictor Terminology*. While some WiFSS studies group their predictors under the subheadings of broader classification terms (e.g., economic, environmental, social, technical, etc.), others do not, or they may use different terms that describe the same sorts of classification. The advantage of classifying is clarification of the role that predictors play in a given WiFSS model. For example, *Distance to Roads* and *Distance to Urban Centers* represent "technical" aspects of wind energy development. Classification is especially useful when comparing WiFSS studies that utilized the same predictors. The terminology used to describe the same predictors (e.g., "Wind Speed" versus

"Wind Potential") is a second example of the importance of uniform vocabulary when conducting WiFSS studies. In their review of modeling water flows, Refsgaard and Henriksen [301, p.76] conclude that adopting standard terminology is important for "bridg[ing] the gap between scientific philosophy and pragmatic modelling". The importance of language choices that have an agreed-upon definition has also been recommended by dietetic [302], ecological [303], and behavioral [304] model developers. Proposing common definitions and vocabulary for the predictors used in different WiFSS models, and enlisting classification terms that group predictors with related effects on WiFSS, should thus be a priority as this modeling discipline expands. Based on the findings of this review, common classifications such as economic, environmental, social, and technical should be made commonplace and be given refined definitions, and uncommon names for predictors (e.g., "Average Wind Blow" rather than "Wind Speed" [273]) should be avoided in future work.

3. *Implementing Predictors for Constraint or Evaluation.* GIS-based approaches to WiFSS tend to incorporate predictors as Boolean restrictions on potential wind farm locations (constraints) and/or as quantified suitability with distance or magnitude (evaluation criteria). While some predictors are used frequently as constraints (e.g., *Wind Speed, Slope,* and *Ocean Depth*), some studies implement the same predictors as evaluation criteria instead, or in addition to being constraints. There is also the concern that logic for the selection and setting of constraints is left unspecified in some studies, meaning that readers must guess whether, for instance, a minimum *Distance to Protected Areas* of 500 meters is appropriate for other study contexts. Not addressing the subjectivity in setting constraints and evaluation criteria can cause serious wind farm planning risks. For example, the five WiFSS studies from Spain in this review all enlisted a *Distance to Protected Areas* constraint; one study did not specify

its constraint [138], three studies prohibited development within protected areas [37,46,259], and the fifth study also applied a 1000-meter buffer distance to these areas [226]. Rodríguez-Rodríguez et al. [305] found that Spain's protected areas are more vulnerable to any form of land development when not surrounded by a buffer distance. For that reason, using a WiFSS study to guide wind energy development in Spain that does not adequately protect vulnerable flora and fauna could cause inadvertent environmental damage, such as heightened avian mortalities [32,33]. It is thus recommended that future WiFSS studies apply predictors for constraint or evaluation in a manner consistent with existing literature, government policy, and expert opinions.

4. *Utilizing Primary and Secondary Data*. There are some commonly used datasets for WiFSS studies that represent predictors with secondary data, such as the Global Wind Atlas for *Wind Speed* [306] and the United States Geological Survey for *Slope* and *Elevation* [79]. However, the familiarity of model developers with certain, often more localized, datasets can result in quite different depictions of WiFSS due to using different datasets, even within the same geographical contexts. Inconsistencies also exist in the representation of predictors with primary data, such as the decision to involve experts in the predictor selection process, whether to use a linguistic scale to capture experts' opinions, and the frequent differences in enlisted predictors. This review found that the decision to use primary and/or secondary data in WiFSS studies often depended on the research question being answered. Studies that focused on developing a model that could construct a continuous WiFSS surface were more likely to rely solely on secondary data, such Mann et al.'s [58] logistic regression-based approach to assessing WiFSS in Iowa. By contrast, primary data were most often incorporated in WiFSS studies that ranked candidate wind farm sites for their development

potential based on expert opinions, as in Deveci et al.'s [307] assessment of potential offshore wind farm locations in Norway. The innovative methods of integrating primary and secondary data presented by this review were of particular interest. Standout examples include Rezaei-Shouroki et al.'s [277] use of expert opinions to construct a land cost predictor alongside secondary sources for population and wind speed, and Xu et al. [124] using secondary spatial data layers to constrain potential sites and then using expert opinions to rank the remaining sites. WiFSS approaches that combine primary and secondary data sources should continue to be pursued, using these existing studies as a basis for standardizing how their predictors are represented.

5. *Data Source and Accessibility.* Facilitating common standards for predictor selection in WiFSS studies depends highly on providing complete and functional citations for enlisted datasets. This review showed, however, that some WiFSS studies do not provide any citation for enlisted secondary datasets; some only specify the dataset provider in-text without a full citation, and sometimes citations lack specific details on how to access the data sources used in these studies. Some WiFSS studies that represented predictors with primary data also did not fully detail their data sources, namely how expert opinions about candidate wind farm sites were collected or the questions that were asked of said experts. At a time when scientific integrity is being questioned by the public and elected officials [102], there is an increasing onus upon model developers to ensure that their work is fully transparent and accessible, such as by making effective use of free online repositories [308] and preparing robust model documentation [309]. This motion toward transparency in part requires documenting how predictors were selected, and how and where each predictor's dataset(s) can be found in a way that both modelers and non-modelers can recognize across separate studies. While

proprietary knowledge and institutional review boards often limit the extent to which data for predictors can be shared, a standard presentation of dataset details in-text, and providing complete citations, would allow modelers to share knowledge of robust datasets more easily, while also garnering trust in their work.

The intent of this systematic review is not to dictate how predictors in WiFSS models should be selected and represented moving forward. Rather, this review serves to identify that an explicit common standard for predictor selection and representation does not yet exist, and that a set of standards for predictors could be adopted by at least recognizing the role played by subjective modeling decisions (e.g., approach to citation, deciding upon important predictors, predictor vocabulary). As such, the construction of WiFSS-LRCA takes cues from each of the five themes presented by the systematic review. Firstly, WiFSS-LRCA uses a comprehensive set of 47 predictors that existing literature have all been shown to influence commercial wind farm siting decisions, with explicit reasoning given for each predictor's inclusion. Secondly, common classification and predictor terminology found throughout the systematic review (e.g., Tables 2 and 3) were applied to WiFSS-LRCA, making this model's vocabulary consistent with previous work. Thirdly, the constraints most frequently used in prior WiFSS studies are used to restrict wind energy development in WiFSS-LRCA's grid cells, with citations given to justify the buffer distance/prohibition applied to each predictor acting as a constraint. Fourthly, while primary data would be useful to directly capture predictors such as social attitudes toward wind energy, the CONUS-wide scale of WiFSS-LRCA's application makes using secondary datasets the more realistic option. Finally, all secondary datasets in WiFSS-LRCA are summarized in the main text, fully cited in the References list, and compiled in a GitHub repository (see Appendix A1), thus clarifying how the data for each predictor were enlisted and where these data can be accessed by

others. Future WiFSS studies will ideally use the five themes presented by this systematic review to similarly justify their methods of predictor selection and representation, hopefully converging toward a standard approach throughout subsequent work.

Chapter 3: Data and Methods – Building WiFSS-LRCA.

3.1. Deciding upon Predictors and Datasets.

3.1.1. Dataset Selection.

Following the systematic review detailed in Chapter 2 [40] and based on predictors deemed relevant to WiFSS by previous work [113,114], the predictors for which datasets were sought for WiFSS-LRCA are presented in Tables 5a and 5b. The tables' columns detail the classification, predictor name (plus each predictor's codename as it appears when running WiFSS-LRCA), year of data preparation, datatype (and units), spatial scale (state or CONUS) and data source, with all datasets being publicly available. The classification and predictor terminology used follow from the systematic review, with the classifications grouping predictors under subheadings with related effects on wind farm siting decisions. These classifications consist of human and natural *environmental* conditions [283], *technical* limitations caused by existing/absent infrastructure and energy sector maturity [52], *economic* viability of wind energy development [286], conduciveness of the *political* landscape for wind energy sector growth [310], and *social* indicators that may be associated with attitudes toward wind energy [138]. Table 5a contains predictors classified for WiFSS-LRCA as either environmental or technical, with Table 5b containing the model's economic, political, and social predictors. Many of the predictors in Table 5b were not readily available as spatial datasets; their tabular data were joined to shapefiles of county [311] and state [312] boundaries for model runs over individual states and the CONUS, respectively. Annual data from each dataset were averaged over the years 2015 to 2019, assuming an installation period for an onshore commercial wind farm in the United States of five years [340]. Additionally, some predictor datasets possess resolutions too coarse to train WiFSS-

Classification	Predictor	Codename	Year	Datatype (unit)	State	CONUS	Source
Pre-established Land Use *(Environmental)*	Critical Habitats	*Critical*	2023	Shapefile, Polygon/Line (Y/N)	✓	✓	[313]
	Historical Landmarks	*Historical*	2014	Shapefile, Polygon/Point (Y/N)	✓	✓	[314]
	Military Installations	*Military*	2019	Shapefile, Polygon (Y/N)	✓	✓	[315]
	Mining Operations	*Mining*	2022	Shapefile, Polygon/Point (Y/N)	✓	✓	[316]
	National Parks	*Nat_Parks*	2022	Shapefile, Polygon (Y/N)	✓	✓	[317]
	Population Density by County	*Dens_15_19*	2019	Tabular* (persons/sq mi)	✓	✓	[311,82]
	Tribal Lands	*Trib_Land*	2018	Shapefile, Polygon (Y/N)	✓	✓	[318]
	Wildlife Refuges	*Wild_Refug*	2022	Shapefile, Polygon (Y/N)	✓	✓	[319]
Condition of the Natural Environment *(Environmental)*	Average Elevation	*Avg_Elevat*	2022	Raster (m)	✓	✓	[79]
	Average Temperature	*Avg_Temp*	2016	Raster (°C)	✓	✓	[320]
	Average Wind Speed	*Avg_Wind*	2017	Raster (m/s)	✓	✓	[321]
	Bat Habitat Range Count	*Bat_Count*	2018	Shapefile, Polygon (number)	✓	✓	[322]
	Bird Habitat Range Count	*Bird_Count*					
	Proportion of Rugged Land	*Prop_Rugg*	2022	Raster (number)	✓	✓	[79]
	Proportion of Undevelopable Land	*Undev_Land*	2019	Raster (number)	✓	✓	[80]
Distance to Infrastructure *(Technical)*	Nearest Airport	*Near_Air*	2022	Shapefile, Point (m)	✓	✓	[323]
	Nearest Hospital	*Near_Hosp*	2022	Shapefile, Point (m)	✓	✓	[324]
	Nearest Major Road	*Near_Roads*	2016	Shapefile, Line (m)	✓	✓	[325]
	Nearest Major Transmission Line	*Near_Trans*	2022	Shapefile, Line (m)	✓	✓	[326]
	Nearest School	*Near_Sch*	2022	Shapefile, Point (m)	✓	✓	[327,328]
Power Plant Maturity *(Technical)*	Age of the Nearest (Non-Wind) Power Plant	*Plant_Year*	2022	Shapefile, Point (year)	✓	✓	[329]
	Age of the Nearest Wind Farm	*Farm_Year*	2022	Shapefile, Point (year)	✓	✓	[73]

Table 5a: Predictors used in WiFSS-LRCA that are classified as Environmental or Technical, along with each predictor's name in the model, year of data preparation, datatype (and units), spatial scale (State/CONUS), and data source. Asterisked entries in the Datatype column are predictors that did not exist as pre-prepared datasets and thus were compiled manually.

Classification	Predictor	Codename	Year	Datatype (unit)	State	CONUS	Source
Competitiveness of Wind Energy Sales *(Economic)*	Electricity Cost	*Cost_15_19*	2021	Tabular ($)		✓	[312,78]
	Independent System Operator(s)	*ISO_YN*	2022	Shapefile, Polygon (Y/N)	✓	✓	[311,330]
	Nearest (Non-Wind) Power Plant	*Near_Plant*	2022	Shapefile, Point (m)	✓	✓	[329]
Land Value by State *(Economic)*	Farmland Value	*Farm_15_19*	2023	Tabular* ($)		✓	[312,331]
	Property Value	*Prop_15_19*	2020	Tabular ($)		✓	[312,332]
Renewable Energy and Energy Efficiency Incentives *(Economic)*	Investment Tax Credits	*In_Tax_Cre*	2023	Tabular* (Y/N)		✓	[312,333]
	Property Tax Exemptions	*Tax_Prop*					
	Sales Tax Abatements	*Tax_Sale*					
	Total Number of Incentives	*Numb_Incen*	2023	Tabular* (number)		✓	
Election Results *(Political)*	Gubernatorial Election Results by State	*Rep_Wins*	2023	Tabular* (number)		✓	[312,334]
	Presidential Election Results by County	*Dem_Wins*	2021	Tabular (number)	✓	✓	[311,335]
Government Legislation in Effect *(Political)*	Interconnection	*Interconn*	2023	Tabular* (Y/N)		✓	[312,333]
	Net Metering	*Net_Meter*					
	Renewable Portfolio Standard	*Renew_Port*					
	Size of Renewable Portfolio Standard Target	*Renew_Targ*	2023	Tabular* (number)		✓	
	Total Number of Statewide Legislative Pieces	*Numb_Pols*					
State Government Lobbies *(Political)*	Fossil Fuel Lobbies	*Foss_Lobbs*	2023	Tabular* (number)		✓	[312,336]
	Green Lobbies	*Gree_Lobbs*					
Condition of the Workforce *(Social)*	Employment Type by County	*Type_15_19*	2021	Tabular (%)	✓	✓	[311,337]
	Unemployment Rate by County	*Unem_15_19*	2021	Tabular (%)	✓	✓	[311,338]
Demographics by County *(Social)*	Percent Female Population	*Fem_15_19*	2019	Tabular (%)	✓	✓	[81,311]
	Percent Hispanic Population	*Hisp_15_19*					
	Percent of Population Under 25	*Avg_25*					
	Percent White Population	*Whit_15_19*					
	Public Support for Renewable Portfolio Standards	*supp_2018*	2022	Tabular (%)	✓	✓	[311,339]

Table 5b: Same as Table 5a but for the predictors that can be broadly classified as Economic, Political, or Social.

LRCA's LR equation over individual U.S. states, only being possible in CONUS model runs, as specified by the "State" and "CONUS" columns in Table 5a/5b. WiFSS-LRCA assumes that all enlisted predictors are relevant to past, present, and future wind farm siting decisions.

3.1.2. Justifications for Each Predictor's Inclusion.

Provided below is the reasoning for why each of the 47 predictors listed in Tables 5a and 5b were included in WiFSS-LRCA, discussed under the same Classification terms given in these two tables. The Codenames for each predictor are given in parentheses throughout this sub-section.

1. Pre-established Land Use (*Environmental Predictors*).

   Wind energy development may be incompatible, or even prohibited, in areas of land designated for specific purposes. Wind turbine installation can disrupt breeding and predation behaviors [341], and cause fragmentation of sensitive ecosystems [342] (*Critical; Wild_Refug*). Federal regulations for renewable energy development prevent impairment or damage to important landmarks, such as National Parks [343] or sites protected under the National Historic Preservation Act [344] (*Nat_Parks; Historical*). Commercial wind farms are also inadvisable in areas with established human activity. Wind turbines pose collision and radar detection hazards for military operations [345], nearby mining operations may compromise a turbine tower's foundations [226], and densely populated areas often lack the space for large wind farm projects [45] (*Military; Mining; Dens_15_19*). Finally, development of wind energy resources in tribal lands has been limited by "federal bureaucratic inefficiencies" [346] (*Trib_Land*).

2. Condition of the Natural Environment (*Environmental Predictors*).

   Suitable areas for wind energy development generally have mean wind speeds over 6.5 meters per second at 80 meters above the ground [347] (*Avg_Wind*). Moderate elevations are

preferable for this development since wind speed increases with height [348], but not over 2000 meters above sea level due to declining air density [349] (*Avg_Elevat*). The natural environment also presents construction challenges, such as damage caused by turbine blade icing in colder climates [350] and the difficulty of building and maintaining wind farms on rugged terrain [218] (*Avg_Temp; Prop_Rugg*). Improper wind farm siting poses a collision risk to birds [351] and bats [352], with bats at an added risk of barotrauma around rotating blades [353] (*Bat_Count; Bird_Count*). Moreover, certain land cover types are considered unsuitable for wind energy development, such as forested land due to lack of open space and downwind turbulence [354] and the sensitive ecosystems of wetlands [355] (*Undev_Land*).

3. Distance to Infrastructure (*Technical Predictors*).

   Wind farm construction is more feasible nearer transmission lines [47], due to lower costs of connecting new wind farms to the electricity grid [356] (*Near_Trans*). Similarly, nearby roads facilitate wind farm maintenance and further reduce costs [357], though roads being too close poses a safety risk [57] (*Near_Roads*). Complaints of wind farms causing shadow flicker, noise, and aesthetic pollution [358] mean that some states now legislate setback distances around specific infrastructure, such as hospitals and schools [359] (*Near_Hosp; Near_Sch*). Wind farms also pose safety and visibility risks to airports [244], along with disruption to radar signals for navigation and weather monitoring [360] (*Near_Air*).

4. Power Plant Maturity (*Technical Predictors*).

   The United States' oldest commercial wind farms were constructed in California in the 1980s [85]. Such maturity suggests well-established infrastructure and protocols for wind energy development and thus less of a technical challenge for future expansion (*Farm_Year*). Furthermore, wind energy could be considered more attractive if nearby non-wind power

plants have aged beyond their life expectancy [361] and are thus more likely to fail [362] or be too inefficient to be worth maintaining [363] (*Plant_Year*).

5. Competitiveness of Wind Energy Sales (*Economic Predictors*).

   Studies from Spain [364] and Germany [365] show that wholesale electricity from wind farms can be cheaper than from other sources. However, wind energy's intermittency comes with a greater risk of negative pricing [366] (*Cost_15_19*). Hence, the monitoring of electricity transmission and prices by Independent System Operators (ISOs) is an asset to WiFSS [367]. ISOs enlist wind forecast data to allow wind energy developers to participate in day-ahead sales markets [368], thus enabling electricity sales [369] and limiting manual wind farm curtailment [370] (*ISO_YN*). Conversely, nearby non-wind power plants that already provide local jobs [371] and sell electricity at a low price [372] can limit a wind farm's economic competitiveness (*Near_Plant*).

6. Land Value by State (*Economic Predictors*).

   State governments may be more willing to subsidize and invest in developing land for agricultural and residential purposes, rather than for wind energy development, especially in states where the former's economic values are high (*Farm_15_19; Prop_15_19*). WiFSS-LRCA does not account for how farmland/property value is itself affected by wind energy development, since studies such as Vyn and McCollough [373], Sampson et al. [374], and Castleberry and Greene [375] suggest a lack of meaningful effect.

7. Renewable Energy and Energy Efficiency Incentives (*Economic Predictors*).

   Wind energy's adoption can be hastened by offering economic incentives, such as rebates [376] and tax incentives [377]. If state governments and utilities offer incentives for renewable energy and energy efficiency technologies, this may suggest an economy that

welcomes commercial wind energy development [378]. The existence of state-level financial incentives (*In_Tax_Cre; Tax_Prop; Tax_Sale*) and the number of current and expired incentives (*Numb_Incen*) [333] are compiled to represent these predictors.

8. Election Results (*Political Predictors*)

   Within American politics, progressive ideals often support decarbonization initiatives [379], with Democrats being generally more proactive in passing legislation that supports the wind energy sector [380]. However, a partisan lens on energy politics can be myopic because some Republicans do recognize wind energy's economic potential [381] and its role in mitigating environmental pollutants [87]. The electorate's attitude toward wind energy is approximated using gubernatorial (*Rep_Wins*) and presidential (*Dem_Wins*) election results since 1976 [338] and 2000 [339], respectively.

9. Government Legislation in Effect (*Political Predictors*).

   Passing legislation theoretically represents citizens' interest in wind farm siting decisions [382]. Example legislation includes Renewable Portfolio Standards (RPS) that establish future targets for total renewable energy capacity [383], and interconnection rules for equitable transmission of renewable electricity [384]. Laws such as RPS vary state-by-state in terms of stringency, and their importance for driving wind energy sector growth over non-political drivers, such as resource abundance [385] and the energy market [310], has been questioned. As with financial incentives (Section 2.1.7), the existence of common state-level regulations (*Interconn; Net_Meter; Renew_Port*), the number of regulations in effect (*Numb_Pols*), and the stringency of RPS (*Renew_Targ*) are compiled to represent these predictors [333].

10. State Government Lobbies (*Political Predictors*).

    While the cost competitiveness and maturity of renewable energy are reducing the influence of interest groups lobbying for/against them [386], fossil fuel companies can still lose revenue if legislators enable wind energy's expansion [387]. Since clean energy and environmental issues are politically intertwined [388], both the number of state-registered renewable energy and environmental lobbies (*Gree_Lobbs*) and the number of registered fossil fuel and mining lobbies (*Foss_Lobbs*) are compiled to represent lobbyism's effects on WiFSS [336].

11. Condition of the Workforce (*Social Predictors*).

    Wind energy development is a force for creating local jobs [389], particularly in rural areas with high unemployment rates [390]. Workers in established sectors often possess skills easily transferable to the renewable energy sector [391], especially those working in manufacturing and construction [392]. However, most jobs created by wind energy development are temporary [393], and large mining and extraction workforces may indicate potential for pushback against wind energy development [88], rather than a potential to retrain workforces. High county-level unemployment rates (*Unem_15_19*) and large utility, construction, and manufacturing workforces (*Type_15_19*) are assumed to be social benefits to WiFSS.

12. Demographics by County (*Social Predictors*).

    While Sokoloski et al. [394] find that older demographics are associated with lower wind energy support, Brannstrom et al. [395] find this link to lack statistical significance. Statistical significance similarly lacks in studies that find greater support for wind energy in the United States among women [396,397], and among Black, Indigenous and People of

Color (BIPOC) [398]. These demographic groups are represented among WiFSS-LRCA's predictors (A*vg_25; Fem_15_19; Hisp_15_19; Whit_15_19*). Variability in actual public opinions is also represented using county-level responses to questions about Renewable Portfolio Standards (*supp_2018*) [339].

3.2. Data Pre-Processing and Aggregation.

The selected datasets come in a variety of data types and resolutions, as shown by Tables 5a and 5b, necessitating their aggregation to a shared format usable by WiFSS-LRCA. This dissertation's process of raw dataset collection, pre-processing, and aggregation takes cues from Plassin et al.'s [77] construction of an integrated geodatabase for the Rio Grande Basin, with a similar purpose of facilitating repeated experiments in the hands of end-users (see Section 1.2). Other WiFSS studies have performed this aggregation by converting all datasets into a raster format for continuous WiFSS evaluation in space [58], or by applying datasets to wind turbine point locations to evaluate their likelihood of acceptance [62]. For WiFSS-LRCA, all datasets were aggregated over a gridded surface covering a single U.S. state or the CONUS, with the suitability for each grid cell containing a commercial wind farm calculated by the model's LR equation. Previous contexts for applying an LR equation to aggregated grid cell data include modeling grassland fire occurrence [399], erosion susceptibility [400], and radon potential in rocks [401]. WiFSS-LRCA enlists hexagonal grid cells for their minimization of edge effects along state and coastal borders [402], and for better approximation of radial distance between wind farms and spatial features (e.g., transmission lines, roads, power stations) compared to square grid cells [403]. Hexagonal surfaces covering a single U.S. state (or the CONUS) that illustrate the probabilities of commercial wind farm occurrence (i.e., WiFSS surfaces), and the grid cells projected to acquire wind farms by the year 2050, are WiFSS-LRCA's primary outputs.

Performing the dataset aggregation at 20 different grid cell resolutions allows WiFSS-LRCA's sensitivity to commercial wind farm capacities (in Megawatts (MW)) to be tested [404]. Typical commercial wind farm density (in acres/MW) in the CONUS varies, quoted as 30 acres/MW by the National Renewable Energy Laboratory [405] and as 85 acres/MW by the University of Michigan [406]. The United States Wind Turbine Database [73] is used to express the CONUS' present range of commercial wind farm capacities as quintiles (20th to 100th percentile). These quintiles and quoted densities are used to define WiFSS-LRCA's 20 grid cell resolutions (in acres), summarized in Table 6. Aggregating the predictors' datasets at multiple resolutions has two benefits: 1) the sensitivity of WiFSS-LRCA to grid cell size can be examined; and 2) the performance of custom model projections can be diagnosed, which is useful to end-users wishing to investigate important predictors and suitable locations for a specific wind farm size.

		Wind Farm Capacity, MW *(Percentile)*				
	x	30 *(20th)*	90 *(40th)*	150 *(60th)*	200 *(80th)*	520 *(100th)*
Wind Farm Density, acres/MW	25	750	2250	3750	5000	13000
	45	1350	4050	6750	9000	23400
	65	1950	5850	9750	13000	33800
	85	2550	7650	12750	17000	44200

Table 6: Area of an individual hexagonal grid cell (in acres) for all 20 aggregations of WiFSS-LRCA's predictors.

Some datasets required pre-processing before aggregation onto the hexagonal grid cells:

- Bat/Bird Habitat Range Count (*Bat_Count; Bird_Count*) – the USGS Gap Analysis Project maintains habitat ranges of bird and bat species as separate shapefiles [322]. Each dataset layer was combined to produce a single shapefile, allowing for counts of the number of species found over each grid cell.

- Average Elevation and Proportion of Rugged Land (*Avg_Elevat; Prop_Rugg*) – Digital Elevation Model rasters with a one arc-second (30-meter) resolution [79] were merged to create one large elevation dataset, from which a land slope raster across the CONUS was also derived. Raster cells with a land slope greater than 7% were classified as unsuitable ("N", otherwise suitable ("Y")) for commercial wind farm construction [55,125].
- Population Density by County (*Dens_15_19*) – besides spatially joining to a shapefile of the CONUS' county borders (Section 3.1.1), the annually-averaged populations from 2015-2019 of each county [82] were divided by its surface area to calculate population density for each county separately.

- Proportion of Undevelopable Land (*Undev_Land*) – WiFSS-LRCA considers forested [354], riverine [254], urban [48], and wetland [355] land cover types to be unsuitable for wind energy development. The National Land Cover Database raster [80] was thus reclassified such that raster cells possessing any of these four land types were assigned 0, and all others were assigned 1.

The applied aggregation method depends on each predictor's datatype (raster, vector (shapefile), or tabular), and whether said predictor is quantitative or Boolean. Predictors of raster datatype (*Avg_Elevat*; *Avg_Temp*; *Avg_Wind*; *Prop_Rugg*; *Undev_Land*) were averaged over the area of

each grid cell to yield a mean cell value. Vector datasets representing proximity to commercial wind farms (i.e., all "Distance to Infrastructure" predictors and *Near_Plant*) are quantitative and were thus aggregated as the distance between a grid cell's centroid and the closest relevant physical feature. The "Power Plant Maturity" predictors (also vector datasets) were aggregated as the age of the closest wind farm and non-wind power plant to this centroid. Vector datasets representing Booleans (*Critical*; *Historical*; *Military*; *Mining*; *Nat_Parks*; *Trib_Land*; *Wild_Refug*; *ISO_YN*) were aggregated as whether or not a grid cell overlapped with each of these features and were thus assigned "Y" (yes) or "N" (no). The vector datasets for *Bird_Count* and *Bat_Count* were unique, with each grid cell acquiring the number of overlapping feature layers (each layer represents a single species). The remaining predictors, all of tabular datatype, were aggregated based on overlapping features, whether Boolean (*In_Tax_Cre*; *Tax_Prop*; *Tax_Sale*; *Interconn*; *Net_Meter*; *Renew_Port*) or quantitative (*Dens_15_19*; *Cost_15_19*; *Farm_15_19*; *Prop_15_19*; *Numb_Incen*; *Rep_Wins*; *Dem_Wins*; *Renew_Targ*; *Numb_Pols*; *Foss_Lobbs*; *Gree_Lobbs*; *Type_15_19*; *Unem_15_19*; *Fem_15_19*; *Hisp_15_19*; *Avg_25*; *Whit_15_19*; *supp_2018*). Figure 8 illustrates the outcome of this aggregation for four of WiFSS-LRCA's 47 predictors, each at a different grid cell resolution. All aggregated data serve as independent variables of WiFSS-LRCA's logistic regression equation.

3.3. The Logistic Regression Equation.

One of the core components of WiFSS-LRCA is its binary logistic regression equation. The model uses this LR equation to calculate the probability of each grid cell of a hexagonal surface containing a commercial wind farm, such that the logit of this probability is a function of the linear combination of the aggregated predictors [93]. The binary dependent variable is the

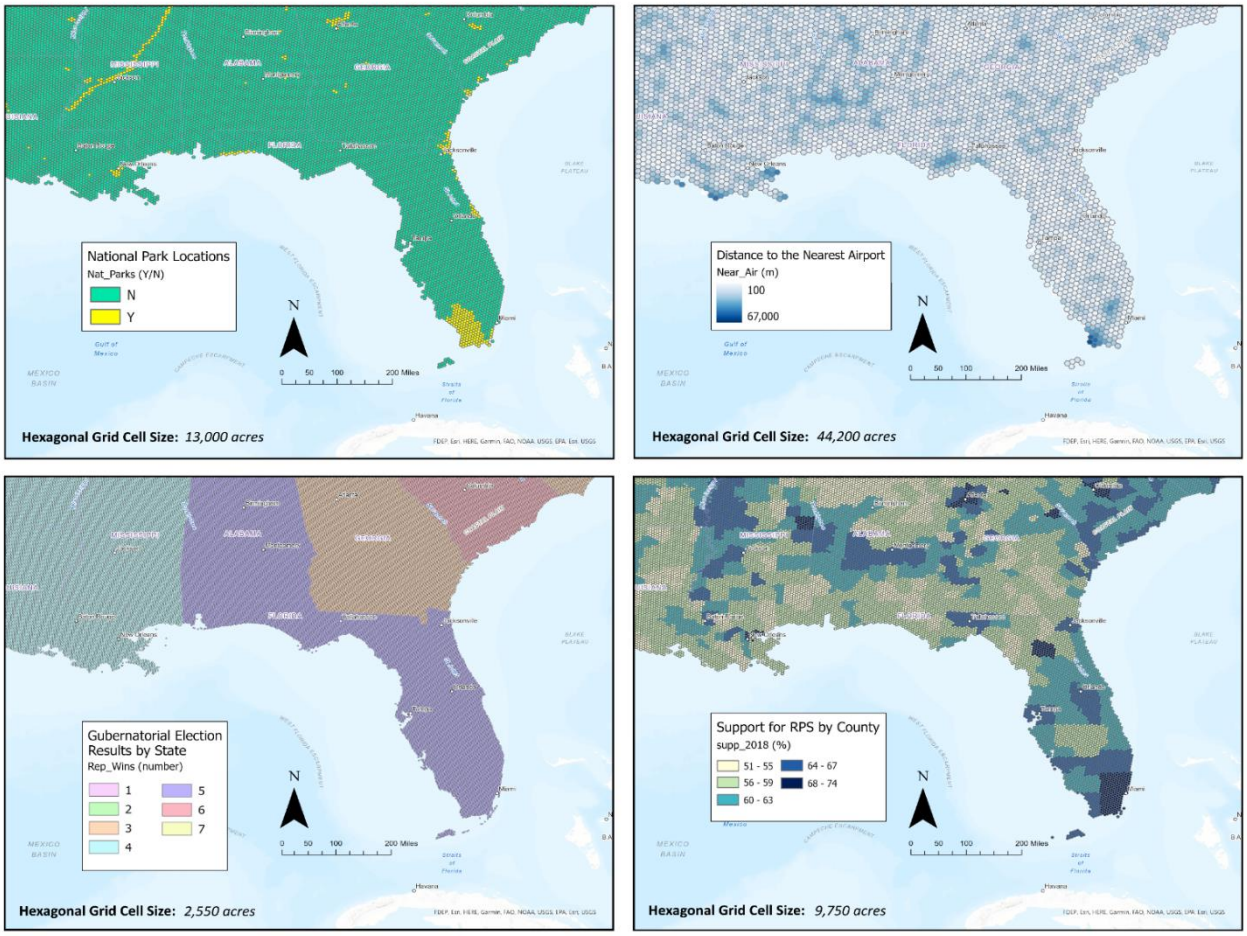


Figure 8: Aggregations of four of the predictors that comprise WiFSS-LRCA's aggregated dataset, for CONUS-level model runs at different resolutions, zoomed into the Southeast United States. From top-left to bottom-right are the aggregations of National Park Locations (*Nat_Parks*), Distance to the Nearest Airport (*Near_Air*), Gubernatorial Election Results by State *(Rep_Wins*), and Support for Renewable Portfolio Standards by County (*supp_2018*). Hexagonal grid cell sizes are given on each map (see Table 6). Basemap from Esri [407].

probabilistic identification of wind turbines contained in the United States Wind Turbine Database [73]. Equation 1 below is the LR equation applied to all grid cells in any study area:

$$P_i = \frac{1}{1 + \exp(Z_i)}, \text{where } Z_i = Y_0 + \sum_{j=1}^{k} Y_j x_{ij} \qquad (\mathbf{1})$$

where P_i is the probability computed for grid cell i, Z_i is the sum of the linear combination of predictors, Y_0 is the intercept, Y_j is the coefficient of predictor j, x_{ij} is the value of predictor j in grid cell i, and k is the number of predictors.

The aggregated predictor datasets (see Section 3.1.1) and the binary dependent variable must satisfy four crucial assumptions before application to Equation 1. Firstly, grid cells containing extreme data values are preemptively identified before training and testing WiFSS-LRCA, thus reducing the risk of these outlying data affecting the model's parameterization [408]. WiFSS-LRCA excludes these identified grid cells based on a Cook's distance test [409]. Secondly, aggregated data for any pair of independent variables must lack multicollinearity, i.e., predictors must not be strongly correlated with each other. WiFSS-LRCA assesses pairwise multicollinearity by computing Variance Inflation Factors (VIFs) for all pairs of predictors [410], such that each pair with a VIF above 10 [411] is removed from the LR equation. Thirdly, WiFSS-LRCA verifies that a linear relationship exists between the continuous predictors and P_i with a Box-Tidwell test [412], with the test's threshold p-value for assessing linearity modified by applying a Bonferroni correction [63]. This linearity assumption is more robust at higher grid cell resolutions and in U.S. states containing more wind farms, since having more grid cells in both classes of the dependent variable reduces the risk of quasi-complete separation [413]. WiFSS-LRCA removes continuous independent variables that breach this assumption. Finally, the aggregated data must be independent, meaning a predictor must not be constant across all grid cells [414], otherwise the predictor has no effect on the binary dependent variable. Predictors lacking independence are again removed prior to training and testing Equation 1. Any attempt at running WiFSS-LRCA applies these four assumptions to the predictor datasets as a first step.

3.3.1. Calibration, Validation, and Analysis of LR Equation Performance.

Meeting this dissertation's purpose of projecting wind farm siting futures (see Section 1.2) requires that WiFSS-LRCA's logistic regression equation can correctly predict the locations of existing commercial wind farms. WiFSS-LRCA thus trains and tests its LR equation in the model's first iteration. The training step samples 75% of all hexagonal grid cells (over a single U.S. state or the CONUS) to calibrate predictor coefficients for Equation 1 ($Y_1, Y_2, \dots, Y_k$) that maximize its goodness-of-fit [415]. Each training grid cell sample contains a stratified number of grid cells that possess a wind farm [416], thus ensuring that the LR equation's calibrated coefficients reflect grid cells that do and do not contain wind farms. Goodness-of-fit assessment requires computing the log-likelihood ratio (λ) of the trained LR equation (all predictors) versus a null version of the same equation (intercept only), computed using Equation 2 [417]:

$$\lambda = -2 \ln\left(\frac{null\ model\ likelihood}{trained\ model\ likelihood}\right) \qquad (\mathbf{2})$$

where the statistical significance ($p < 0.05$) of λ assumes a chi-square distribution [418]. McFadden's Adjusted Pseudo R-squared (R^2_{MFA}) re-expresses λ as a bounded value; WiFSS-LRCA uses McFadden's version for its numerator adjustment based on the number of predictors (k) [419], as shown in Equation 3:

$$R^2_{MFA} = 1 - \frac{\ln(trained\ model\ likelihood) - k}{\ln\ (null\ model\ likelihood)} \qquad (\mathbf{3})$$

Random selection of training grid cells for logistic regression necessitates repeating this calibration process 30 times [420], a standard sample size for deriving sample statistics [421]. These statistics are the median coefficients that maximize the LR equation's goodness-of-fit, as well as the number of times the trained model outperforms the null model.

WiFSS-LRCA illustrates the range of coefficients obtained from these repeated calibrations by computing Odds Ratios (ORs) for each predictor. ORs quantify how strongly associated unit changes in a predictor are with the likelihood of a binary event (i.e., a grid cell containing a wind farm) taking place [422]. Sperandei [423] notes ORs to be a standard approach of representing effects on a binary event within an LR equation, with an odds ratio below (above) one signifying a lower (greater) chance of the binary event's occurrence [424]. As such, ORs provide important insight into the geographical properties (e.g., wind speed, state legislature, infrastructure, etc.) associated with WiFSS-LRCA's projected wind farm siting futures, one of this dissertation's purposes (see Section 1.2). Tables 7a and 7b summarize expected median ORs for each predictor, and reasons for each expectation, to be compared to the ORs produced by WiFSS-LRCA.

Calibration of the LR equation is followed by WiFSS-LRCA using the remaining 25% of grid cells to validate this equation's ability to correctly predict which grid cells contain commercial wind farms. Validation using these testing grid cells gives credence to the WiFSS surfaces later constructed by WiFSS-LRCA. The model enlists two metrics to perform this validation:

- *The Receiver Operating Characteristic (ROC)* – a plotted summary of the LR equation's ability to correctly classify grid cells as containing a wind farm across multiple classification thresholds [425]. Testing thresholds from 0 to 1 identifies changing proportions of Type 1 errors, as well as the probability that maximizes correct grid cell classifications [426]. Accompanying ROC is the Area Under Curve (AUC) statistic, with a greater AUC suggesting a high rate of correct grid cell classification at more thresholds [427].
- *Confusion Matrices* – a presentation of the number of (in)correctly classified testing grid cells at a given threshold [428]. The probability that maximized grid cell classification accuracy when constructing the ROC curves is used by WiFSS-LRCA to present the number

Predictor *(Codename)*	Greater Than 1?	Less Than 1?	Explanation
Critical Habitats *(Critical)*		✓	Wind farms present risks to breeding and ecosystem fragmentation.
Historical Landmarks *(Historical)*		✓	Spaces and objects protected from development by federal law.
Military Installations *(Military)*		✓	Greater risk of collision and radar detection hazards.
Mining Operations *(Mining)*		✓	Weaker foundations for turbine towers.
National Parks *(Nat_Parks)*		✓	Natural spaces protected from excessive human development.
Population Density by County *(Dens_15_19)*		✓	Too much urbanization to allow for large commercial wind farms.
Tribal Lands *(Trib_Land)*		✓	Bureaucratic inefficiency can limit wind energy development.
Wildlife Refuges *(Wild_Refug)*		✓	Wind farms present risks to breeding and ecosystem fragmentation.
Average Elevation *(Avg_Elevat)*	✓		Greater wind speed at higher altitudes.
Average Temperature *(Avg_Temp)*	✓		Lower blade icing risk.
Average Wind Speed *(Avg_Wind)*	✓		Greater energy source to capture.
Bat Habitat Range Count *(Bat_Count)*		✓	Habitat disruption risk (collisions, predation behavior, barotrauma).
Bird Habitat Range Count *(Bird_Count)*			
Proportion of Rugged Land *(Prop_Rugg)*		✓	Lack of flat open spaces for wind farm construction.
Proportion of Undevelopable Land *(Undev_Land)*		✓	Harder to build wind farms in forests, wetlands, rivers, and urban areas.
Nearest Airport *(Near_Air)*	✓		Less concern for navigation disruption.
Nearest Hospital *(Near_Hosp)*	✓		Setback distance laws are less of a concern.
Nearest Major Road *(Near_Roads)*		✓	Transporting construction and maintenance materials becomes more expensive.
Nearest Major Transmission Line *(Near_Trans)*		✓	More costly to connect wind farms to the electricity grid.
Nearest School *(Near_Sch)*	✓		Setback distance laws are less of a concern.
Age of the Nearest (Non-Wind) Power Plant *(Plant_Year)*	✓		Wind farms are a candidate to replace old/failing power plants.
Age of the Nearest Wind Farm *(Farm_Year)*	✓		Wind energy infrastructure is more mature and established.

Table 7a: Expected odds ratios produced by WiFSS-LRCA, assuming a unit increase in each predictor. Note that Boolean predictors (Y/N see Table 5a/5b) are re-expressed by the model such that N = 0 and Y = 1.

Predictor *(Codename)*	Greater Than 1?	Less Than 1?	Explanation
Electricity Cost *(Cost_15_19)*	✓		Electricity from wind provides a cheaper alternative.
Independent System Operator(s) *(ISO_YN)*	✓		Monitoring of fair electricity prices and transmission is in effect.
Nearest (Non-Wind) Power Plant *(Near_Plant)*	✓		Reduced competition for providing local jobs and cheap electricity.
Farmland Value *(Farm_15_19)*		✓	Wind farm development is a less profitable means of using the land.
Property Value *(Prop_15_19)*			
Investment Tax Credits *(In_Tax_Cre)*	✓		Economic supports from government and utilities make wind energy development more affordable.
Property Tax Exemptions *(Tax_Prop)*			
Sales Tax Abatements *(Tax_Sale)*			
Total Number of Incentives *(Numb_Incen)*			
Gubernatorial Election Results by State *(Rep_Wins)*		✓	Republicans tend to be less supportive of wind energy development.
Presidential Election Results by County *(Dem_Wins)*	✓		Democrats tend to be more supportive of wind energy development.
Interconnection *(Interconn)*	✓		Political will exists to ensure that growth of the wind energy sector is supported.
Net Metering *(Net_Meter)*			
Renewable Portfolio Standard *(Renew_Port)*			
Size of Renewable Portfolio Standard Target *(Renew_Targ)*			
Total Number of Statewide Legislative Pieces *(Numb_Pols)*			
Fossil Fuel Lobbies *(Foss_Lobbs)*		✓	More organizations are influencing politics against wind energy development.
Green Lobbies *(Gree_Lobbs)*	✓		More organizations are influencing politics in favor of wind energy development.
Employment Type by County *(Type_15_19)*	✓		A greater number of people possess employable, transferable skills.
Unemployment Rate by County *(Unem_15_19)*	✓		Opportunity exists to create new jobs, especially in poverty-stricken areas.
Percent Female Population *(Fem_15_19)*	✓		Younger, marginalized demographics tend to be more supportive of renewable energy.
Percent Hispanic Population *(Hisp_15_19)*			
Percent of Population Under 25 *(Avg_25)*			
Percent White Population *(Whit_15_19)*		✓	Support for renewable energy tends to be higher among BIPOC.
Public Support for Renewable Portfolio Standards *(supp_2018)*	✓		Support from communities for wind energy growth is likely to be higher.

Table 7b: Same as Table 7a but for the remaining predictors.

of true/false positive/negative testing grid cell classifications. Confusion matrices provide a non-spatial summary of accuracy by comparing expected and observed grid cell states [429].

As with the calibration step, these two metrics are computed by WiFSS-LRCA 30 times to account for randomness in testing grid cell samples [420], thus validating the LR equation's performance based on median confusion matrices, ROC curves, and AUC statistics. Following validation is use of the LR equation to calculate the probability of commercial wind farm occurrence across all grid cells in the selected study area. WiFSS-LRCA constructs four boxplots, with one plot for each of the four grid cell classifications (true/false positive/negative). If the equation's grid cell classifications are robust, then these boxplots will illustrate a statistically significant difference in the rank order of the true positive and false positive (true negative and false negative) plots ($p < 0.05$), according to a Mann-Whitney U-test [430]. Statistically significant differences suggest a non-random reason for WiFSS-LRCA correctly classifying the highest (lowest) probability grid cells as containing (not containing) a commercial wind farm. Possible reasons come from interpreting the ORs obtained from calibrating the LR equation, with large OR values suggesting predictors having strong associations with predicted grid cell states, thus providing a geographical basis for WiFSS-LRCA's outcomes. It should be noted that this calibration, validation, and analysis of model performance are done in WiFSS-LRCA's first iteration, meaning that they are only representative of the model's ability to construct wind farm site suitability surfaces that reflect present conditions.

3.3.2. WiFSS Surface Construction and Model Caveats.

The probabilities computed for all grid cells by Equation 1 are used by WiFSS-LRCA to create a map of present WiFSS across the selected study area. Doing so is beneficial to end-users interested in assessing present suitability for wind energy development in addition to future

suitability, while also presenting the outcomes of the trained and tested LR equation in a more accessible format, again being purposes laid out for this dissertation (see Section 1.2). Furthermore, a high rate of correct classification suggested by the present WiFSS map, as well as the aforementioned metrics, should mean that future wind farm locations projected by WiFSS-LRCA should reflect the predictors most strongly associated with wind farm siting decisions. Constructing a map of present WiFSS yields three key outputs (as shown in an example for Wyoming in Figure 9):

1. The probability (WiFSS) of a given grid cell containing a commercial wind farm in WiFSS-LRCA's first iteration (*"Probab*").
2. The classification of a grid cell as true positive, false positive, true negative, or false negative ("*Cell_State*"), as classified when creating the boxplots (see Section 3.3.1).
3. The Getis-Ord (Gi*) statistic [431] computed for each grid cell, to assess statistically significant ($p < 0.05$) clustering of grid cells classified as true positive or false positive ("*GiZScore*", "*GiPValue*"). Clusters of grid cells with high Gi* scores suggest high-probability grid cells being scored as such due to commonly suitable conditions for wind farm siting. Applications of Gi* to identify clustered probabilities generated from LR range from juvenile delinquency [432] to leptospirosis infection [433].

There exist two caveats to training and testing an LR equation and then using it to construct suitability surfaces. Firstly, running the LR equation at the state-level is only possible in U.S. states that contain at least two grid cells that overlap with commercial wind farms, otherwise the stratified training and testing datasets cannot be produced [416]. As such, unless performing a CONUS model run, present WiFSS cannot be assessed using WiFSS-LRCA for the following

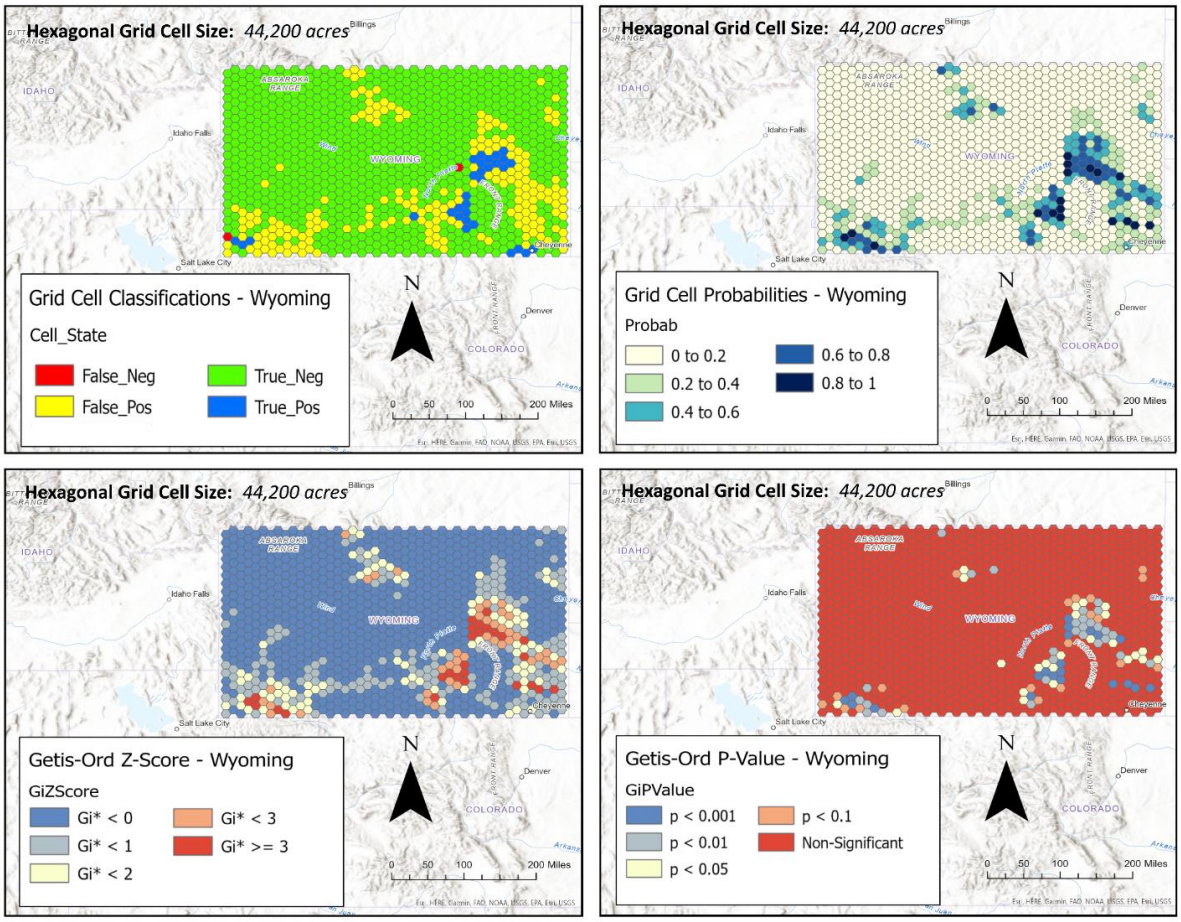


Figure 9: Example fields of the hexagonal gridded surface produced by WiFSS-LRCA's first iteration over the state of Wyoming. From top-left to bottom-right are the grid cell classifications (*Cell_State*), grid cell probabilities (*Probab*), and the Getis-Ord z-scores and p-values (*GiZScore; GiPValue*). Hexagonal grid cell sizes are given on each map (see Table 6). Basemap from Esri [407].

U.S. states: Alabama, Arkansas, Connecticut, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, New Jersey, Rhode Island, South Carolina, Tennessee, and Virginia. Secondly, WiFSS-LRCA excludes certain predictors when applied over an individual U.S. state rather than the CONUS, as summarized in Table 5a/5b. Of the 47 predictors enlisted in CONUS runs of WiFSS-LRCA, 32 are used in state-level runs. The other 15 predictors consist of datasets too coarse (e.g., state-averaged farmland values (*Farm_15_19*), number of state government lobbies (*Foss_Lobbs; Gree_Lobbs*)) to vary over a single U.S. state. Including these 15 predictors would invalidate LR's assumption of independent observations [414].

3.4. Projections using Logistic Regression-Cellular Automata.

The other core component of WiFSS-LRCA is its cellular automata, the means of updating a hexagonal grid using a set of transition rules [67]. The CA iteratively convert grid cells to acquire a commercial wind farm (excluding cells that already possess one) to thus project patterns of future wind energy development. WiFSS-LRCA performs six iterations after its first one (see Section 3.3), each one representing a five-year wind farm installation timestep [340] from the years 2025 to 2050. The grid cells that experience binary land-use change in each iteration are those with the greatest probability of obtaining a wind farm. The number of grid cells that change state in each timestep is based on projected wind energy capacity gains from the United States Department of Energy's *Wind Vision* report [92] and the grid cell sizes defined in Table 6, although the overall results may vary since these gains are customizable by end-users. Using CA to modify grid cell states based on a prescribed demand, i.e., gained wind energy capacity every five years, has precedent in studies by Barredo et al. [66] and Shu et al. [75], both of whom imposed an external increase in urban area to determine the number of cells that became urbanized in each iteration. Equation 4 summarizes the transition rules applied by the CA to all grid cells during each iteration of WiFSS-LRCA:

$$Prob_i^t = Const_i * Neighb_i^t * P_i \qquad (\mathbf{4})$$

$Prob_i^t$ is the probability of grid cell i obtaining a wind farm during timestep t, $Const_i$ is whether wind energy development is constrained in the grid cell, $Neighb_i^t$ is the neighborhood effect factor of the grid cell computed using surrounding grid cell states, and P_i is computed using Equation 1, the latter representing the CA's "equation-based" transition rule [434] ($Const_i$ and $Neighb_i^t$ are described below). If grid cell i already contains a commercial wind farm, Equation

4 is not applied. WiFSS-LRCA assumes that current constraints on wind energy development persist in future, hence $Const_i$ does not update with each iteration.

WiFSS-LRCA accounts for unfeasible wind energy development in certain grid cells using constraints ($Const_i$). Representing predictors as Boolean constraints on wind energy development is common in GIS-MCDA approaches to WiFSS (see Section 2.1.1), such that constructing the composite suitability surface illustrates areas in which this development is impossible or not recommended, such as wildlife refuges [233], areas too far from transmission lines [236], or places with low average wind speeds [48]. In modeling studies that use CA, a constraint transition rule assigns a value of 0 to all grid cells that violate at least one constraint, while assigning 1 to those that do not [435], with the constraints similarly representative of conflicting land-use [436]. Table 8 details the predictors used as WiFSS-LRCA's default Boolean constraints. The selected predictors and the (numerical) values that define each constraint follow from the systematic review [40] (see Section 2.2) and from the wider literature. If the aggregated predictor data in grid cell i violate just one of these constraints, $Const_i$ is set as 0 (otherwise set as 1), meaning $Prob_i^t$ equals 0 in all timesteps. These constraints are fully customizable as part of this dissertation's purpose of preparedness for experimentation by end-users (see Section 1.2).

The neighborhood effect term of Equation 4 ($Neighb_i^t$) represents the influence that nearby grid cells have on the binary state of grid cell i, intended to approximate a decaying influence on land-use change with increasing distance [69]. A CA model's output is sensitive to both the selected neighborhood size (e.g., 2x2 cells versus 3x3 cells) [437] and the prescribed neighborhood shape (e.g., Moore versus von Neumann) [438]. This dissertation uses a hexagonal neighborhood shape, given WiFSS-LRCA's use of hexagonal grid cells. Example applications of hexagonal neighborhoods in CA studies range from simulating traffic flows [439] to floodwater

Predictor	Constraint Value	Source(s)
Nearest Airport	< 2,500 meters	[44,47,233]
Nearest Hospital	< 2,500 meters	[359]
Nearest Major Road	< 500 meters AND >10,000 meters	[216,236,240]
Nearest Major Transmission Line	< 250 meters AND >10,000 meters	[233,236,272]
Nearest School	< 2,500 meters	[359]
Nearest (Non-Wind) Power Plant	> 10,000 meters	[222,255]
Critical Habitats	Prohibited (N)	[125,265]
Historical Landmarks	Prohibited (N)	[47,257]
Military Installations	Prohibited (N)	[191,272]
Mining Operations	Prohibited (N)	[24,246]
National Parks	Prohibited (N)	[125,265]
Tribal Lands	Prohibited (N)	[346]
Wildlife Refuges	Prohibited (N)	[233,265]
Average Elevation	> 2,000 meters	[110,216,233]
Average Temperature	< 0 °C	[350]
Average Wind Speed	< 4 m/s	[48,56,111]

Table 8: Default Boolean constraints of WiFSS-LRCA when applied over a given grid cell. Sources are given for the value that defines each constraint.

dynamics [440], with Nugraha et al. [441] noting an equivalent (if not improved) simulated positional accuracy of hexagonal CA neighborhoods versus other prescribed shapes. Values of $Neighb_i^t$ in each iteration are calculated by WiFSS-LRCA using Equation 5, which defines the neighborhood effect transition rule. The equation comes from the formulation given by Shu et al. [76], altered for a hexagonal neighborhood shape:

$$Neighb_i^t = \frac{\sum_n (Cell_i = wind\ farm)}{3n(n+1)} \qquad (\mathbf{5})$$

The numerator is the total hexagons in grid cell i's neighborhood that contain a wind farm, the denominator is the total number of neighboring cells, and n is the neighborhood size. $Neighb_i^t$

takes a range of values from 0 to 1, depending on how many neighboring grid cells contain a wind farm. Since applying Equation 4 causes the highest-probability grid cells to obtain wind farms up to the prescribed demand, $Neighb_i^t$ values for all grid cells are updated before the next model iteration. It is the neighborhood effect term that allows WiFSS-LRCA to account for cluster-like patterns of land use change [72], i.e., the assumption that future commercial wind farms will continue to be installed in common geographical regions across the CONUS [73].

Completing all six iterations yields a map of grid cells projected to obtain commercial wind farms across a given U.S. state or the CONUS by the year 2050. Projections may cease before the sixth iteration if all remaining grid cells violate the set constraints, or none are left within the neighborhoods of current or projected wind farms. Example maps produced by running WiFSS-LRCA out to the year 2050 over Oklahoma are shown in Figure 10, which like the maps of present WiFSS shown in Figure 9 possess several key fields:

1. The grid cells eligible to gain a commercial wind farm, based on WiFSS-LRCA's constraint rule and the default constraints given in Table 8 (*"Constraint"*).
2. The value of $Neighb_i^t$ computed for each grid cell using Equation 5, presented for the first (*"Neighborhood"*) and final (*"Neighb_Update"*) iterations of WiFSS-LRCA.
3. The WiFSS surfaces constructed in every iteration of WIFSS-LRCA, consisting of the values of $Prob_i^t$ calculated for each grid cell (*"Probab_2025"*, *"Probab_2030"*, *"Probab_2035", "Probab_2040"*, *"Probab_2045", "Probab_2050").*
4. The grid cells projected to gain commercial wind farms by WiFSS-LRCA's final iteration, with grid cell color identifying the iteration in which land-use change takes place (*"Wind_Turb_Fut"*).

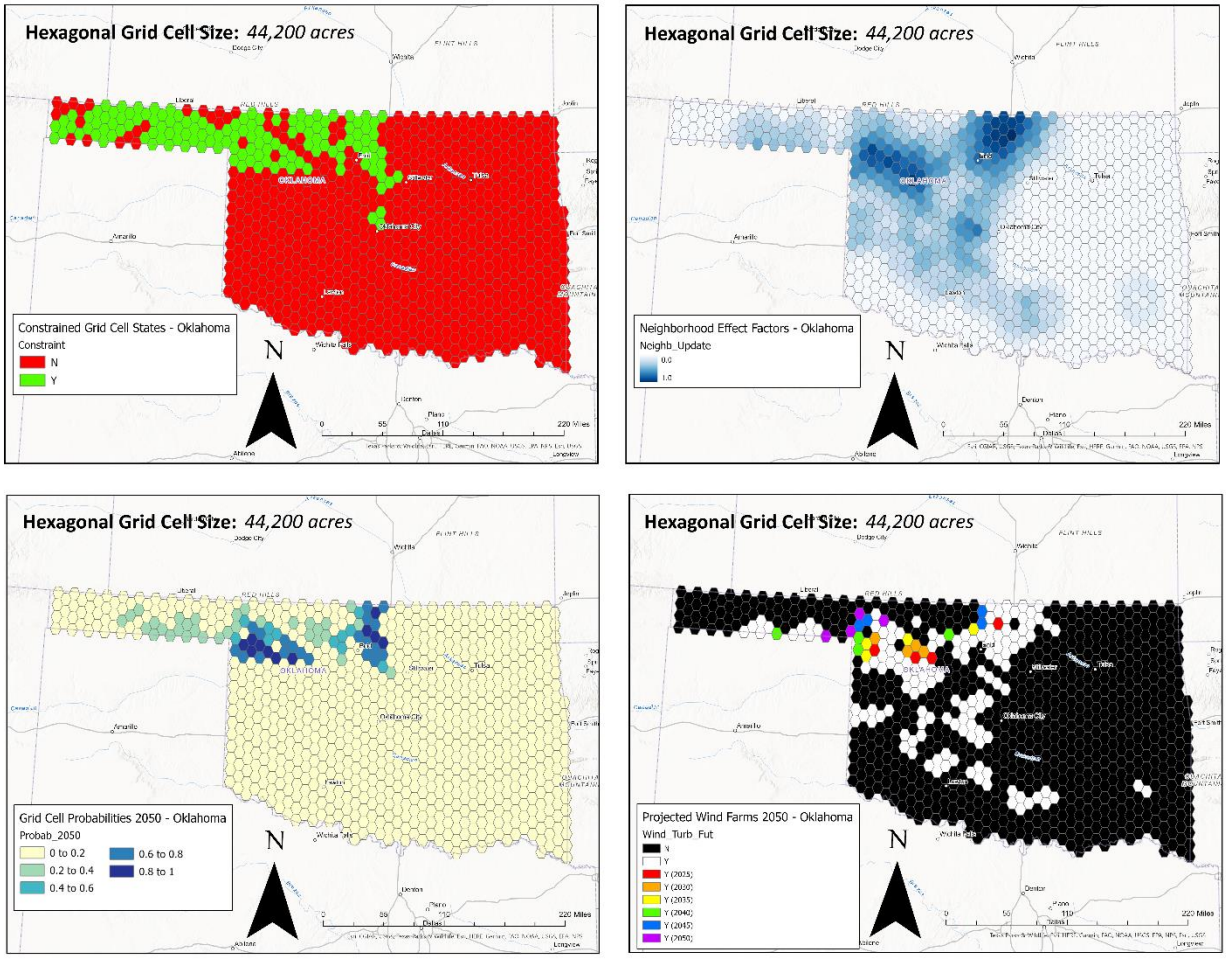


Figure 10: Example fields of the hexagonal gridded surface produced by WiFSS-LRCA's iterations out to 2050 over the state of Oklahoma. From top-left to bottom-right are the grid cells constrained for future wind energy development (*Constraint*), neighborhood effect factors by the final iteration (*Neighb_Update*), the constructed WiFSS surface (*Probab_2050*), and projected wind farm locations by the year 2050 (*Wind_Turb_Fut*). Hexagonal grid cell sizes are given on each map (see Table 6). Basemap from Esri [407].

Maps like those in Figure 10 can be produced by WiFSS-LRCA for any study area (U.S. state or CONUS), or grid cell resolution (defined in Table 6) of interest and are the primary output of this model. It is through creating these maps that this dissertation realizes its purposes (see Section 1.2), namely to project and illustrate wind farm siting futures to assist decision-makers and to show that models that incorporate cellular automata can be developed for spatial scales larger than in LRCA's previous applications.

3.5. Sensitivity Analysis.

3.5.1. Scenario Building and Sensitivity to Model Parameters.

The many components of WiFSS-LRCA (e.g., grid cell size, neighborhood size, predictors, etc.) mean that its projections may be sensitive to how these components are defined. Sensitivity analysis aims to "establish overall behavior of a system or model to the variation of a parameter" [442, p.720], which WiFSS-LRCA addresses partially through incorporation of scenarios that modify the fitted coefficients obtained from training the model's LR equation (see Section 3.3). Some of the predictors enlisted by WiFSS-LRCA are likely to change between now and 2050, meaning their influences on WiFSS are unlikely to be static, such as transmission line networks expanding to meet electricity demand [356], continued changes in overall attitudes toward renewable energy [23], and climate change's impacts on wind energy generation [430]. Hence, modifying these predictors' coefficients between iterations, and by extension their association with whether a grid cell contains a wind farm, is a means of representing this non-static influence. Urban growth applications of LRCA modeling often enlist scenarios that assume various future growth dynamics [66,70], which WiFSS-LRCA emulates by prescribing arbitrary percent changes to the coefficients in Equation 1 between timesteps. Ten scenarios constructed for WiFSS-LRCA are summarized in Table 9, with each scenario representing a group of predictors that have related effects on wind farm siting decisions. A default percent change of ±10% to each predictor coefficient is used in each scenario, unless otherwise specified, with signs derived from the expected ORs for each predictor (OR > 1 means positive sign, and vice versa; Tables 7a and 7b). Average Elevation (*Avg_Elevat*) and Proportion of Rugged Land (*Prop_Rugg*) are absent from these conceived scenarios, because their rate of change is negligible over a 30-year study period.

Scenario	Description	Predictor	% Change (+/-)
Changing Energy Economies	Older forms of energy generation age out, and a demand grows for green energy and green jobs.	Dens_15_19	-
		Farm_Year	+
		ISO_YN	+
		Plant_Year	+
		Type_15_19	+
		Unem_15_19	-
Climate Change	Temperature and wind speed increase, and bird and bat habitats are increasingly threatened.	Avg_Temp	+
		Avg_Wind	+
		Bat_Count	-
		Bird_Count	-
Demographic Changes	Demographics that are statistically more supportive of wind energy projects comprise a greater amount of local populations.	Avg_25	+
		Hisp_15_19	+
		Fem_15_19	+
		Whit_15_19	-
Nationwide	In model runs performed for the CONUS, predictors with effects at a nationwide level, such as legislation in effect, lobbyism, and land value are implemented as an extra scenario.	Cost_15_19	+
		Farm_15_19	-
		Foss_Lobbs	-
		Gree_Lobbs	+
		In_Tax_Cre	+
		Interconn	+
		Net_Meter	+
		Numb_Incen	+
		Numb_Pols	+
		Prop_15_19	-
		Renew_Port	+
		Renew_Targ	+
		Rep_Wins	-
		Tax_Prop	+
		Tax_Sale	+
Natural and Cultural Protection	Protection of land that is historically, culturally, or environmentally significant is prioritized as commercial wind energy development continues.	Critical	-
		Historical	-
		Nat_Parks	-
		Trib_Land	-
		Undev_Land	-
		Wild_Refug	-
New Infrastructure	Roads and transmission lines are built to support development of new commercial wind farms.	Near_Roads	+
		Near_Trans	+
Sociopolitical Landscape	Support for wind energy development among politicians and the electorate increases.	Dem_Wins	+
		supp_2018	+
Urban Protection	Wind energy development continues at a distance set far enough away from industrial and domestic activities.	Military	-
		Mining	-
		Near_Air	+
		Near_Hosp	+
		Near_Plant	+
		Near_Sch	+
Custom	A unique set of coefficient changes can be set by an end user.	Any	Any
Default	The coefficients of predictors remain constant, meaning projections are driven only by neighborhood effects.		

Table 9: Scenarios constructed to iteratively modify Equation 1's coefficients. Standard percent change is ±10%. The *Nationwide* scenario cannot be used in model runs over individual U.S. states. If the *Default* or *Custom* scenario is used, no other scenarios can be selected. Percent changes other than ±10% can be specified only in the *Custom* scenario.

In addition to modifying predictor coefficients, several of WiFSS-LRCA's other parameters are also customizable. Listed below are the four customizable parameters of WiFSS-LRCA and evidence from the literature of how each of them has been shown to influence CA-based models:

1. Sensitivity to grid cell size. Pan et al. [443] found that small grid cell sizes can cause land use change projected by CA to remain localized to existing patches of the same land type. Grid cell size modifications are limited to those listed in Table 6.

2. Sensitivity to neighborhood size. Wu et al.'s [444] study of CA-Markov modeling of land use change in China showed that a grid cell's neighborhood size more greatly impacts land use change potential than grid cell size does. Modification to n in Equation 5 alters WiFSS-LRCA's neighborhood size.

3. Sensitivity to constraints. By loosening or disabling the default constraints in Table 8, more grid cells become candidates to acquire wind farms in any model iteration, i.e., $Const_i = 1$ for more grid cells, as similarly noted by Li and Yeh [435].

4. Sensitivity to scenario setups. Allowing certain coefficients to change between iterations (i.e., selecting only some of the scenarios in Table 9), rather than none or all of them, would affect computed P_i values (and thus of $Prob_i^t$) in each iteration, and thus potentially affect which grid cells acquire wind farms first.

The value of integrating such sensitivity analysis into WiFSS-LRCA's design is twofold: 1) determination of which components of the model are the most consequential to projected commercial wind farm locations should they be modified; and 2) allowing end-users to perform model projections that meet their specifications and interests in projecting wind energy development (e.g., excluding certain constraints, selecting a larger wind farm capacity).

This dissertation will examine the consequences to WiFSS-LRCA's projected commercial wind farm locations of modifying each of these parameters, thus determining whether the sensitivities observed in the literature above also apply in the present context. Doing so is key to this dissertation's purpose of expanding the use of LRCA modeling to new and larger contexts (see Section 1.2), since existing sensitivity analyses of similar models have mostly been performed on smaller spatial scales (e.g., city, county) than a U.S. state or the CONUS [435,443,444]. A response to model parameters that is consistent with those in the literature would accredit a state or national-level scope for future studies using WiFSS-LRCA or similar models. Beyond visual inspection, WiFSS-LRCA computes metrics to aid its own sensitivity analysis. Firstly, Getis-Ord statistics [430] convey changes in the statistical significance ($p < 0.05$) of projected clusters of future wind farms produced by different parameter setups. For instance, disabling certain constraints may allow new commercial wind farms to be projected closer together, thus altering the cluster-like growth driven by $Neighb_i^t$. Secondly, WiFSS-LRCA computes a Quantity and Allocation Disagreement Index (QADI) to numerically represent differences in the number and location of grid cells that gain wind farms [445], as an alternative to the now-deprecated Kappa index [446]. QADI values are based on comparing projected wind farm locations against those produced using a "null" version of Equation 4 that excludes all predictors from P_i. As such, a QADI closer to 0 means lower disagreement, indicating P_i's lack of importance in Equation 4 as well as $Neighb_i^t$ effecting tighter future wind farm clusters. Answering this dissertation's research questions requires knowing the robustness of the supporting evidence, particularly when proposing geographical influences on WiFSS-LRCA's projections (see Section 1.2); sensitivity analysis is crucial in the composition of said evidence.

3.5.2. Setting Predictor Configurations.

Projections made by WiFSS-LRCA for a U.S. state or the CONUS are also likely to be sensitive to the predictors used to train and test its LR equation. WiFSS-LRCA therefore includes four predictor configurations to examine this sensitivity, while also facilitating bespoke projections by end-users interested in the effects of specific predictors on the model's output. The *Full* configuration incorporates all predictors into Equation 1 whose datasets satisfied the assumptions of logistic regression (see Section 3.3). Since Wind Speed is considered an important predictor in existing WiFSS studies, as it represents the energy resource itself [44,124,218], WiFSS-LRCA possesses two more configurations named *Wind_Only* (Wind Speed is the only predictor in Equation 1) and *No_Wind* (all predictors except for Wind Speed are used in Equation 1). The *Reduced* configuration is a refined set of predictors that maximizes the number of correctly predicted grid cell states and true to false positive ratio when assessing WiFSS-LRCA's performance during its validation step (see Section 3.3.1). The *Reduced* configuration is the most complex of the four to construct, created by WiFSS-LRCA as part of its calibration and validation in the following steps:

- In each of the 30 repeats of its calibration, WiFSS-LRCA samples all predictors with replacement, with each sample removing the training grid cell data for a different predictor, *k*. As such, WiFSS-LRCA assesses its goodness-of-fit and log-likelihood ratio (see Equation 2) *k* times, with each assessment yielding coefficients, Y_1 to Y_{k-1}.
- WiFSS-LRCA compares the log-likelihood ratios obtained used each sampled set of $k-1$ predictors against the ratio obtained using the *Full* predictor configuration. This comparison counts the number of times (out of 30 calibration repeats) that the *Full* configuration yields the greater median log-likelihood ratio. In essence, WiFSS-LRCA counts how many times its

goodness-of-fit is worsened by removing a given predictor. This approach of refining model predictors to obtain a median worsened model performance is based on log-likelihood quantization in signal processing studies [447].

- WiFSS-LRCA also counts the number of times that this worsened model performance is significant, assuming a chi-square distribution and a stopping criterion of $p < 0.5$ [448] to determine said significance. WiFSS-LRCA then orders the predictors by how consequential each one's removal is to the model's goodness-of-fit, based on the number of times each predictor's removal significantly worsened model performance.
- Finally, in the validation 30 repeats, WiFSS-LRCA creates confusion matrices using the testing grid cell data for all sets of predictors, starting from the predictor whose removal was the most consequential and adding predictors consecutively. The predictor set that maximizes WiFSS-LRCA's median correctly predicted grid cell states, based on the median true to false positive ratios yielded by the confusion matrices, defines the *Reduced* configuration.

There are two benefits to constructing the *Reduced* predictor configuration, the first being that its removal of nonessential predictors reduces the risk of overfitting WiFSS-LRCA to the aggregated predictor data. Having too many degrees of freedom, i.e., too many predictors, makes effects of individual predictors on model outputs difficult to distinguish, while also exaggerating data variations [192]. Secondly, this removal of predictors allows those most pertinent to WiFSS over a selected study area to be identified; the *Reduced* configuration may not retain Wind Speed when running WiFSS-LRCA over a U.S. state that generally has weaker winds. As with assessing sensitivity to selected scenarios and model parameters, this dissertation examines impacts of selecting specific predictors as part of its sensitivity analysis, realized through comparing model projections produced using these four predictor configurations.

Chapter 4: Results – Prediction Using the LR Equation.

This chapter covers the outcomes of training and testing WiFSS-LRCA's Logistic Regression equation, with the chapter focusing specifically on this equation for two reasons. Firstly, given this dissertation's purpose of projecting wind farm siting futures that can be explained geographically (see Section 1.2), doing so requires determining first that the model can capture present wind farm locations. This idea of evaluating WiFSS-LRCA's performance against true wind farm locations is analogous to comparing hindcast climate model runs against historical observations before using said climate model for projection [449]. A model that produces limited error or uncertainty when compared to observations can be regarded as better at capturing present system behavior [450], making the model's projections of future behavior more trustworthy [451]. It is thus important to determine that this model's LR equation can correctly predict present wind farm locations before asserting possible reasons for these predictions and before allowing the CA component to project locations for future wind energy development. Secondly, this dissertation's second research question concerns identifying locations across the CONUS that are currently the most suitable for building commercial wind farms. Identifying suitable locations requires analysis only of WiFSS-LRCA's output in its first iteration, i.e., that of the trained and tested LR equation. Focusing only on the LR equation's predictions allows for questions to be answered about the regions of the CONUS in which WiFSS-LRCA can most accurately capture present wind farm locations, which predictors are strongly associated with these predictions, and how predictions are impacted by chosen grid cell resolutions and predictor configurations. Using a selection of U.S. states and the CONUS as case examples, this chapter will address all these questions using figures and tables produced by WiFSS-LRCA in its first iteration. An example console output from running WiFSS-LRCA's LR script is in Appendix A3.

4.1. Results from Calibrating the Logistic Regression Equation.

4.1.1. Goodness-of-Fit.

According to Smith and McKenna [452], an LR equation has a good fit if its classification of a binary outcome statistically significantly improves (assuming a chi-square distribution [418]) over a null model. Across all four predictor configurations and most study areas, WiFSS-LRCA's goodness-of-fit showed consistent improvements over the null model, as summarized in Table 10. The log-likelihood ratio (λ, Equation 2) was frequently statistically significant when using larger sets of predictors, particularly the *Full* configuration, and for study areas with more grid cells. The least improvement in goodness-of-fit occurred with the *Wind_Only* configuration, especially in U.S. states with lower wind speeds, such as Maine and New Hampshire [321]. Although an R^2_{MFA} above 0.2 is a recommended standard for judging a good fit [453], Table 10 shows that a negative R^2_{MFA} can occur despite λ being statistically significant. Negative values may result from large log-likelihood (LL) scores (e.g., North Dakota), or from the adjustment for k predictors (Equation 3) having a greater effect on small differences in LL scores (e.g., Maine). Generally, WiFSS-LRCA's goodness-of-fit was most robust under the *Full* and *Reduced* configurations.

4.1.2. Observed and Expected Odds Ratios.

The coefficients obtained from maximizing WiFSS-LRCA's log-likelihood ratio convey the strength of association between each predictor and the binary existence of commercial wind farms [422]. Each coefficient's exponent yields an Odds Ratio (OR), as shown from a CONUS model run of WiFSS-LRCA (Figure 11), with green bars signifying a positive association (OR > 1), and red bars a negative association (OR < 1). The association of Wind Speed (*Avg_Wind*) with a training grid cell containing a wind farm was the strongest, possessing a median OR of

State	Predictor Configuration	Median LL, (Trained - Null)	LL Ratio, λ	Times Improved (Sig.)	Median Pseudo-R^2, R^2_{MFA}
Maine	*Full*	31.22	62.44*	30 (30)	-0.02
	No_Wind	33.56	67.11*	30 (30)	-0.03
	Wind_Only	1.62	3.25	30 (3)	0.00
	Reduced	26.7	53.40*	30 (30)	-0.02
New Hampshire	*Full*	10.2	20.39	30 (1)	0.11
	No_Wind	8.11	16.22	30 (0)	0.12
	Wind_Only	0.49	0.99	30 (3)	0.02
	Reduced	1.7	3.39	20 (3)	0.03
North Dakota	*Full*	100.27	200.54*	30 (30)	-0.36
	No_Wind	77.95	155.89*	30 (30)	-0.24
	Wind_Only	49.67	99.34*	30 (30)	-0.25
	Reduced	76.91	153.82*	30 (30)	-0.35
Oklahoma	*Full*	190.09	380.18*	30 (30)	0.49
	No_Wind	112	223.99*	30 (30)	0.25
	Wind_Only	123.15	246.30*	30 (30)	0.36
	Reduced	140.77	281.54*	30 (30)	0.40

Table 10: Results from calibrating WiFSS-LRCA's logistic regression equation over four states (grid cell size = 13,000 acres, see Table 6), using all four predictor configurations. Median log-likelihood (LL) and Pseudo-R^2 values were obtained from the 30 calibration repeats. "Times Improved" shows the number of times the trained model's LL score exceeded the null model's score, and the number of times the difference was statistically significant ($p < 0.05$).
* = Statistically significant LL Ratio ($p < 0.05$).

9.74. Other strongly positively associated predictors included Total Number of Statewide Legislative Pieces (*Numb_Pols;* OR = 2.64), Size of RPS Target (*Renew_Targ*; OR = 1.86), and Percent White Population (*Whit_15_19*: OR = 1.42). Important negatively associated predictors included Nearest Major Transmission Line (*Near_Trans*; OR = 0.36), Proportion of Undevelopable Land (*Undev_Land*; OR = 0.38), and Green Lobbies (*Gree_Lobbs*; OR = 0.49). Five of these seven predictors (except for *Gree_Lobbs* and *Whit_15_19*) possessed OR values that agreed with expectations presented in Tables 7a and 7b; 26 (55.3%) of all 47 predictors

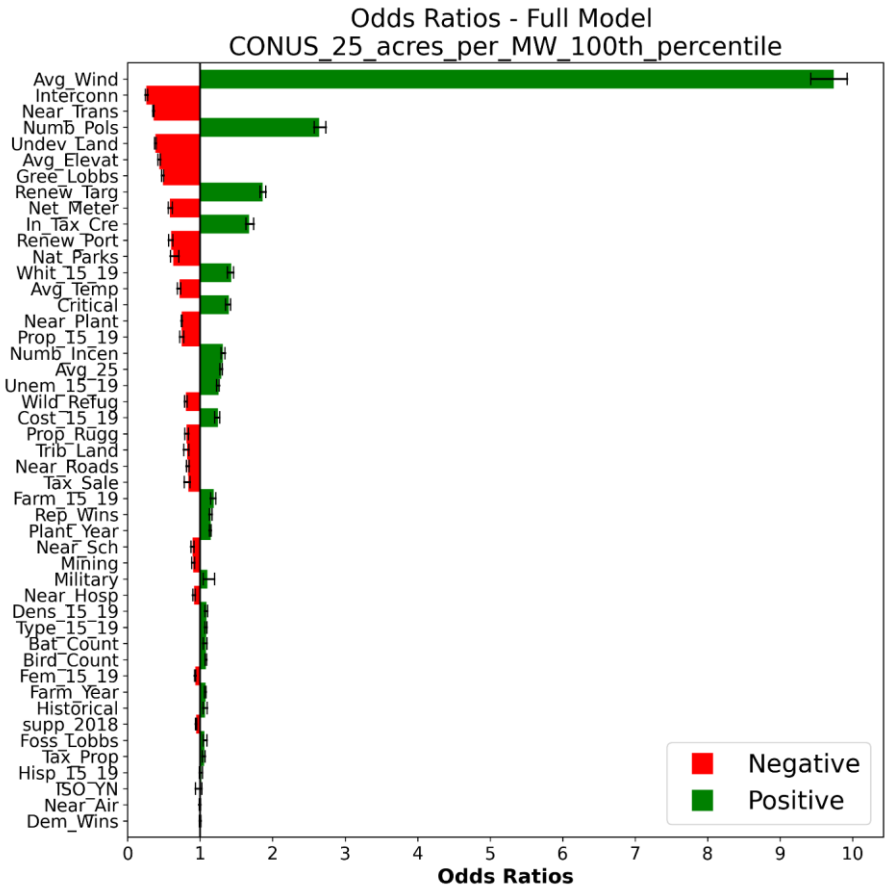


Figure 11: Odds Ratio (OR) chart produced from running WiFSS-LRCA over the CONUS using the *Full* predictor configuration (grid cell size = 13,000 acres, see Table 6). Each bar presents a positive (green, OR = 1) or negative (red, OR = 1) median OR for each predictor, with the error bars showing the lower quartile and upper quartile ORs obtained from the 30 repeats of the model calibration step.

agreed with these expectations. However, over-interpretation of OR values should be avoided when associations are weak (i.e., closer to one) [95], and because of confounding predictors potentially influencing each other and/or the dependent variable [96] resulting in spurious relationships that cannot be easily explained. Figure 11 typifies runs of WiFSS-LRCA over the CONUS, such that predictors like *Avg_Wind*, *Near_Trans*, *Undev_Land*, and *Numb_Pols* are frequently among the predictors most strongly associated with whether or not training grid cells contain commercial wind farms.

Running WiFSS-LRCA over different states (or locations), such as a model run over Indiana that produced Figure 12, shows that these computed ORs are sensitive both to the selected predictor configuration and spatial scale, with several notable results:

- Unlike for the CONUS, *Avg_Wind* was not the most positively associated predictor when running WiFSS-LRCA over Indiana. Public Support for RPS (*supp_2018*; OR = 7.46), Employment Type by County *(Type_15_19*; OR = 5.20), *Whit_15_19* (OR = 3.40), and Critical Habitats (*Critical;* OR = 2.12) were all more positively associated than *Avg_Wind* (OR = 1.87) in WiFSS-LR's *Full* (top-left) predictor configuration.
- The OR of *Avg_Wind* was much greater in WiFSS-LRCA's *Wind_Only* (bottom-left) configuration (OR = 6.00). This and the *Full* configuration suggest that Wind Speed is a less important predictor for wind farm siting decisions in Indiana than across the CONUS, a difference perhaps attributable to a state versus nationwide spatial scale and/or the smaller number of predictors used in state-level model runs (see Tables 5a and 5b)
- WiFSS-LRCA's *Reduced* (bottom-right) configuration retained five of the 32 predictors available for state-level runs (*Undev_Land*, *supp_2018, Avg_25, Near_Air, Near_Plant*). These five predictors maximized WiFSS-LRCA's ability to correctly classify testing grid cells as (not) containing a wind farm. However, *Avg_25* lacked the expected OR listed in Table 7b, unlike the other four predictors.

The removal of wind speed for the *No_Wind* configuration (top-right) did little to the other predictors' ORs, unlike the reduction of predictors to five when applying the *Reduced* configuration to Indiana. Comparing these two configurations shows large changes in median ORs, especially for *supp_2018* and *Undev_Land*, with the former decreasing from 7.56 to 4.31

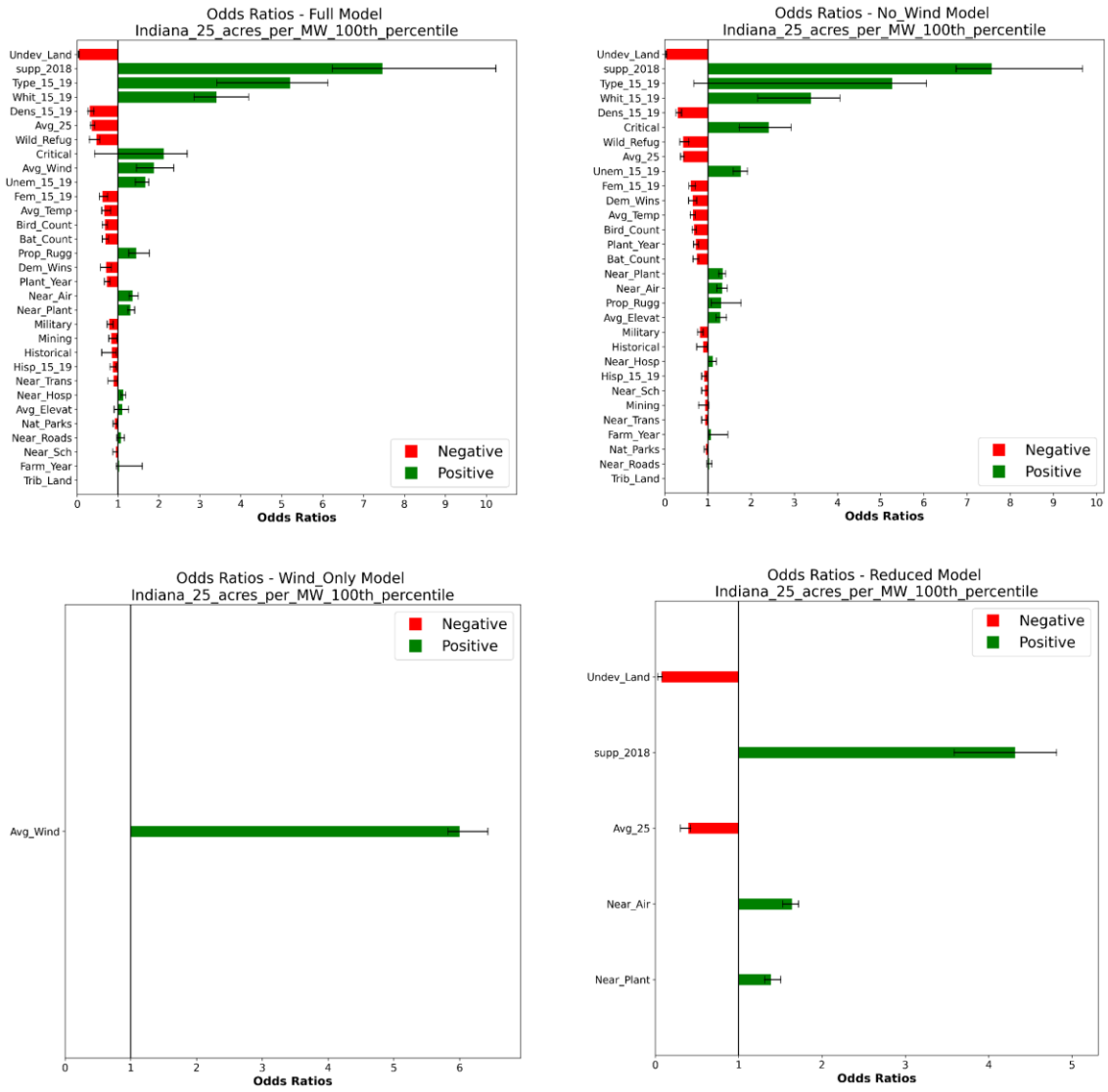


Figure 12: Same as Figure 11, but for runs of WiFSS-LRCA over Indiana using all four predictor configurations. From top-left to bottom-right are the results from using the *Full*, *No_Wind*, *Wind_Only*, and *Reduced* configurations. Note that each of these charts possesses a different x-axis.

(1.75 times weaker association) and the latter increasing from 0.03 to 0.07 (2.33 times weaker association). WiFSS-LRCA's computed ORs were thus sensitive to the number of predictors, the selected spatial scale, and somewhat sensitive to the specific predictors added/removed.

4.1.3. Effects of Randomness on the "Reduced" Predictor Configuration.

Although WiFSS-LRCA mitigates impacts of random grid cell sampling by repeating its calibration 30 times [420], and by stratifying its training and testing data [416], the *Reduced* predictor configuration was still sensitive to this randomness. The combination of predictors that maximized the number of correctly classified grid cell states (see Section 3.5.2) changed greatly in each training grid cell sample, meaning that separate runs of WiFSS-LRCA retained different sets (and different numbers) of predictors in their *Reduced* configurations. Table 11 illustrates this sensitivity to randomness based on four separate WiFSS-LRCA runs over California (grid cell size = 17,000 acres). Despite the number of predictors ranging from three to 21 (out of 32 predictors), each configuration produced a similar median accuracy, with 89-93% of California's testing grid cells correctly classified as (not) containing a wind farm. Moreover, the predictors most consequential to model performance were consistent between separate WiFSS-LRCA runs, with removal of *Avg_Wind*, *Whit_15_19*, Independent System Operators (*ISO_YN*), and Tribal Lands (*Trib_Land*) consistently worsening its goodness-of-fit. As such, while the *Reduced* configuration's predictor combinations varied widely in separate model runs, impacts on WiFSS-LRCA's performance were minimal. The *Reduced* configuration is therefore robust in its intent to retain only the predictors most pertinent to wind farm siting decisions.

4.2. Results from Validating the Logistic Regression Equation.

4.2.1. Receiver Operating Characteristics and Sensitivity to Resolution.

Classification of binary events by an LR equation is better than random chance if the computed AUC statistic for these events is greater than 0.5 [454], which WiFSS-LRCA assesses by constructing ROC curves. Figure 13 shows ROC curves produced from 30 repeats of WiFSS-LRCA's validation over Texas (grid cell size = 23,400 acres) for all four predictor configurations.

Predictors	Reduced_Fit	Stop_Criterion
Avg_Wind	30	30
Whit_15_19	30	28
Near_Plant	30	8
Near_Air	28	10
Trib_Land	28	7
Undev_Land	27	14
Military	27	14
Unem_15_19	27	13
Mining	26	15
ISO_YN	26	15
Dem_Wins	26	13
Prop_Rugg	26	12
Bird_Count	25	9
Farm_Year	25	9
Near_Hosp	24	15
Fem_15_19	24	13
Avg_25	24	10
Wild_Refug	24	10
Near_Roads	24	6
Nat_Parks	24	5
Plant_Year	23	13
Median Prediction Accuracy:	0.89	
Median True-False Positive Ratio:	0.09	

Predictors	Reduced_Fit	Stop_Criterion
Avg_Wind	30	30
Whit_15_19	24	14
ISO_YN	21	7
Trib_Land	21	6
Near_Hosp	21	4
Undev_Land	20	6
Unem_15_19	20	6
Dem_Wins	20	4
Historical	19	7
Nat_Parks	19	5
Fem_15_19	18	6
Farm_Year	18	3
Plant_Year	18	2
Bat_Count	18	1
Type_15_19	18	1
Near_Air	17	3
supp_2018	17	3
Critical	17	3
Median Prediction Accuracy:	0.93	
Median True-False Positive Ratio:	0.13	

Predictors	Reduced_Fit	Stop_Criterion
Avg_Wind	30	30
Whit_15_19	28	14
Mining	25	7
Median Prediction Accuracy:	0.91	
Median True-False Positive Ratio:	0.09	

Predictors	Reduced_Fit	Stop_Criterion
Avg_Wind	30	30
Whit_15_19	28	7
Dem_Wins	22	3
ISO_YN	20	6
Trib_Land	20	5
Median Prediction Accuracy:	0.90	
Median True-False Positive Ratio:	0.08	

Table 11: Results from identifying the predictors to be used in the *Reduced* predictor configuration of four separate model runs over California. Each table shows the number of times the removal of each predictor worsened WiFSS-LRCA's goodness-of-fit ("Reduced_Fit"), the number of times this worsening was significant based on a $p < 0.5$ stopping criterion [447], and the median proportion of testing grid cells that were correctly classified based on 30 repeats with this set of predictors.
Grid cell size = 17,000 acres.

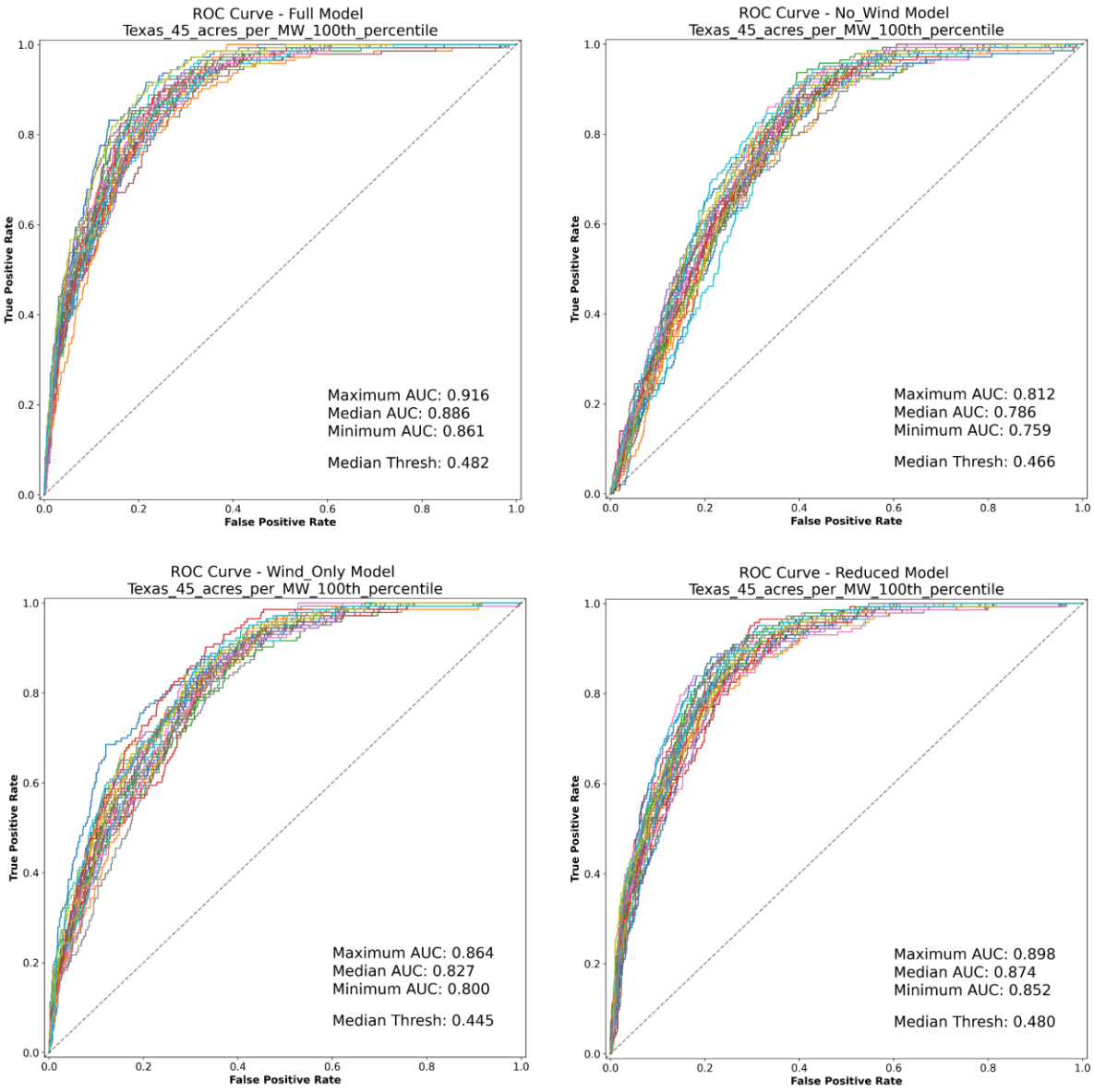


Figure 13: Receiver Operating Characteristic (ROC) curves produced from running WiFSS-LRCA over Texas. Each colored line represents one of the 30 repeats of the model's validation step. The diagonal (gray) line denotes the ROC curve should the correct classification of testing grid cells be equal in probability to random chance (Area Under Curve (AUC) = 0.5). From top-left to bottom-right are the results from using the *Full*, *No_Wind*, *Wind_Only*, and *Reduced* predictor configurations. Grid cell size = 23,400 acres.

These curves came closer to the top-left corner (0,1), and thus possessed a greater AUC [455], when enlisting the *Full* (Median AUC = 0.886) or *Reduced* (Median AUC = 0.874) configuration. WiFSS-LRCA thus seemed to distinguish between testing grid cells with and without wind farms more successfully when using a complete (*Full*) or more refined (*Reduced*; nine out of 32 predictors) predictor set, much like how WiFSS-LRCA's goodness-of-fit generally maximized under these two configurations (see Section 4.1.1). The median threshold probability for correctly classifying grid cells was similarly sensitive to predictor configuration, with this threshold maximizing under the *Full* configuration (0.482) and being lowest under the *Wind_Only* configuration (0.445). Across most U.S. states, this threshold probability was similarly highest when WiFSS-LRCA enlisted the *Full* or *Reduced* configuration, suggesting that this model commits more Type 1 (false positive) errors (i.e., incorrectly classifies more grid cells as containing wind farms) when using an incomplete predictor set.

Constructing ROC curves also revealed WiFSS-LRCA's lower predictive accuracy when run over U.S. states with fewer commercial wind farms. Construction of ROC curves from validating WiFSS-LRCA over Vermont (Figure 14; *Reduced* predictor configuration) at two different grid cell resolutions (left = 23,400 acres; right = 1,350 acres) presented two key results:

- The lack of grid cells containing wind farms in the testing dataset increases the risk of WiFSS-LRCA's 30 validation repeats not outperforming random chance (i.e., AUC < 0.5). At the lower grid cell resolution (Figure 14, left), having only two testing grid cells that contained wind farms resulted in some repeated validations vastly underperforming, with a Minimum AUC of 0.119 observed.
- In study areas with a smaller number of commercial wind farms, such as Vermont, the AUC statistic becomes more sensitive to changes in grid cell size. Indeed, the higher grid cell

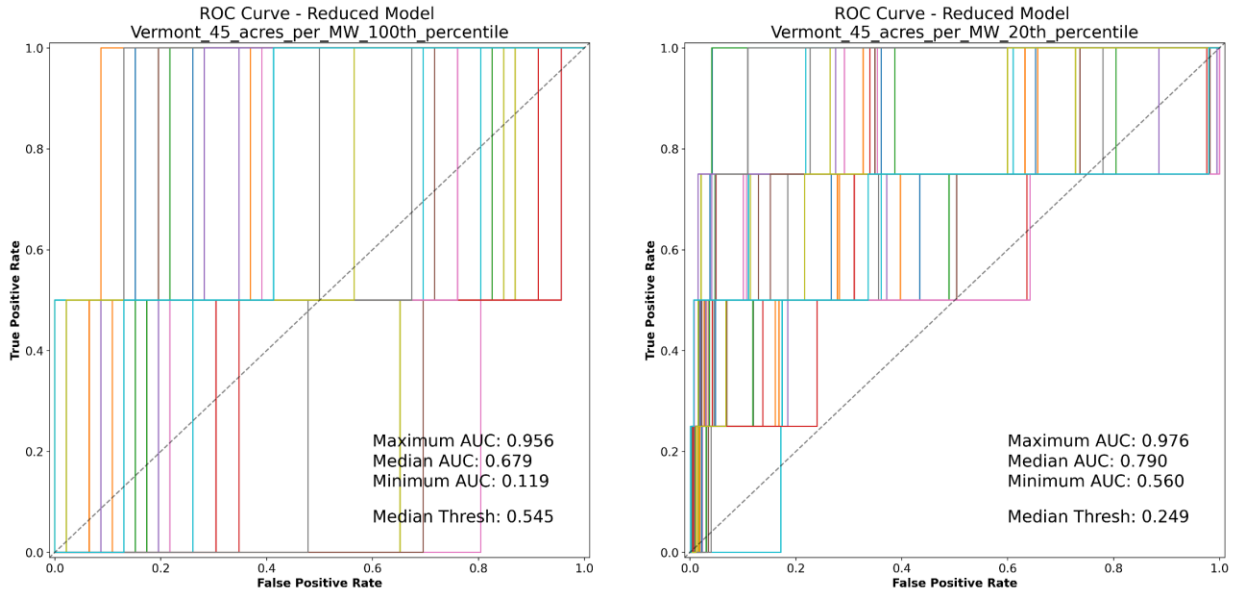


Figure 14: Same as Figure 13, but from runs of WiFSS-LRCA over Vermont using the *Reduced* predictor configuration only. Grid cell size, left = 23,400 acres. Grid cell size, right = 1,350 acres.

resolution in Figure 14 (right) resulted in the number of testing grid cells that contain wind farms increasing from two to four, accompanied by a greater Area Under Curve statistic (Median AUC = 0.790), along with no incidents of performing worse than random chance (Minimum AUC = 0.560). Furthermore, this sensitivity of AUC statistics to grid cell size was lower in states possessing more commercial wind farms, such as Texas.

Validating WiFSS-LRCA using ROCs is limited by these curves not conveying which testing grid cells are reclassified at different threshold probabilities [456]. Nevertheless, WiFSS-LRCA's sensitivity to predictor configuration, grid cell size, and the number of existing commercial wind farms is apparent. However, the sensitivity to grid cell size should not be overstated, since this sensitivity is partially an artifact of a selected study area having few commercial wind farms. Low prevalence of a binary outcome (i.e., a lack of testing grid cells available to predict as

ontaining wind farms) is a known limitation of using ROCs to evaluate a predictive model's performance [457], as highlighted by Figure 14.

4.2.2. Using Confusion Matrices to Summarize Type 1 and Type 2 Errors.

Confusion matrices constructed by WiFSS-LRCA illustrate the actual number of testing grid cells correctly and incorrectly classified as containing a commercial wind farm. This validation enlists the median threshold probability computed from the prior ROC analysis, such that grid cells with a P_i above this threshold value and do contain a wind farm [73] are classified as true positive, increasing the true positive count by one. Similarly, grid cells with a P_i below the median probability threshold that do not contain a wind farm are classified as true negative, and so on for false positive (Type 1 error) and false negative (Type 2 error) classifications [458]. WiFSS-LRCA consistently achieves high rates of correct grid cell classification across most states and grid cell sizes. Figure 15 shows that, for all four predictor configurations, more than 75% of grid cells in a model run over New Mexico (grid cell size = 33,800 acres) were classified as true positive (bottom-right quadrant) or true negative (top-left quadrant), with the biggest difference between configurations being the number of Type 1 errors (top-right quadrant). In both state-level and CONUS model runs, the fewest Type 1 errors consistently occurred when using the *Full* or *Reduced* predictor configuration, again suggesting that more robust assessments using WiFSS-LRCA came from using a complete or more refined predictor set. WiFSS-LRCA was also not susceptible to frequent Type 2 errors, as evidenced by the small numbers of false negative testing grid cells (bottom-left quadrant).

Much like the ROCs made by WiFSS-LRCA, the confusion matrices revealed the model's sensitivity to selected grid cell size. Comparing model runs using the *Full* configuration over

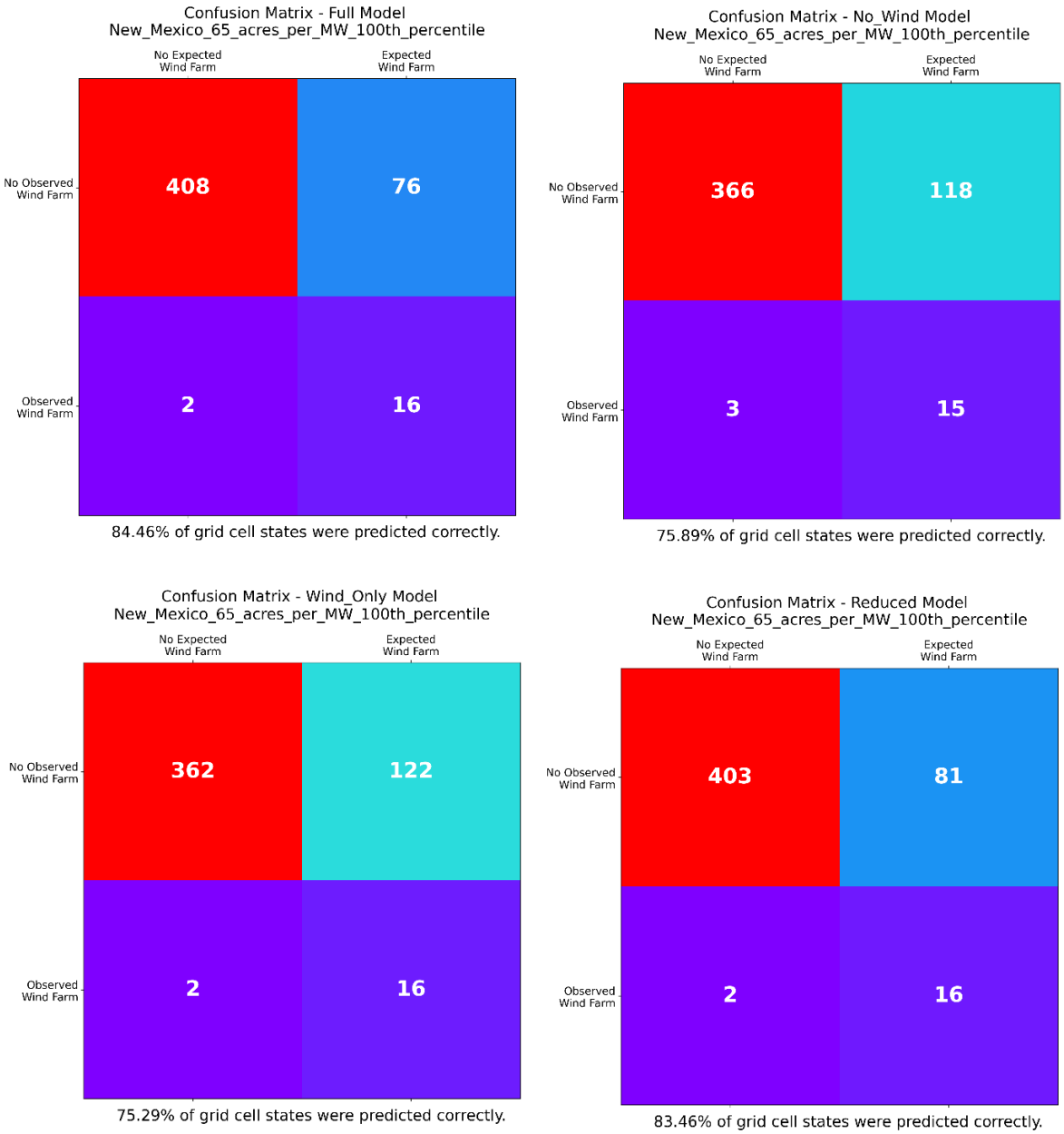


Figure 15: The median confusion matrices produced from running WiFSS-LRCA over New Mexico, produced by 30 repeats of the model's validation step. The color ramp from purple to red designates the relative number of testing grid cells in each quadrant. From top-left to bottom-right are the results from using the *Full*, *No_Wind*, *Wind_Only*, and *Reduced* predictor configurations. Grid cell size = 33,800 acres.

New York (Figure 16), the percentage of correctly predicted testing grid cell states was 64.83% at the 33,800-acre resolution (left), compared to 76.22% at the 9,750-acre resolution (right). As with Figure 15, the biggest difference between these matrices was the smaller proportion of Type 1 errors at the higher grid cell resolution. This result suggests that WiFSS-LRCA can correctly identify a larger number of existing commercial wind farm locations when enlisting a smaller grid cell size. As such, WiFSS-LRCA's practical applications may be more suited to predicting sites for smaller commercial wind farm projects, rather than wind farms comparable to the highest-capacity wind farms in the CONUS. The lack of Type 2 errors in Figures 15 and 16 means that WiFSS-LRCA rarely misclassifies grid cells that do contain a commercial wind farm, a result that pervades across study areas, grid cell sizes, and predictor configurations. The high number of true positives versus false negatives therefore suggests that predictors with strong associations (Figures 11 and 12) are indeed important for identifying and siting commercial wind farms, which Yun [459] similarly observed in their LR-based study of algal blooms.

4.3. Applying the LR Equation to All Grid Cells.

4.3.1. Boxplot Construction.

After having calibrated the equation's coefficients (see Section 4.1) and having validated its ability to correctly classify grid cell states (see Section 4.2), WiFSS-LRCA's logistic regression equation is then ready for application to all grid cells in the selected study area to create a WiFSS surface. WiFSS-LRCA first constructs boxplots using its LR equation to assess model performance across all grid cells, as in the boxplots produced by a CONUS model run in Figure 17 (*Full* configuration, grid cell size = 44,200 acres). Each of the four boxplots illustrates the range of probabilities of the grid cells within each classification (true/false positive/negative), compared to the median threshold probability derived from the ROC curves constructed by this

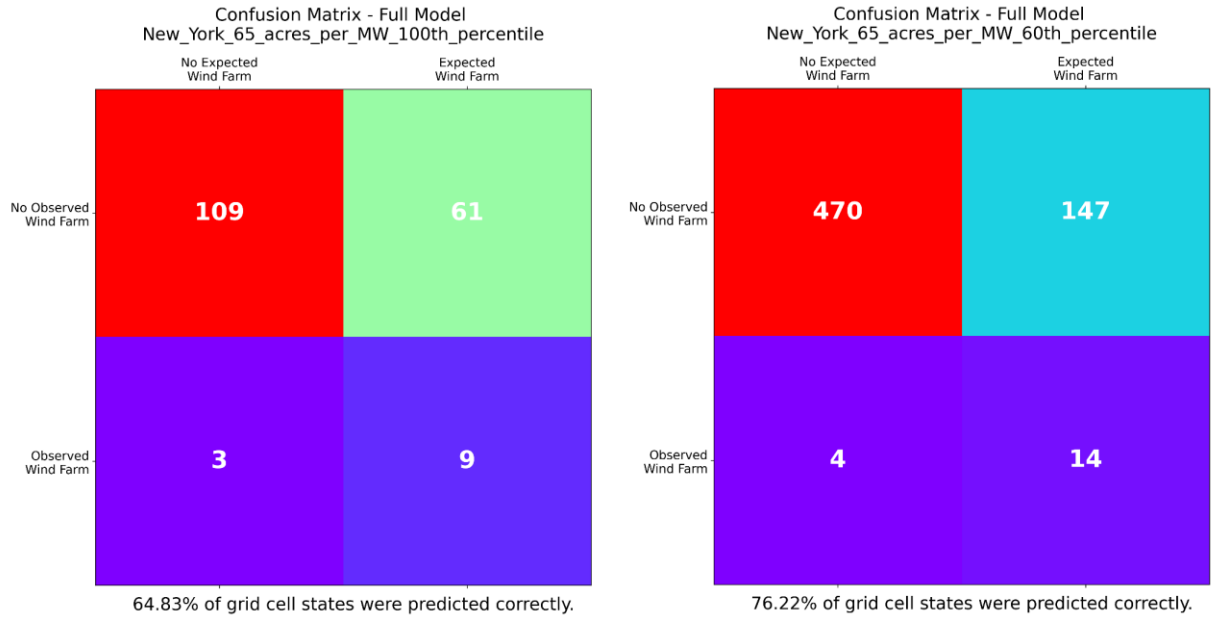


Figure 16: Same as Figure 15, but from runs of WiFSS-LRCA over New York using the *Full* predictor configuration only. Grid cell size, left = 33,800 acres. Grid cell size, right = 9,750 acres.

same model run. In this example, grid cells classified as true positive have a greater median probability of containing a commercial wind farm (0.827) than those classified as false positive (0.716). Similarly, true negative grid cells possessed the lowest median probability of wind farm occurrence (0.041), being much lower than that for false negative grid cells (0.355). If WiFSS-LRCA can make accurate predictions of where wind farms are currently located across the CONUS, then grid cells with the greatest P_i values should be those that contain wind farms, and vice versa. Figure 17 shows that of the 2055 grid cells in this configuration that contain a wind farm across the CONUS (true positive plus false negative), 1725 (83.9%) of them were classified as true positive, with 33,102 grid cells that do not contain a wind farm (74.9%) similarly classified as true negative (rather than false positive). Such high rates of correct classification

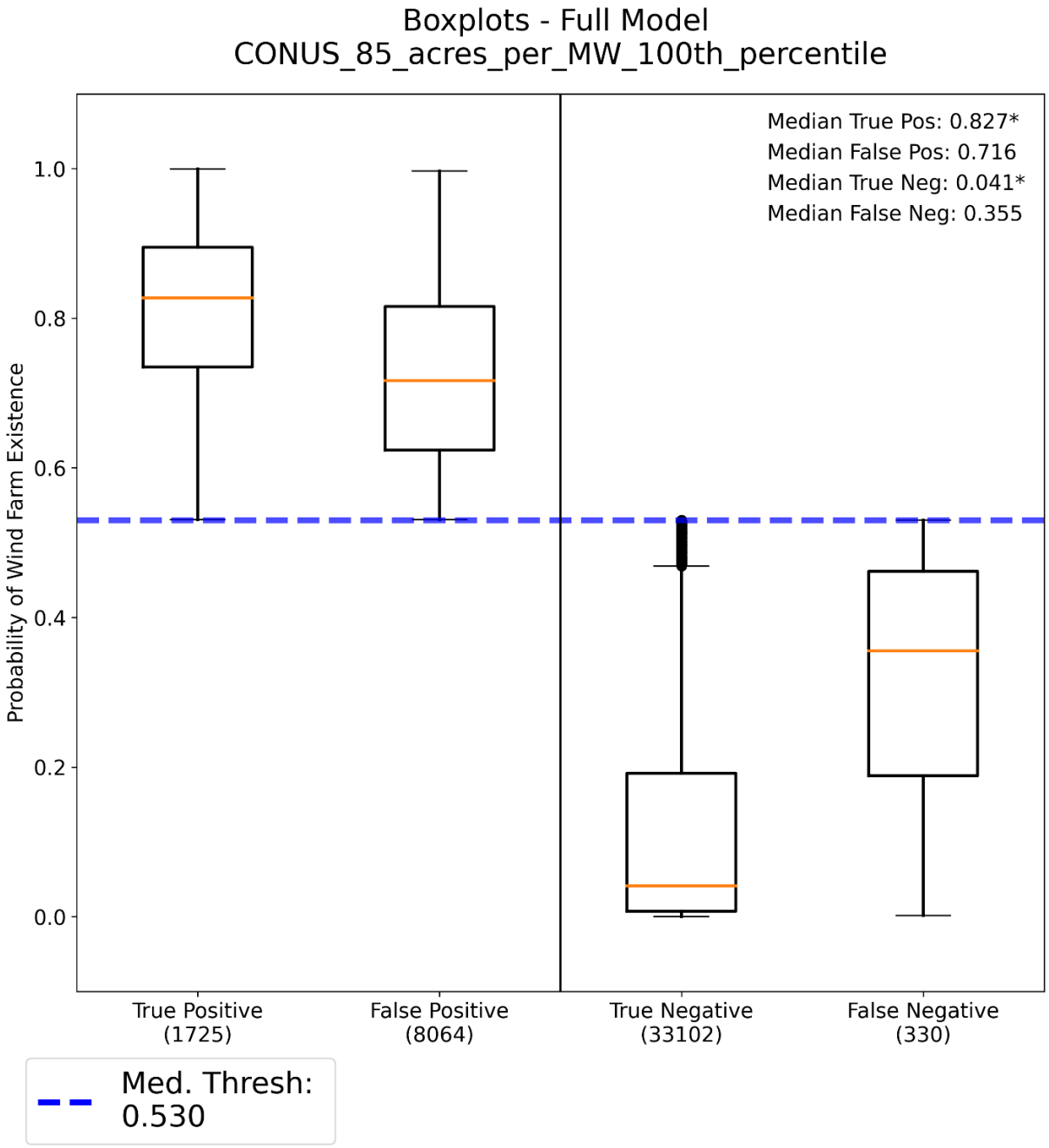


Figure 17: Boxplots of the probability of all grid cells in the CONUS study area (*Full* configuration, grid cell size = 44,200 acres) containing a commercial wind farm, with one boxplot for each of the four grid cell classifications. Orange lines indicate median probabilities. The blue dashed line indicates the median probability threshold derived from WiFSS-LRCA's validation step, against which grid cell probabilities were classified. Asterisks against the True Positive and True Negative probabilities designate their statistical significance based on a Mann-Whitney U-test ($p < 0.05$).

mean that this trained and tested LR equation can make mostly accurate predictions of commercial wind farm locations based on computed P_i values, making the constructed WiFSS surface more trustworthy.

WiFSS-LRCA checks the robustness of these computed probabilities by conducting a Mann-Whitney U-test [430] on two pairs of classifications: true positive and false positive; true negative and false negative. Doing so ensures that the classification of grid cells as (not) containing commercial wind farms is not random, based on a rank order comparison of the probabilities that comprise each pair of boxplots. The asterisked median probabilities in Figure 17 indicate statistically significant ($p < 0.05$) differences in rank order, the reason for which can be suggested based on the computed ORs for each predictor in the *Full* configuration. Figure 11 shows that predicted grid cell states over the CONUS (that use a *Full* predictor configuration) are strongly associated with aggregated data values for *Avg_Wind, Near_Trans, Numb_Pols, Undev_Land*, *Renew_Targ*, and several others. Based on these predictors' associations, grid cells classified as true positive are thus more likely to possess greater wind speeds, be closer to transmission lines, benefit from many renewable energy policies, not be faced with competing land use, and be located in areas with more ambitious RPS targets, respectively. False positive grid cells may possess all these qualities as well, but do not currently contain a commercial wind farm. While interpreting a Mann-Whitney U-test in this manner provides geographical context to WiFSS-LRCA's predictions, there are two caveats to doing so. Firstly, Mann-Whitney U-tests are less robust for boxplots composed of fewer than 30 data points [460], which mostly affects false negative boxplots to WiFSS-LRCA's infrequent Type 2 errors (see Section 4.2.2). Secondly, interpreting ORs is complicated by the associations produced by WiFSS-LRCA not always agreeing with the expectations laid out in Tables 7a and 7b, as well as ORs being strictly non-

causal depictions of the link between predictors and the outcome of interest [422]. Nonetheless, WiFSS-LRCA is proficient at computing probabilities that reflect both geographical context and the locations of present commercial wind farms across the CONUS.

4.3.2. Interpreting the Wind Farm Site Suitability Surface.

The final step of WiFSS-LRCA's first iteration is application of all probabilities to the hexagonal grid, thus generating a map that depicts how the suitability for commercial wind farm occurrence varies in space, i.e., a WiFSS surface. Figure 18 shows WiFSS surfaces generated by running WiFSS-LRCA over Wisconsin (grid cell size = 5,000 acres) for all four predictor configurations, which highlight the impacts made on the final WiFSS surface by an end-user's predictor choices. All four configurations predicted high suitability for commercial wind energy development on the state's southern and eastern sides, areas of Wisconsin with dense road and transmission line networks [325,326], fewer protected natural areas [317,318], and higher average wind speed due to the proximity of Lake Michigan [461]. The importance of wind speed for predicting WiFSS over Michigan stands out when using the *Wind_Only* configuration (bottom-left), given the concentration of high probabilities in the Green Bay area. Interestingly, the *Wind_Only* configuration also resulted in WiFSS-LRCA predicting higher suitability along Wisconsin's northern border than under the other three configurations, an area of the state with higher average wind speeds than to the south [321]. The *No_Wind* configuration (top-right) similarly resulted in unique high-probability regions, with one such region in Central Wisconsin that was less pronounced under the *Full* (top-left) and *Reduced* (bottom-right) configurations. The ORs generated alongside Figure 18 showed that these predictions were most strongly associated with Wind Speed and Military Installations (*Military*), with very few of the latter existing across Central Wisconsin [315], perhaps explaining the results under the *No_Wind* configuration.

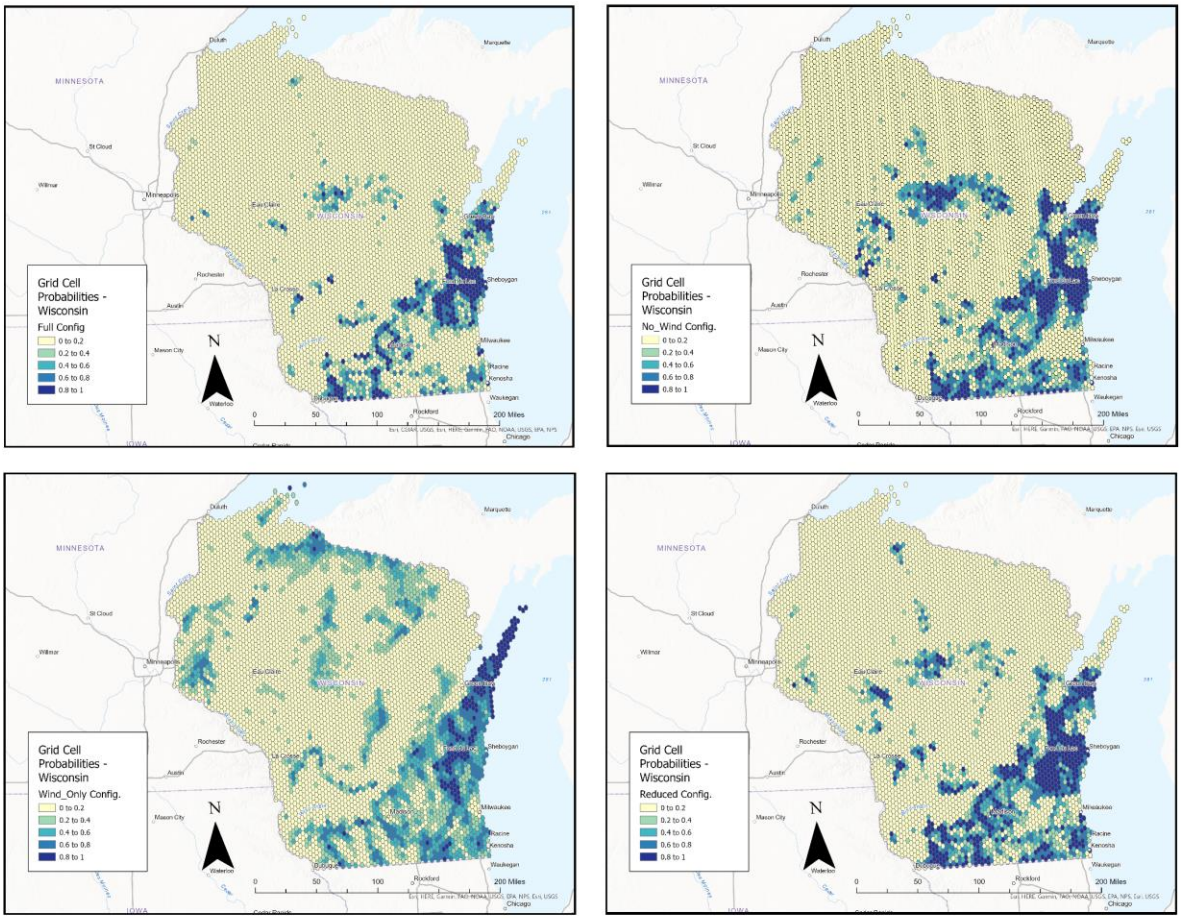


Figure 18: Wind Farm Site Suitability (WiFSS) surfaces constructed from running WiFSS-LRCA over Wisconsin, with each grid cell assigned a probability between 0 and 1. From top-left to bottom-right are the results from using the *Full*, *No_Wind*, *Wind_Only*, and *Reduced* predictor configurations. Grid cell size = 5,000 acres. Basemap from Esri [407].

Harper et al. [62] similarly found that ORs obtained from different combinations of predictors can provide spatial context to the outputs of an LR equation. These four maps ultimately attest to how the enlisted predictors can impact the WiFSS surface generated by WiFSS-LRCA, and moreover WiFSS-LRCA's sensitivity to predictor choices made by end-users.

Projecting believable wind farm siting futures depends on whether the suitability surfaces generated by WiFSS-LRCA are reflective of present commercial wind farm locations (see Sections 1.2 and 3.3.2). As such, present WiFSS surfaces like those in Figure 18 can be

compared against the classified existence or absence of wind farms in each grid cell, and clusters of high-probability regions indicated by Getis-Ord statistics [431], as summarized in Figure 19. A greater number of true positive grid cells, and fewer false positives, occurred in South and East Wisconsin under the *Full* configuration (top) versus the *Wind_Only* configuration (bottom). As such, the WiFSS surface created using the *Full* configuration (Figure 18, top-left) was more reflective of Wisconsin's existing commercial wind farms, meaning that constructing robust model outputs for Wisconsin required more predictors than Wind Speed alone. This result reflected in the validation of WiFSS-LRCA over Wisconsin, which classified 87.19% of testing grid cells correctly under the *Full* configuration, compared to 84.54% under the *Wind_Only* configuration, the biggest difference being fewer Type 1 errors under the former (see Section 4.2.2). Further context for grid cells classified as false positive comes from Figure 19's mapped Getis-Ord statistics. Many of the high-probability grid cells in Figure 18 are shown to exist in statistically significant clusters ($p < 0.05$), with these clusters containing almost all grid cells classified by WiFSS-LRCA as true positive. For that reason, the false positive grid cells in these clusters could be candidates for future commercial wind energy development.

4.3.3. Contrasting CONUS and State-Level Suitability Surfaces.

Comparison of ORs generated by U.S. state and CONUS-level model runs showed that WiFSS-LRCA's outputs are sensitive to selected spatial scale (see Figure 4.1.2), potentially translating into inconsistencies in the WiFSS surfaces constructed at these different scales. This potential is considered by examining maps presented in Figure 20 (*Full* configuration, grid cell size = 33,800 acres): classified grid cell states by WiFSS-LRCA for the CONUS (top) and for model runs over Iowa (bottom-left) and West Virginia (bottom-right). These two U.S. states were selected to convey regional limitations of constructing a WiFSS surface for the CONUS. For instance, while

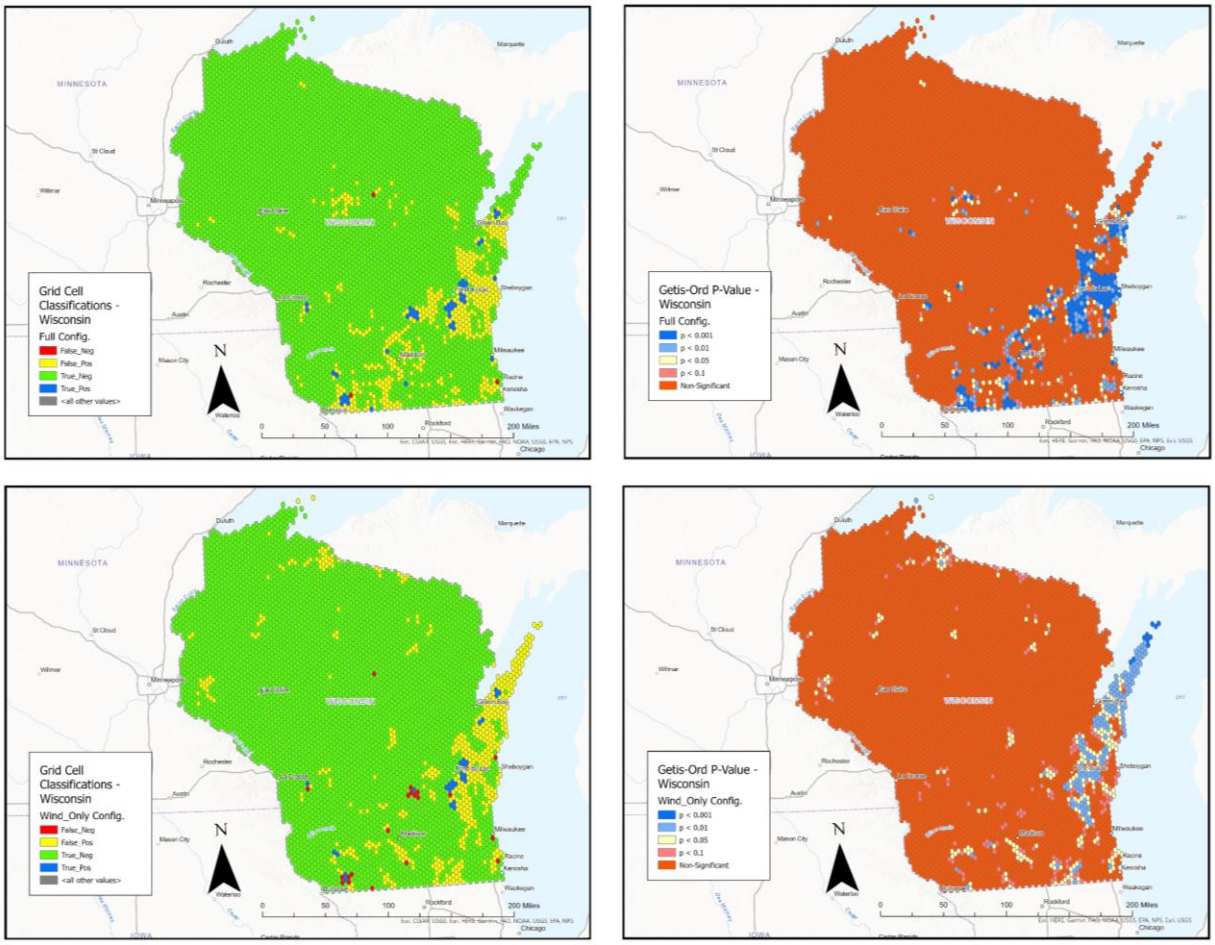


Figure 19: Outputs from the same model run that produced Figure 18, showing the classified grid cell states (left) and cluster analyses of the WiFSS surfaces using the Getis-Ord statistic (right). Results are shown from the *Full* (top) and *Wind_Only* (bottom) predictor configurations. Grid cell size = 5,000 acres. Basemap from Esri [407].

the CONUS model run does capture present wind farms in the Central Plains and Great Lakes, hence the many true positive (blue) grid cells in these regions, comparing this map to the WiFSS surface generated for Iowa shows a greater number of false positive (yellow) grid cells in the former. Runs of WiFSS-LRCA over the CONUS seem to overestimate the number of existing wind farms more than in runs over individual U.S. states, meaning CONUS model runs are more susceptible to Type 1 errors. Conversely, while the state-level run of WiFSS-LRCA over West Virginia classified no grid cells as false negative (red), several false negative classifications

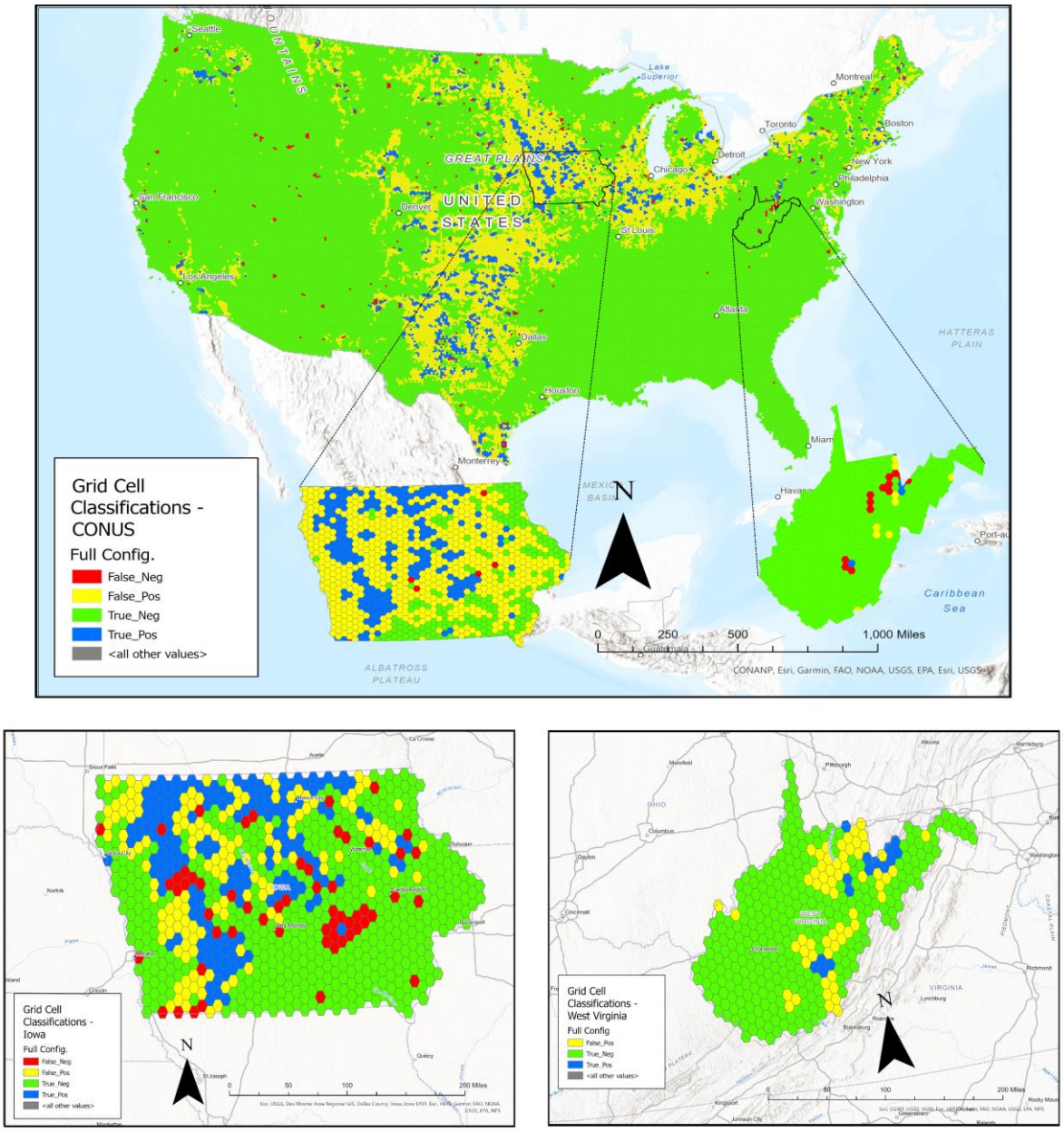


Figure 20: Classified grid cell states produced by running WiFSS-LRCA (*Full* configuration) over the CONUS (top), as well as over Iowa (bottom-left) and West Virginia (bottom-right). The grid cell states obtained for Iowa and West Virginia within the CONUS model run have been enlarged for comparison with those obtained from running WiFSS-LRCA over these two states separately. Grid cell size = 33,800 acres. Basemap from Esri [407].

occurred over West Virginia in the CONUS model run. WiFSS-LRCA similarly classified many grid cells in the Western United States as false negative, despite the successful identification of wind farms in state-level model runs over this region (Figures 9 and 15). A greater proportion of

Type 2 errors seems to occur in CONUS model runs versus individual U.S. states, though these errors are located further away from the Central Plains and Great Lakes, in contrast to the large number of Type 1 errors in these two regions. Recall from Figures 11 and 12 that spatial scale possessed a link to the predictors retained by WiFSS-LRCA and changes in their ORs, thereby altering computed probabilities for each grid cell and by extension their classifications.

The WiFSS surfaces associated with Figure 20 provide more context to the inconsistencies between state-level and CONUS-level runs of WiFSS-LRCA. The WiFSS surfaces produced by these same model runs are given in Figure 21, showing that the CONUS model run assigned higher probabilities to Iowa (and the Central Plains more broadly) compared to running WiFSS-LRCA over Iowa alone. With these higher probabilities comes more grid cells exceeding the median probability threshold for classication as containing a wind farm (see Section 4.2), hence the larger number of false positive grid cells over the Central Plains and Great Lakes compared to state-level model runs. The opposite is true when comparing outputs over West Virginia, with WiFSS-LRCA assigning lower P_i values across much of the state in CONUS model runs, hence the larger number of false negative classifications. The implication is that WiFSS surfaces generated by a CONUS-level model run are less representative of true commercial wind farm potential than when running WiFSS-LRCA over an individual U.S. state, with the highest-probabilities in the regions most populated by wind farms.

Despite WiFSS-LRCA's lower predictive accuracy in identifying wind farm locations and constructing WiFSS surfaces for the CONUS than individual U.S. states, the model still possesses high performance when run at a CONUS-level spatial scale. This performance is evident from the metrics produced from the calibration and validation of WiFSS-LRCA that subsequently yielded Figures 20 and 21; Figure 22 presents the ORs (top-left), ROC curves

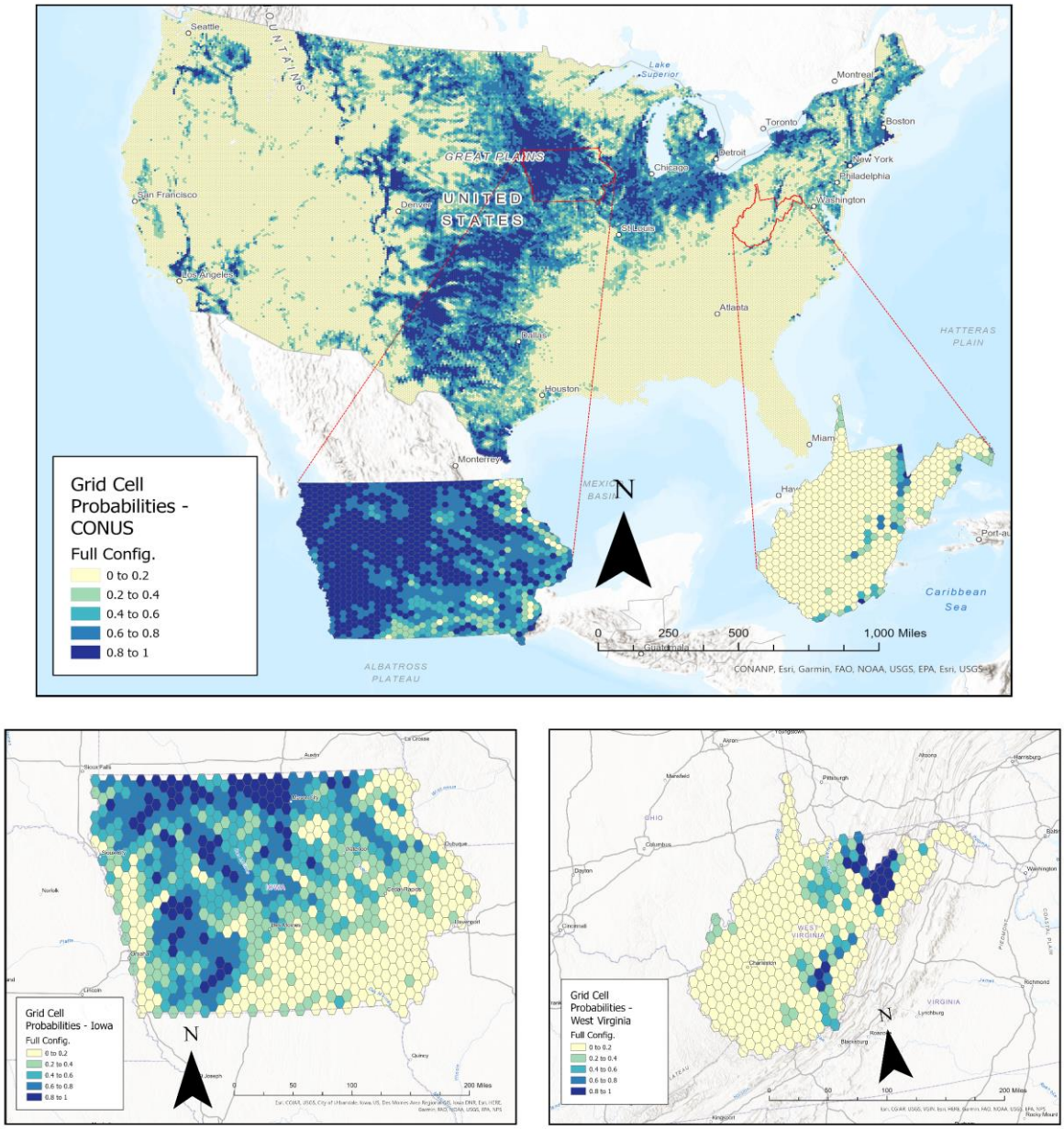


Figure 21: Same as Figure 20 but for the constructed WiFSS surfaces from the same model runs. Basemap from Esri [407].

(top-right), median confusion matrix (bottom-left) and boxplot (bottom-right) produced from this model run. The predictors with the strongest association with WiFSS-LRCA's predicted outcomes are consistent with those presented earlier in this dissertation (see Section 4.1.2), including *Avg_Wind*, *Near_Trans*, *Undev_Land, Numb_Pols,* and *Renew_Targ,* suggesting that

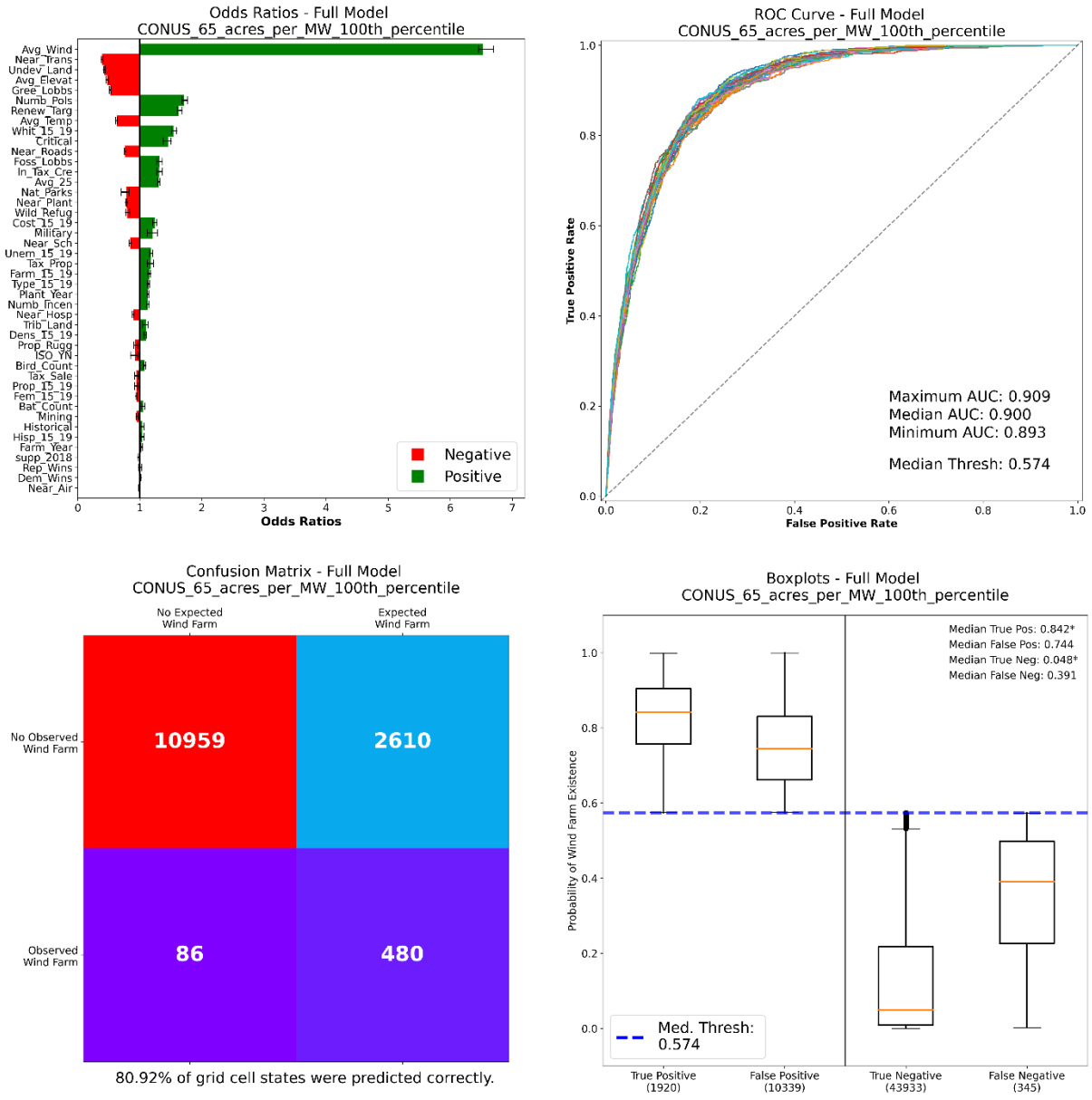


Figure 22: Results from the calibration and validation of WiFSS-LRCA that produced Figures 20 and 21. From top-left to bottom-right are the Odds Ratios of each predictor retained by the model, the Receiver Operating Characteristic curves produced by each repeated validation, the median confusion matrix, and the boxplots from applying the trained and tested model to all grid cells.

the most suitable areas for wind energy development are likely those with greater average wind speeds, nearer to transmission lines, few competing land uses, a legislature that is supportive of renewable energy, and ambitious RPS. The high-probability regions in Figure 21 are often indeed

those with high wind speeds [321], dense transmission line networks [326] and an abundance of flat, undeveloped land [80]. However, of the 44 predictors that were retained by this model run, only 23 of them (52.2%) possessed OR consistent with the expectations in Tables 7a and 7b, meaning geographical interpretations of Figures 20 and 21 should not be overstated, especially outside of the Central Plains and Great Lakes regions.

The ROC curve and confusion matrix illustrate a high rate of correct grid cell classification across the CONUS, with 80.92% of testing grid cells being classified on average (median) as either true positive or true negative. Despite the Type 1 and Type 2 errors presented in the above WiFSS surfaces, these metrics suggest that the predictors used in WiFSS-LRCA can still correctly identify the majority of commercial wind farm locations. Finally, the boxplots and Mann-Whitney U-test illustrate statistically significant differences in the median probabilities associated with each grid cell classification (true positive = 0.842, false positive = 0.744, true negative = 0.048, false negative = 0.391). This result gives credence to geographical explanations for the constructed WiFSS surface, such as high wind speeds over the Central United States and support from the legislature. WiFSS-LRCA is ultimately able to construct suitability surfaces in its first iteration that reflect present patterns of commercial wind energy development, at both U.S. state and CONUS-level scales (though the former with greater accuracy). These predicted wind farm locations are geographically explicable to a point, and are also sensitive to the predictor configuration and grid cell size selected by an end-user. This ability to capture existing conditions should make future sites for development projected by the model's CA component more trustworthy [450,451], thereby legitimizing this dissertation's purposes of projecting wind farm siting futures and demonstrating a larger scale application of LRCA than used in previous work (see Section 1.2).

Chapter 5: Results – Projection Using the CA Component.

Whereas the previous chapter focused on assessing currently suitable wind energy development locations, this chapter addresses which of the most suitable locations could gain commercial wind farms first. The focus of this chapter pertains to the third research question of projecting wind farm siting futures using WiFSS-LRCA, including their connections to geographic properties and the sensitivity to WiFSS-LRCA's own parameters (see Section 1.2). This dissertation distinguishes between prediction of current wind farm locations in WiFSS-LRCA's first iteration and projection of future wind farm locations in its subsequent iterations. Lemos and Rood [463] pose that prediction is deterministic, meaning it is concerned with whether a specific event will happen based on current conditions, such as whether high average wind speeds suggest a wind farm's presence. By contrast, projection communicates a range of possible future states that accommodate dynamics in independent variables, which are often temporally explicit [464]. Uncertain time horizons and multiple simultaneous processes (e.g., constructing roads and transmission lines, legislated setback distances, electricity cost fluctuations) mean that expressing modeled futures as ranges of possibilities is more appropriate than a deterministic approach [465]. This uncertainty is why WiFSS-LRCA has been constructed to accommodate multiple predictor configurations, grid cell sizes, constraints, neighborhood sizes, and scenario setups, allowing the model to project a range of futures for commercial wind energy development across the CONUS. Inclusion of constraints and neighborhood effects in WiFSS-LRCA represents application of CA at a spatial extent larger than most previous examples [442]. According to Wu et al. [444, p.1041], "spatial extent is often associated with a given study case, such that it does not have a universal feature of spatial scale sensitivity during CA-based land use change simulation", meaning that CA is not limited to applications of simulating city-level

urbanization [69-71]. As such, this chapter presents examples of WiFSS-LRCA's outputs over U.S. states and the CONUS to show that CA can project commercial wind farm siting futures. An example console output from running the CA component of WiFSS-LRCA is in Appendix A4.

5.1. Core Projections from Running WiFSS-LRCA Simulations.

5.1.1. Suggestion of Geographical Patterns.

As exemplified by Figure 10 (see Section 3.4), each iteration of WiFSS-LRCA projects a set number of grid cells to experience binary land use change (i.e., "yes" or "no" to a future commercial wind farm), based on a default capacity increase in Megawatts [92] or the increase set by an end-user. Geographical influences on common locations for projected wind farms can be suggested based on ORs, like in Chapter 4, along with how these influences change with spatial scale (U.S. state versus CONUS). Figure 23 presents maps from running WiFSS-LRCA over the CONUS (parameter setup is given in the figure caption), with grid cell colors corresponding to the year (iteration) in which grid cells acquire a wind farm. In both maps, future wind farms appear mostly in the South-Central Plains and Midwest. Of particular note is the concentrated wind energy development projected from 2040 onward in the Texas Panhandle, South Kansas, and East Colorado. As mentioned in Section 4.3.3, the Central United States region is populated by many wind farms [73] and possesses high wind speeds (*Avg_Wind*) [83,321], political support for renewable energy (*Numb_Pols; Renew_Targ*) [87,333], and large tracts of undeveloped land (*Undev_Land*) [80]. Furthermore, based on computed ORs, the predictors most associated with WiFSS-LRCA's projected outcomes included these predictors, with ORs of 23.53, 2.87, 2.38, and 0.20, respectively. Although ORs are not causal [422] and their magnitudes should not be overstated [95,96], those with the strongest associations appear geographically sound, while also consistent with the expected OR values in Table 7a/7b.

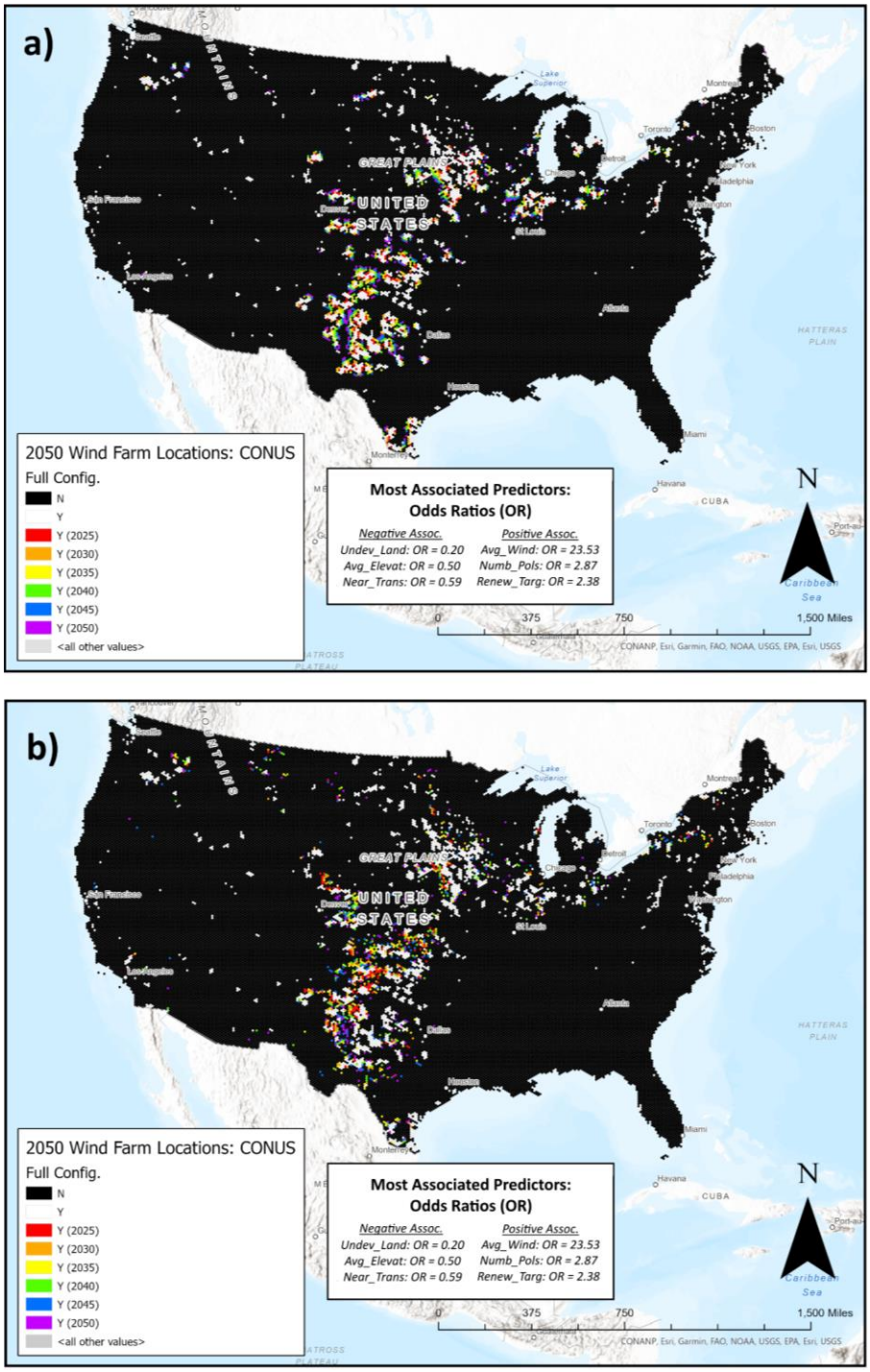


Figure 23: Outputs from running WiFSS-LRCA, showing projected wind farm sites out to the year 2050 for the CONUS. The maps display the year in which grid cells are projected to obtain a wind farm. Odds ratios of the most strongly associated predictors are also given. Default gained wind energy capacity [92]; *Full* predictor configuration; default constraints (*Mining* and *Historical* are switched off); all applicable scenarios used; neighborhood size = 2 (Neighborhood effects are on in Figure 23a, and off in Figure 23b). Grid cell size = 44,200 acres. Basemap from Esri [407].

Much like other CA-based model studies that observed cluster-like patterns of land use change [72], such clustering appears in Figure 23a due to the inclusion of $Neighb_i^t$ in Equation 4 (see Section 3.4). The influence of neighborhood effects on WiFSS-LRCA's projections can be assessed by removing $Neighb_i^t$, as in Figure 23b. While the South-Central Plains and Midwest remain the areas projected to undergo the greatest future wind energy development, smaller clusters also appear in Washington, Montana, and New York, especially after 2035. Another impact of switching off neighborhood effects is that wind farms in the Oklahoma and Texas Panhandles proliferate during the 2020s, earlier than most other locations. A westward shift of projected wind farm locations toward the Rocky Mountains also occurs; given this model run's strong associations with *Renew_Targ* and *Undev_Land*, New Mexico and Colorado's RPS [149] and undeveloped spaces perhaps influence projected wind farm locations more so than when not limited to neighboring grid cells. The biggest difference between these two simulation runs is that switching neighborhood effects off causes wind farms in WiFSS-LRCA's later iterations to appear farther from present clusters (white grid cells). For instance, East Iowa, East Oklahoma, and West Wisconsin all gain fewer wind farms in Figure 23b, implying that these regions have lower probabilities of wind farm occurrence based on the LR equation alone, as also suggested by Figure 21. Switching off the neighborhood effect transition rule reduces the tightness of future wind farm clustering around the CONUS' present wind farms, with the timing of cluster formation also changing slightly. These maps show the importance of including $Neighb_i^t$ in WiFSS-LRCA for characterizing wind farm siting potential around present wind farm clusters.

Figure 23 shows that regions of the CONUS where commercial wind farms presently exist are those with the greatest probability of future wind energy development, meaning few projected new wind farm locations exist in the Western and Southeastern United States. Running WiFSS-

LRCA over individual U.S. states identifies suitable future wind farm sites at smaller spatial scales, revealing both regional WiFSS (Figure 24a) and differences in WiFSS from CONUS model runs (Figure 24b) for Florida and Illinois, respectively. Future wind farm clusters in Florida are located mostly on the periphery of infrastructure, particularly Florida's cities and interstates, perhaps explaining why *Undev_Land* (OR = 0.20) and *Near_Trans* (OR = 0.59) have strong negative associations with model outputs. Projected wind farm locations also have a strong positive association with Percent of Population Under 25 (*Avg_25*: OR = 1.67), suggesting that WiFSS in Florida may be greater in areas with younger populations, which is important given the high median age of Floridians [466]. As for Illinois, the state-level run of WiFSS-LRCA that is shown in Figure 24b projects more future wind farms on the state's west side than does Figure 23a. Furthermore, predictors like Critical Habitats (*Critical*: OR = 1.84) and Bat Habitat Range Count (*Bat_Count:* OR = 0.49) have stronger associations with projected outcomes when running WiFSS-LRCA over Illinois only, predictors that are most prevalent on the state's east side [313,322]. However, both the state and CONUS-level simulations by WiFSS-LRCA agree on few future wind farms appearing to the north and south of Illinois' present wind farms, possibly due to lower average wind speeds given the strong positive association of *Avg_Wind* (OR = 2.43). Running WiFSS-LRCA at these two spatial scales shows how broader nationwide trends in potential future wind farm locations differ from those over individual U.S. states, along with corresponding differences in the predictors associated with localized versus broader trends in wind energy development. Recall from Section 3.3.2 that states lacking two or more grid cells that contain wind farms (such as Florida) use an LR equation trained and tested for the CONUS (rather than the individual U.S. state), though WiFSS-LRCA is nevertheless able to project future wind farm locations that can be explained geographically.

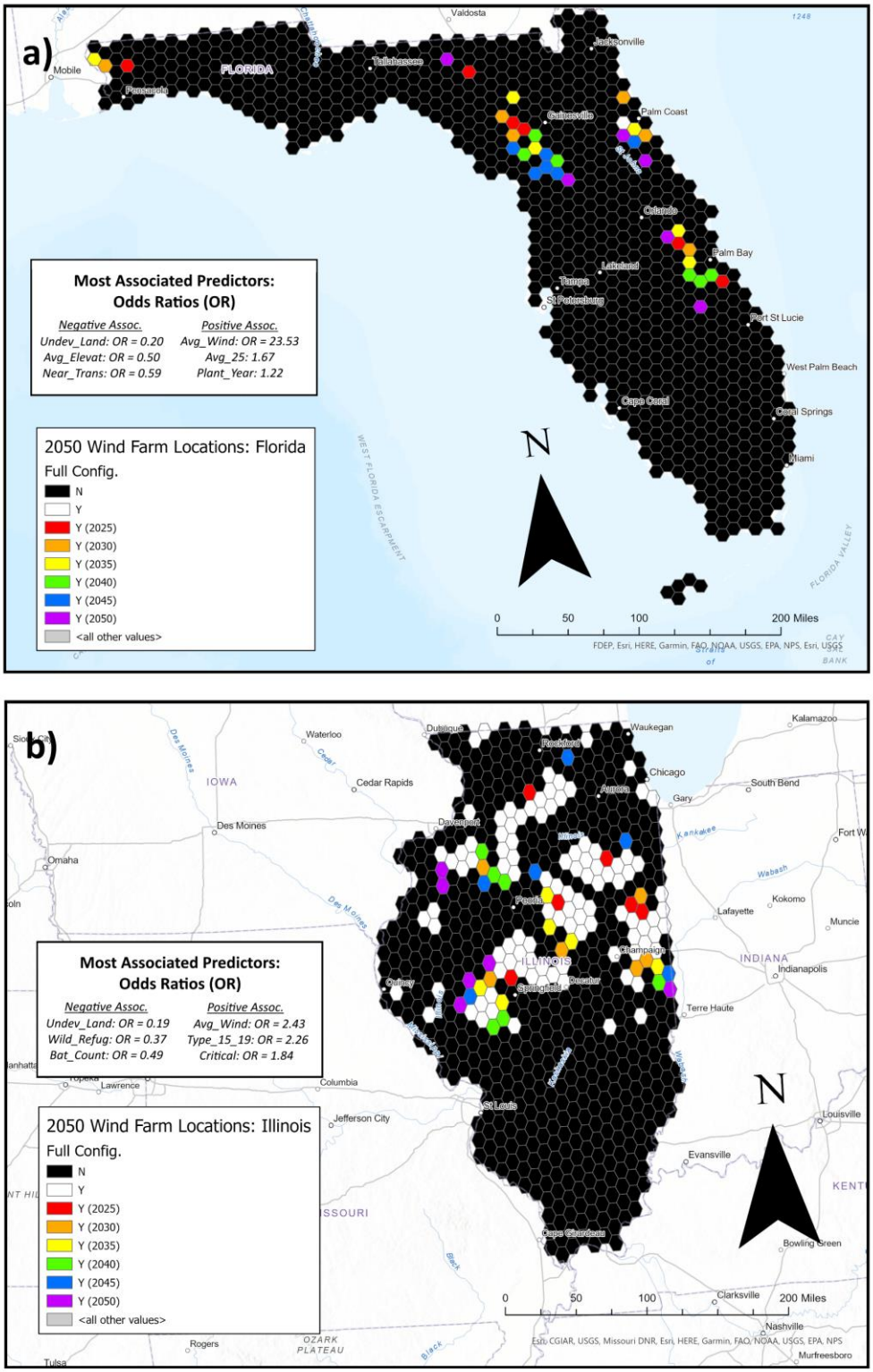


Figure 24: Same as Figure 23 but for runs of WiFSS-LRCA over Florida (Figure 24a) and Illinois (Figure 24b), and a 2,000 Megawatt gained capacity in each model iteration. Basemap from Esri [407].

5.1.2. Analysis of Projected Wind Farm Clusters.

Figures 18 and 19 (see Section 4.3.2) showed that grid cells classified as true positive during WiFSS-LRCA's first iteration are frequently part of high-probability clusters of the constructed WiFSS surface. The false positive grid cells in these clusters that also possessed statistical significance ($p < 0.05$) were proposed as being candidates for future wind energy development, something that WiFSS-LRCA can identify in its subsequent iterations. Indeed, the role of the neighborhood effect transition rule in Equation 4 is to represent decaying suitability for land-use change with distance from the physical feature of interest (i.e., commercial wind farms) [69]. One would thus expect that the grid cells projected to gain wind farms should be those that were initially classified as false positive and also part of these high-probability clusters. Figure 25 illustrates projected wind farm locations by 2050 for Iowa, New York, and Texas, along with how their grid cells were initially classified (false positive or true negative) during WiFSS-LRCA's first iteration. Of the 120 grid cells set to gain a wind farm during each model run (20 cells per iteration, see figure caption for details), 113 (94.2%) of those cells over Texas, 111 (92.5%) over Iowa, and 95 (79.2%) over New York were initially false positive.

Getis-Ord statistics [431] of the P_i values from WiFSS-LRCA's first iteration showed that 71 of Texas' 113 false positive grid cells (62.8%) that gained a wind farm were initially part of statistically significant clusters of high P_i values, with a similar result having occurred for New York (56 out of 95, or 58.9%). Iowa stands out for only seven of its 111 false positive grid cells (6.3%) being part of such clusters, for which there are two possible explanations. Firstly, based on Figure 20, Iowa's present commercial wind farms are large in number and spread out fairly evenly across the state, unlike in states such as West Virginia, Texas, or New York. Since standardized Getis-Ord statistics are a function of the number of physical features [467], more

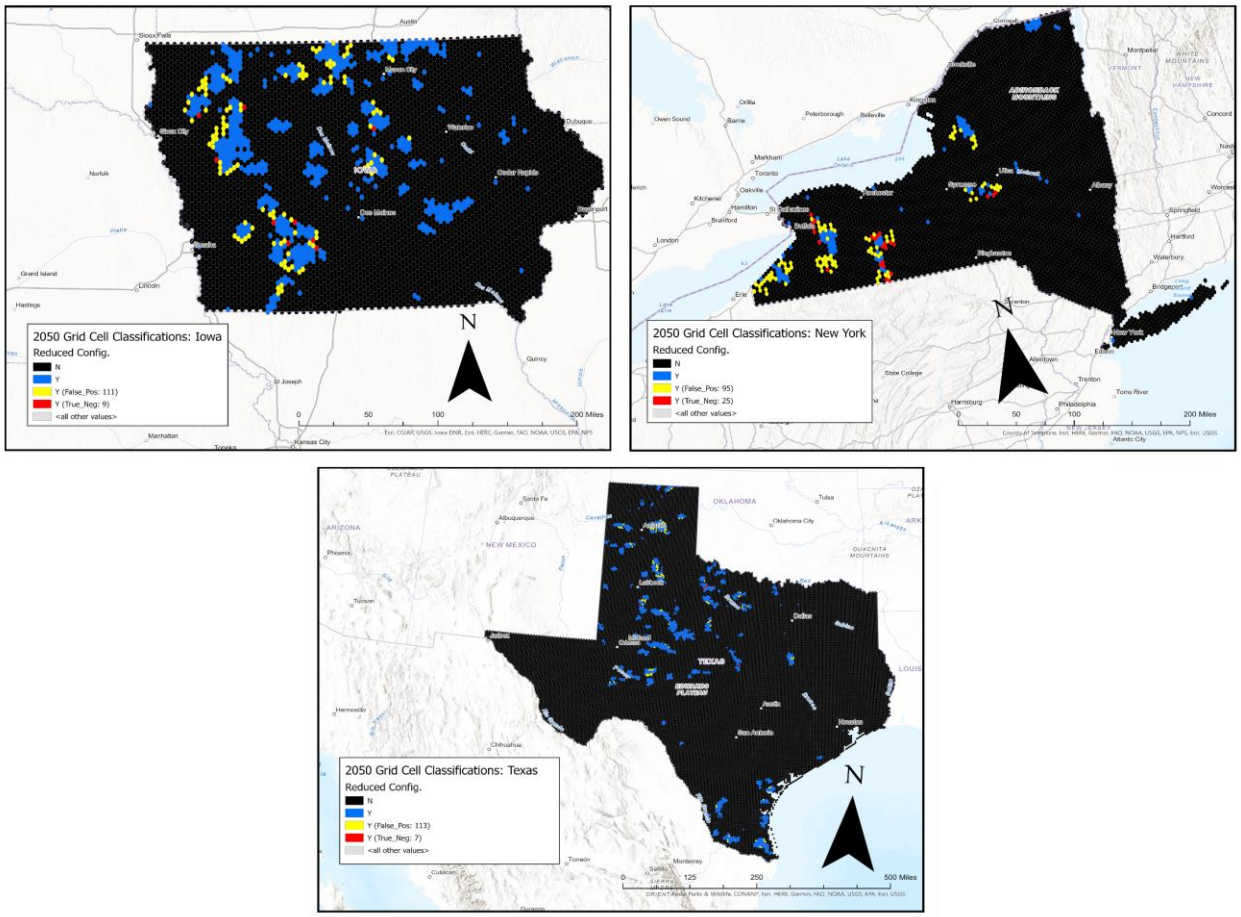


Figure 25: Outputs from running WiFSS-LRCA at the state-level, showing projected wind farm sites out to the year 2050 for Iowa (top-left), New York (top-right), and Texas (bottom). Grid cells initially classified as false positive (yellow) and true negative (red) are also illustrated. 4,000 Megawatt gained capacity in each iteration; *Reduced* predictor configuration; default constraints; all applicable scenarios used; neighborhood size = 2. Grid cell size = 5,000 acres. Basemap from Esri [407].

features mean lower scores and thus fewer clusters. Secondly, although Figure 21 does overestimate WiFSS across the Central United States in CONUS model runs (as discussed in Section 4.3.3), high probabilities pervade across much of Iowa, suggesting that WiFSS-LRCA's predictors indicate suitability for wind energy development (e.g., high wind speeds, flat land, productive legislature, etc.). With more grid cells possessing high P_i values, $Neighb_i^t$ becomes

less of a limiting term for computing $Prob_i^t$ (Equation 4), allowing grid cells to gain wind farms further away from any statistically significant clusters. Despite the large spatial scales of WiFSS-LRCA's application, in contrast to the city and county-level scales of previous LRCA studies [68,76,442], the influence of neighborhood effects on the likeliest locations for wind energy development is evident from Figure 25, also giving credence to the cluster-like patterns discussed in Figures 23 and 24.

5.1.3. Impact of Predictor Configurations on Model Projections.

It was previously established that certain predictor configurations allow WiFSS-LRCA to capture the number of grid cells that presently contain commercial wind farms more accurately than others, namely the *Full* and *Reduced* configurations (see Section 4.2). Since predictor configurations impact computed P_i values for each grid cell, the final values of $Prob_i^t$ and thus the projected grid cell states would by extension be altered, meaning that future WiFSS surfaces could be sensitive to predictor configurations. Projected wind farm locations across North Dakota are shown in Figure 26, using the *No_Wind* (Figure 26a) and *Full* (Figure 26b) predictor configurations. The wind farm locations projected by both configurations share some similarities; wind energy development in the 2020s is concentrated on the state's southwest side (where many wind farms currently exist), along with a new wind farm cluster appearing near the border with Minnesota during the 2040s. However, this cluster to the east is larger, forms earlier, and extends to the southeast corner under the *Full* configuration, with the *No_Wind* configuration instead adding to the larger cluster in the southwest after 2035. The only difference in constructing these two maps is the exclusion of *Avg_Wind* under the *No_Wind* configuration, with Figure 26 thus showing that Wind Speed's inclusion greatly alters the spatial distribution and timing of future wind energy development. Previous studies have frequently justified Wind

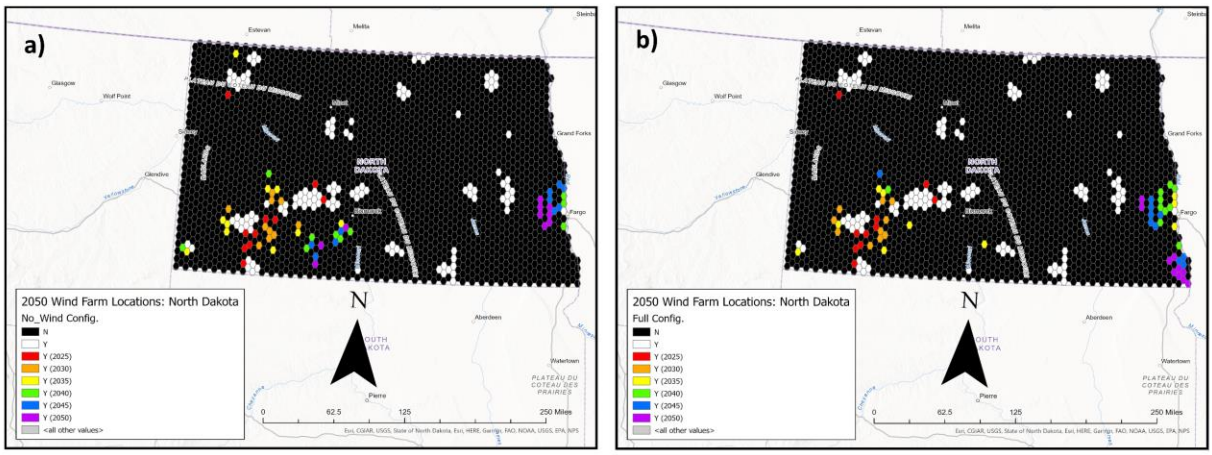


Figure 26: Outputs from running WiFSS-LRCA at the state-level, showing projected wind farm sites out to the year 2050 for North Dakota. The maps display the year in which grid cells are projected to obtain a wind farm. 4,000 Megawatt gained capacity in each iteration; *No_Wind* predictor configuration used for Figure 26a; *Full* predictor configuration used for Figure 26b; default constraints; all applicable scenarios used; neighborhood size = 3. Grid cell size = 13,000 acres. Basemap from Esri [407].

Speed's inclusion in WiFSS assessments because of its assumed importance compared to other predictors [44,124,218]. WiFSS-LRCA goes one step further by mapping out the sensitivity of future wind farm siting decisions to whether Wind Speed is included as a predictor, and the persistence of this sensitivity regardless of the model's constraints and neighborhood effects.

Constraints and predictor configurations can, however, have compounding influences on projections made by WiFSS-LRCA. Since $Const_i$ places a Boolean restriction on whether a grid cell can gain a commercial wind farm (see Section 3.4), one could imagine both maps in Figure 26 looking quite different should the constraints have been modified. Furthermore, because of the impacts that the enlisted predictor configuration has on P_i values, a new set of constraints may prohibit wind energy development for grid cells in which P_i is high, thus yielding a different future WiFSS surface. The dual impact of constraints and predictor configurations is shown in

Figure 27, future WiFSS surfaces produced by running WiFSS-LRCA over Kentucky using the *Wind_Only* (Figure 27a) and *Reduced* (Figure 27b) configurations. Because the two predictor configurations yield a different set of P_i values for all grid cells, the highest-probability grid cells that gain wind farms first are not the same in both maps, being concentrated in two locations under the *Reduced* configuration compared to four under the *Wind_Only* configuration. Added to this, many grid cells in West Kentucky violate WiFSS-LRCA's default constraints, particularly Mining Operations and Wildlife Refuges (Table 8), limiting the growth of future wind farm clusters in this area by preventing neighboring grid cells that violate said constraints from gaining wind farms. Consequently, no grid cells gain wind farms during the final iteration (2050) under the *Reduced* configuration because all neighboring grid cells in this area that lack a wind farm violate at least one constraint. These results demonstrate WiFSS-LRCA's response to predictor configuration choices in two ways: 1) excluding certain predictors from WiFSS-LRCA may obscure pathways to wind energy capacity targets should those predictors conflate with the constraint transition rule; and 2) the collection of projected wind farms in West Kentucky under the *Reduced* configuration illustrates the important role that predictors other than Wind Speed have on future wind farm siting decisions.

5.2. Sensitivity Analysis of WiFSS-LRCA's Parameters.

5.2.1. Sensitivity to Grid Cell Size.

According to Ménard and Marceau's [442, p.710] CA-based study of agriculture encroachment on forested land: "finer exploration of cell size sensitivity suggests that even small variations in cell size can produce significant divergence in results when scale thresholds are crossed." It is therefore expected that different grid cell sizes could drastically alter the pattern of projected future wind farm locations across a selected study area. Alongside catering to the interest of end-

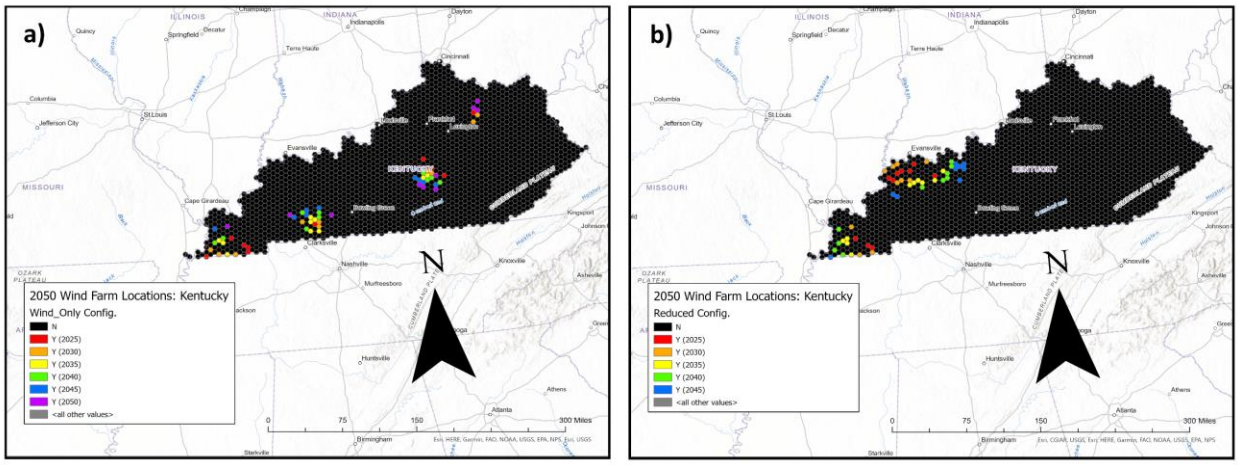


Figure 27: Same as Figure 26 but for applications of WiFSS-LRCA over Kentucky, with the *Wind_Only* predictor configuration used for Figure 27a and the *Reduced* predictor configuration used for Figure 27b. Basemap from Esri [407].

users in experimenting with different commercial wind farm capacities, preparing aggregated predictor data at 20 different resolutions (Table 6) allows the sensitivity of WiFSS-LRCA's projections to grid cell size to be tested. An example of this sensitivity is presented in Figure 28, future wind farm locations projected for Maine at two different grid cell sizes: 23,400 acres (low-resolution; Figure 28a/c) and 6,750 acres (high-resolution; Figure 28b/d). Some agreement exists between the two projections, namely the earliest future wind farms (2025 and 2030) appearing in Southeast Maine, an area with overall less rugged land [79] and a relatively dense transmission line network [326]. WiFSS-LRCA runs at both resolutions also agree on future wind energy development being limited on the state's west side, perhaps due to the area's mountainous terrain complicating commercial wind farm installation [218]. Grid cell size therefore does not seem to impact the general areas in which future wind energy development is projected to occur.

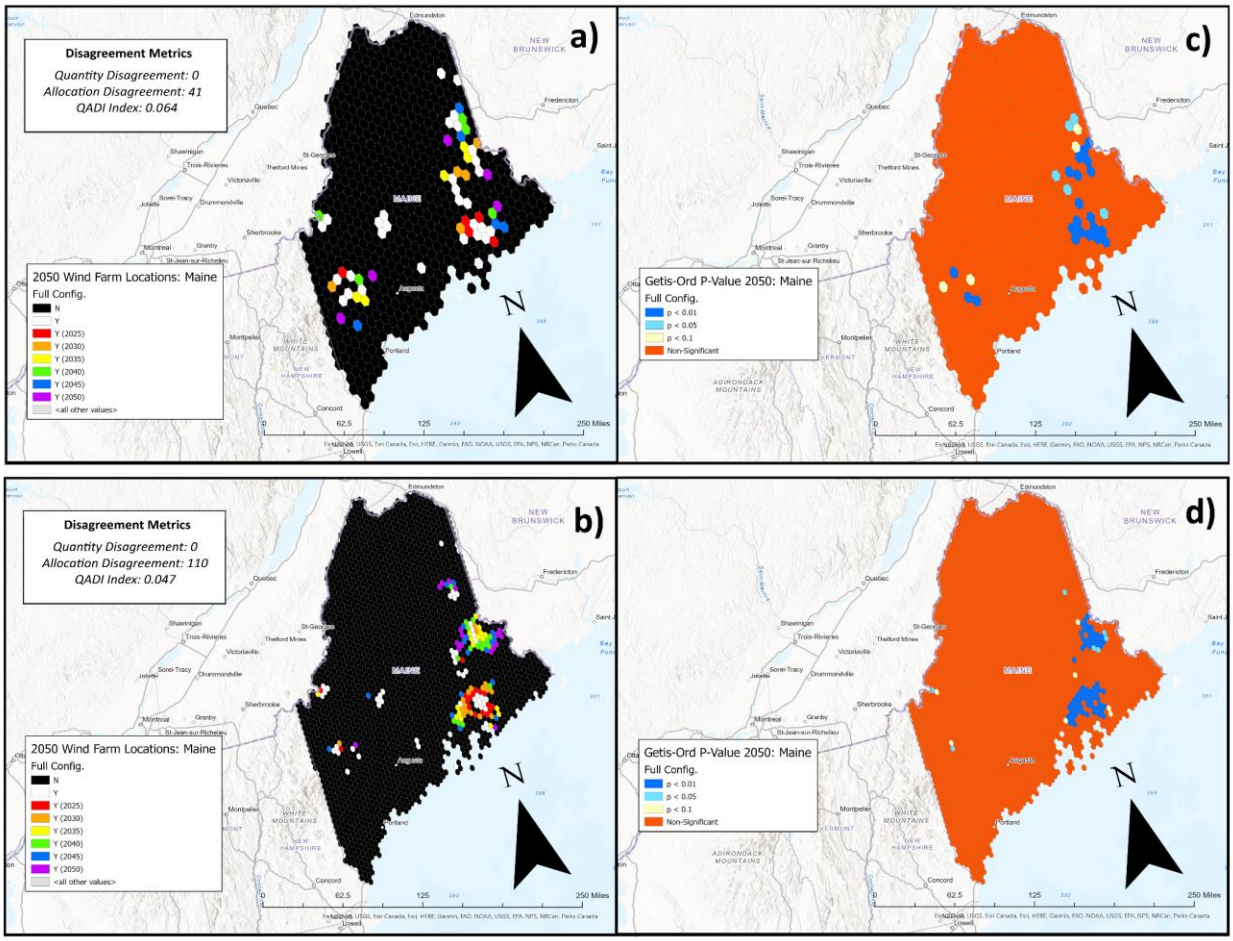


Figure 28: Outputs from running WiFSS-LRCA at the state-level, showing projected wind farm sites out to the year 2050 for Maine at a lower (Figure 28a, 23,400 acres) and a higher (Figure 28b, 6,750 acres) grid cell resolution. Also given are p-values obtained from computed Getis-Ord statistics (Figures 8c and 8d), based on the $Prob_i^t$ values that produced Figures 28a and 28b. 2,500 Megawatt gained capacity in each iteration; *Full* predictor configuration; default constraints (*Critical*, *Historical,* and *Mining* are switched off, *Near_Trans* increased to 15,000 meters); *Climate Change*, *Demographic Changes* and *New Infrastructure* scenarios are used; neighborhood size = 3. Basemap from Esri [407].

Differences between WiFSS-LRCA's outputs at different resolutions present themselves when comparing the timing of land use change and the tightness of the formation of projected wind farm clusters. Firstly, the north edge of Maine's eastern cluster expands at different times depending on the grid cell size, expanding from 2030 onward at the lower resolution (Figure 28a) compared to 2040 onward at the higher resolution (Figure 28c). More interesting is the clusters of high-probability grid cells being tighter at the higher resolution, as evidenced by

computing Getis-Ord statistics. At both resolutions, statistically significant ($p < 0.01$) hotspots of $Prob_i^t$ values exist in two large clusters, though these clusters are less dense and displaced northward at the lower resolution (Figure 28c) versus the higher resolution (Figure 28d). QADI statistics [445] further highlight the tighter clustering that occurs when enlisting smaller grid cell sizes, with the higher resolution map's lower statistic (QADI = 0.047, versus 0.064 at the lower resolution) conveying less disagreement with the null version of WiFSS-LRCA (see Section 3.5.1). Smaller grid cell size therefore increases the influence of WiFSS-LRCA's constraint and neighborhood effect transition rules on projections, compared to the influence of the LR equation's predictors, resulting in denser, more localized patterns of projected land-use change. Pan et al. [443] came to a similar conclusion that smaller grid cell sizes cause CA-based models to project land-use change that sprawls over shorter distances, though WiFSS-LRCA's sensitivity to grid cell size has less impact on the timing of these projections.

5.2.2. Sensitivity to Neighborhood Size.

An increased neighborhood size (i.e., a greater value of *n*; Equation 5) means that the decaying influence on land-use change with distance stretches further from present physical features [68]. In other words, grid cells that lack wind farms are more likely to gain one if present wind farms have a larger defined influence. Like with grid cell size, the sensitivity of WiFSS-LRCA's outputs to neighborhood size is explored based on differences in timing and location of projected wind energy development, presented again for Maine in Figure 29 (Figure 29a: $n = 1$, Figure 29b: $n = 3$, Figure 29c: $n = 5$). As in Figure 28, two clusters of wind energy development are projected in East and Southeast Maine, but neighborhood size greatly changes the timing of these clusters' expansion. At the largest neighborhood size ($n = 5$), Southeast Maine's cluster is the only location projected to gain wind farms in 2025, with only two grid cells outside of this

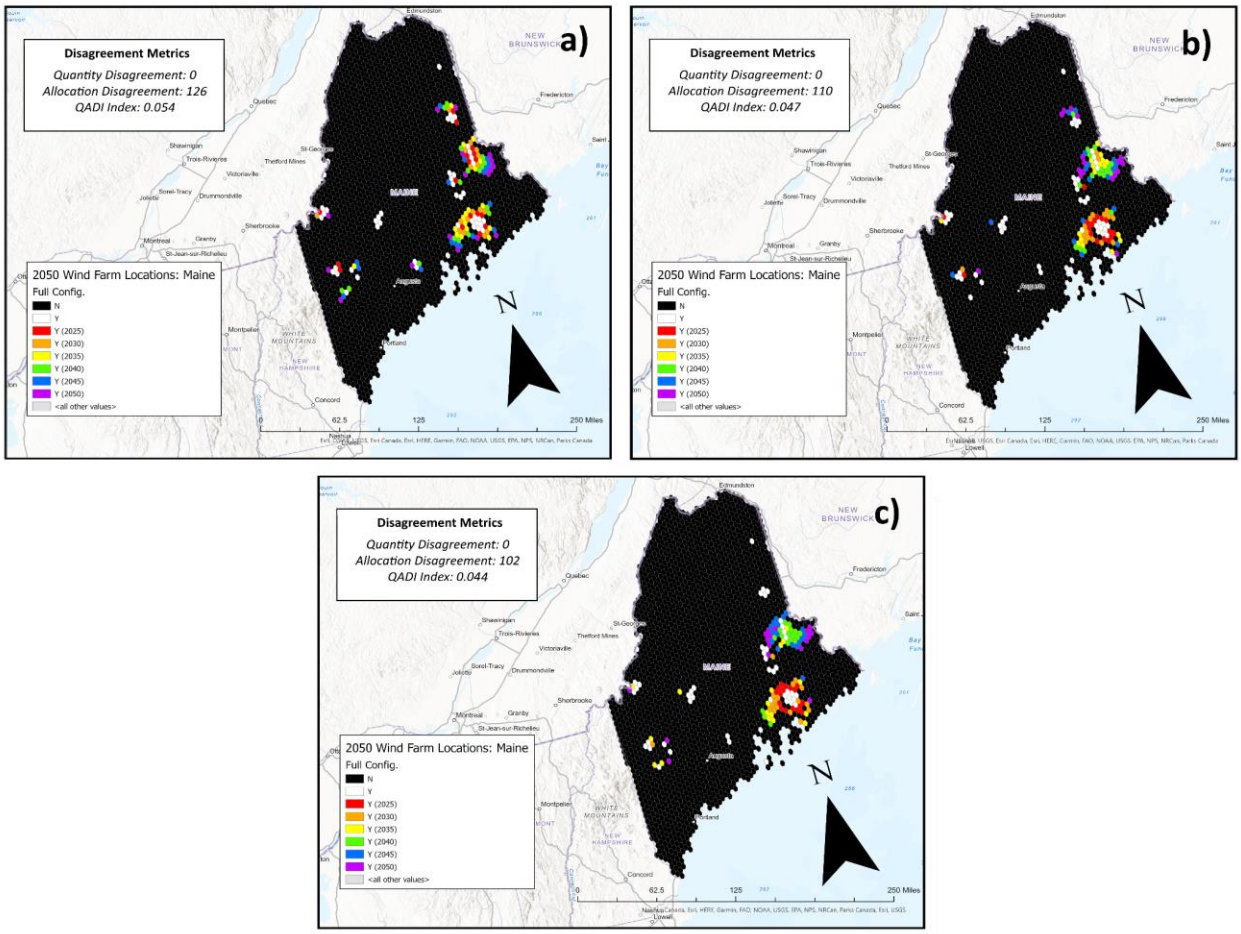


Figure 29: Same as Figure 28 but with each map possessing a different neighborhood size, being n = 1 (Figure 29a), n = 3 (Figure 29b), and n = 5 (Figure 29c). Grid cell size = 9,000 acres. Basemap from Esri [407].

cluster experiencing land-use change in 2030. Furthermore, East Maine's cluster expands progressively later as neighborhood size increases, gaining no commercial wind farms in 2025 (and only one in 2030) in Figure 29c. The construction of Equations 4 and 5 explain why Southeast Maine increasingly gains wind farms first at larger neighborhood sizes. The likelihood of $Neighb_i^t$ equaling zero is higher for smaller neighborhood sizes, since there are fewer chances for grid cells to already contain a wind farm, meaning $Prob_i^t$ will also equal zero regardless of

the value of P_i. Consequently, a larger neighborhood means fewer grid cells are assigned a $Prob_i^t$ value of zero, thus allowing the high P_i values of these grid cells to influence WiFSS-LRCA's projections. It was noted in Section 5.2.1 that the model's predictors were conducive for wind energy development in Southeast Maine [79,326]. Increasing WiFSS-LRCA's prescribed neighborhood size thus increases the suitability for larger present wind farm clusters being the first to expand in the model's early iterations.

As well as the above changes in timing, some evidence also exists for neighborhood size affecting the locations of projected future wind energy development. Based on visual inspection, the East and Southeast Maine clusters become slightly larger at greater neighborhood sizes, again due to fewer grid cells being assigned a $Neighb_i^t$ value of zero in WiFSS-LRCA's first few iterations. Of more interest, however, is the decline in both the QADI statistic and the allocation disagreement score as neighborhood size increases; the smallest neighborhood size (Figure 29a) yields the greatest disagreement with projections made by the null WiFSS-LRCA model (QADI = 0.054, allocation disagreement = 126). Projected wind farm locations more closely resemble the null model's output as neighborhood size increases, meaning that the neighborhood effect transition rule ($Neighb_i^t$) supersedes the influence of the equation-based transition rule (P_i) when a larger neighborhood is prescribed. However, it was previously noted that a larger neighborhood size is precisely what allows grid cells with high P_i values to gain wind farms. It therefore stands to reason that a neighborhood size that is too large may reduce WiFSS-LRCA's ability to accurately project future wind farm locations, an observation that Li et al. [468] similarly made in their study of trend-adjusted LRCA modeling of urban sprawl. Based on these results, the sensitivity of CA-based models to neighborhood size noted by Kocabas and Dragicevic [437] is also produced when running WiFSS-LRCA. Furthermore, neighborhood

size's strong impact on the timing of future wind energy development suggests that Wu et al.'s [444] conclusion of higher CA sensitivity to neighborhood size than grid cell size equally applies to WiFSS studies that take this modeling approach.

5.2.3. Sensitivity to Constraints.

Predictor configurations and the constraint transition rule have been shown to have compounding effects on WiFSS-LRCA's projections (see Section 5.1.3), such that otherwise suitable grid cells cannot gain future wind farms should they, for instance, be too far from transmission lines or overlap with military installations (Table 8). Long et al. [469] found that trialing different sets of constraints is key to preparing a CA-based model that captures urban forms that are in line with planning preferences. In the same way, running WiFSS-LRCA while varying its default constraints allows an end-user to project wind farm siting futures that may or may not be limited by these constraints. Cluster-like expansion of commercial wind farms could thus occur in different locations should WiFSS-LRCA's default constraints be modified, hence the model's potential sensitivity to its own constraints. Figure 30 illustrates outputs from running WiFSS-LRCA over Montana using three constraint setups: the default set from Table 8 (Figure 30a), a loosened setup that deactivates *Wild_Refug* and increases the distance for *Near_Trans* to 15,000 meters (Figure 30b), and all constraints deactivated (Figure 30c). A more restrictive constraint setup limits the grid cells that gain wind farms, plus the iteration of WiFSS-LRCA after which future wind energy development cannot occur; the default constraints prevent grid cells from gaining wind farms after 2035, and the loosened constraints extend development over Montana to 2040. Although WiFSS-LRCA's constraints are defined based on the literature and this dissertation's systematic review [40] (see Section 2.2), Figure 30 shows that the default constraints are restrictive enough over some U.S. states to prevent completion of the model's

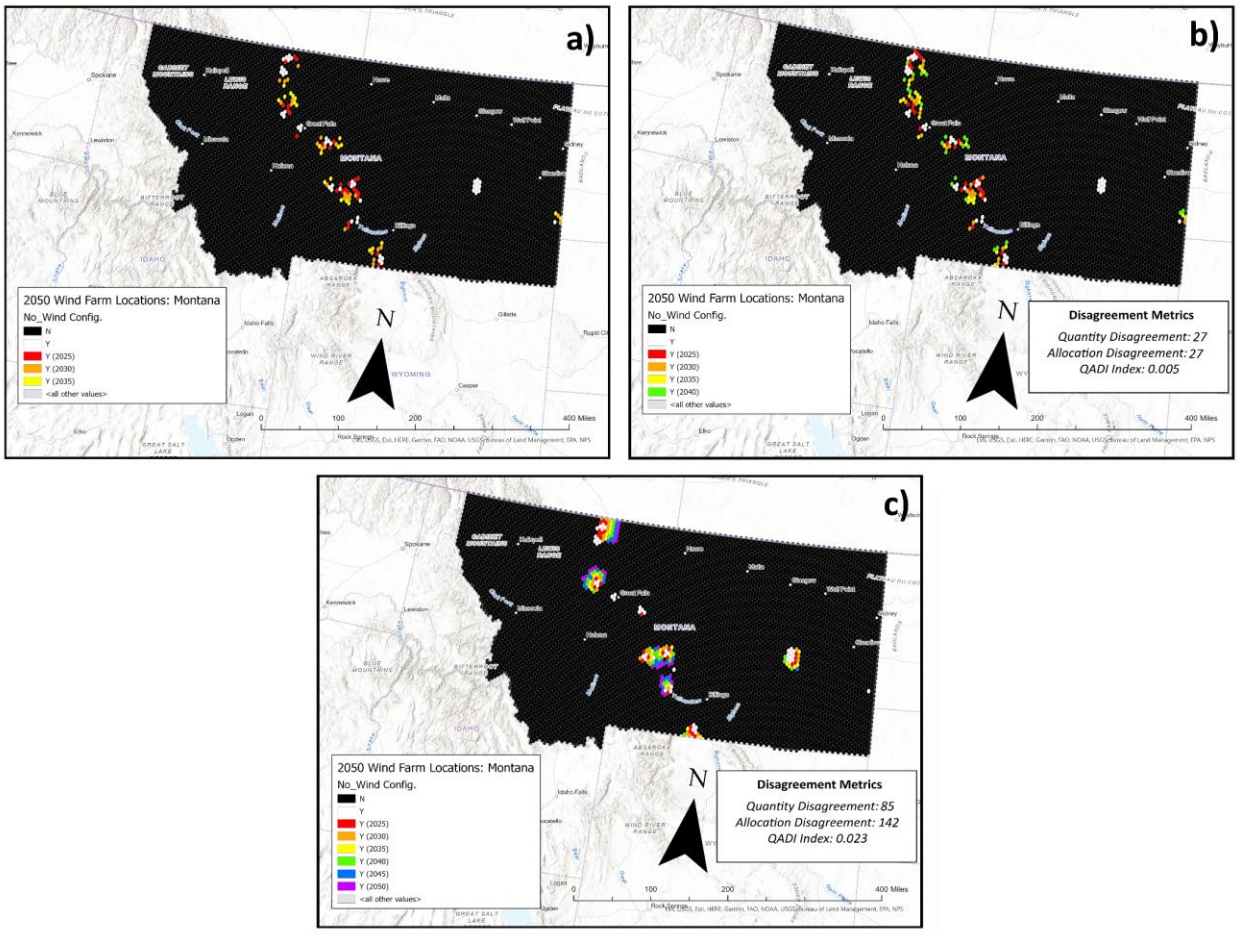


Figure 30: Outputs from running WiFSS-LRCA at the state-level, showing projected wind farm sites out to the year 2050 for Montana using the default constraints (Figure 30a), a loosened constraint setup (Figure 30b, *Wild_Refug* turned off and *Near_Trans* extended from 10,000 to 15,000 meters) and all constraints switched off (Figure 30c). QADI statistics are computed using Figure 30a as a point of comparison, rather than the output from the null model. 4,000 Megawatt gained capacity in each iteration; *No_Wind* predictor configuration; *Changing Energy Economies, New Infrastructure,* and *Urban Protection* scenarios are used; neighborhood size = 2. Grid cell size = 9,750 acres. Basemap from Esri [407].

iterations. Furthermore, as Figure 10 showed, many grid cells that contain present wind farms should be restricted based on WiFSS-LRCA's default constraints, meaning that the constraint transition rule is an inherently imperfect depiction of where future wind energy development is feasible. Deactivating WiFSS-LRCA's constraints (Figure 30c) allows iterations to complete out to 2050, though the cluster pattern of grid cells that gain wind farms is quite different. The

cluster of present wind farms in East Montana does not expand at all in Figures 30a and 30b, though this cluster grows in almost all directions without the constraint transition rule in effect, gaining wind farms in all but the final iteration. An MCDA-based WiFSS study by Díaz-Cuevas [226] found that constraints that were too restrictive rendered all of Córdoba (Spain) unsuitable for wind energy development. The expansion of East Montana's wind farm cluster only when deactivating WiFSS-LRCA's constraints attests to a similar result, exemplifying the model's sensitivity to its own constraints. Switching off constraints means that only WiFSS-LRCA's neighborhood effect transition rule is able to restrict future wind energy development (i.e., automatically set $Prob_i^t$ to equal zero), causing projections to occur in tighter clusters around present developments, similar to the clustering in Figures 28 and 29. This result highlights the interdependence of a CA-based model's constraint and neighborhood effect transition rules. For WiFSS-LRCA, the tendency of neighborhood effects to limit land-use change to compacted areas of development [470] often works against that of constraints to place exogenous limits on where said development is allowed to occur [69]. An important difference between these two transition rules, however, is that modifying constraints more easily reduces the number of grid cells that gain commercial wind farms, thus affecting both quantity and allocation disagreements between WiFSS-LRCA's generated maps [445] (changes to neighborhood size only affected the allocation disagreement; Figures 28 and 29). As Figure 30b and 30c show, loosening the constraints increased their quantity disagreement with Figure 30a, thus increasing the computed QADI statistic and quantifying WiFSS-LRCA's sensitivity to how its constraints are defined. However, using looser constraints would require an end-user to accept potentially detrimental impacts of future wind energy development, such as impacts on Wildlife Refuges [341,342].

5.2.4. Sensitivity to Scenario Setups.

Whereas the previous sections pertain to WiFSS-LRCA's sensitivity to its neighborhood effect (Section 5.2.2) and constraint (Section 5.3.3) transition rules, using the scenarios defined in Table 9 to modify Equation 1's coefficients allows sensitivity to the model's equation-based transition rule to be considered. Modifying the LR equation in this way allows for consideration of how changes in WiFSS-LRCA's predictors alter the projected sites for wind energy development, and subsequently the importance of each predictor for devising wind energy policy. As discussed in Section 3.5.1, modifying predictor coefficients to construct scenarios assumes that ORs represent how strongly associated changes in each predictor are with the suitability for land-use change, recognizing that associations are not causal [422]. Projections made by WiFSS-LRCA using three different scenario setups over Washington are shown in Figure 31, one using the *Default* scenario (Figure 31a), one using all eight available scenarios for a state-level model run (Figure 31b), and one using a *Custom* scenario that modifies the coefficients of the five most and five least associated predictors (based on OR values in WiFSS-LRCA's first iteration) by ±50% in each iteration (Figure 31c). All three maps exhibit a similar cluster pattern: future wind energy development is concentrated around three present clusters on the state's east side, with development elsewhere nonexistent due to WiFSS-LRCA's constraint and neighborhood effect transition rules. Equation 4 sets the value of $Prob_i^t$ to zero for a grid cell that violates at least one constraint ($Const_i = 0$) or is too far away from other grid cells that contain a wind farm ($Neighb_i^t = 0$), limiting scenario setup to mostly influencing the timing of expansion of existing wind farm clusters without greatly changing their spread or location. It can thus be inferred that scenario setup (and by extension modifications to P_i) has less influence than constraints and neighborhood effects on location differences in projected wind farm sites.

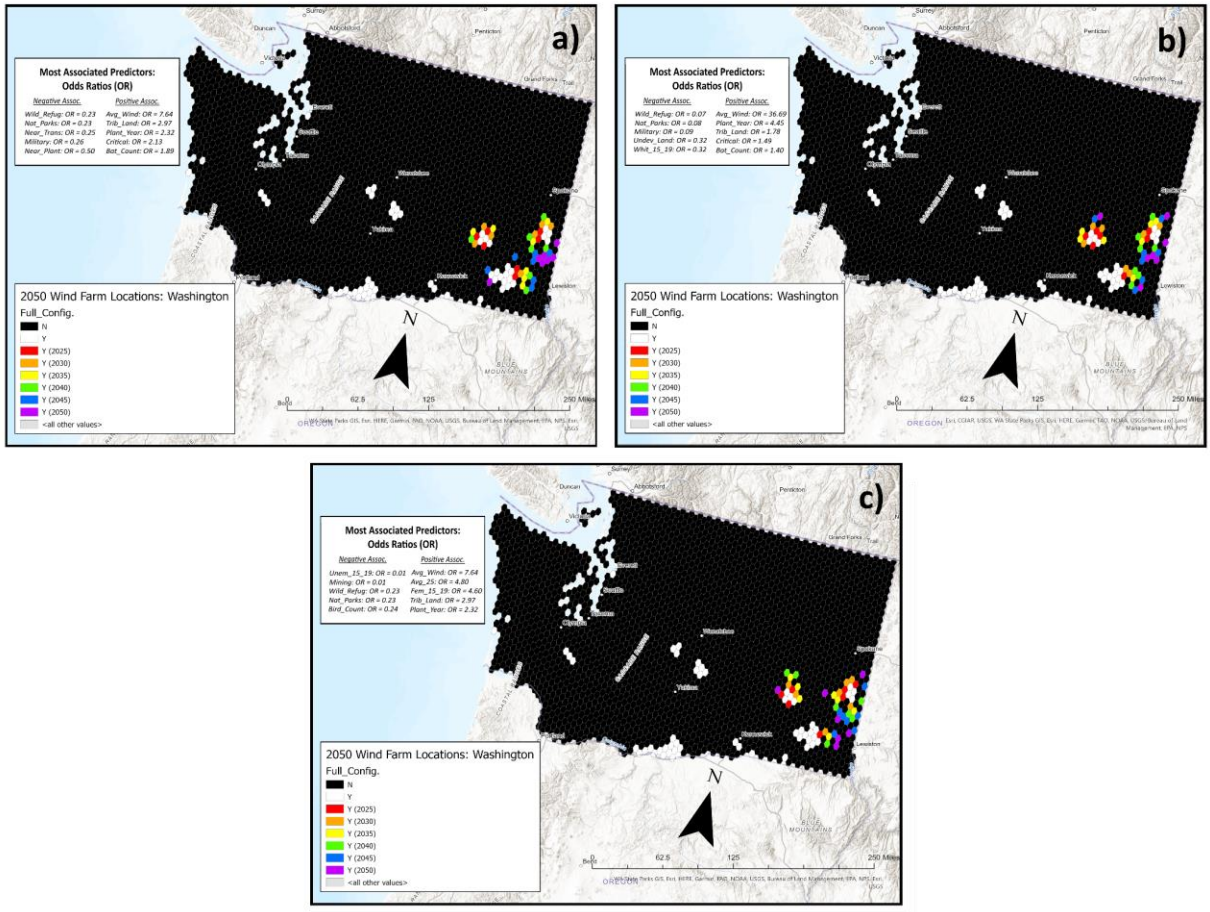


Figure 31: Outputs from running WiFSS-LRCA at the state-level, showing projected wind farm sites out to the year 2050 for Washington using the *Default* scenario (Figure 31a), all eight scenarios available for state-level model runs in Table 9 (Figure 31b) and a Custom scenario (Figure 31c). The custom scenario modified the coefficients of the 5 most and 5 least associated predictors based on ORs by ±50% in accordance with Table 9. 3,500 Megawatt gained capacity in each iteration; *Full* predictor configuration; default constraints (*Mining* and *Near_Plant* switched off); neighborhood size = 2. Grid cell size =13,000 acres. Basemap from Esri [407].

The expansion of Washington's three easternmost wind farm clusters can be understood by comparing them against the ORs (listed on each map in Figure 31) of WiFSS-LRCA's predictors. East Washington is an area that is characterized by a lack of vulnerable species habitats (*Critical*) [313], few designated tribal lands (*Trib_Land*) [318], higher wind speeds compared to areas further west (*Avg_Wind*) [321], and a denser transmission line network than the areas immediately north and south of these clusters (*Near_Trans*) [326]. All four of these features

correspond to some of the most positively and negatively associated predictors listed on all three maps, with *Avg_Wind* consistently having the most positive association. While the ORs can offer suggestions for why future wind energy development in East Washington is suitable, they cannot explain the differences in timing and spread of each cluster's projected expansion. For instance, the westernmost of these three clusters is projected to expand the fastest in Figure 31b (all scenarios), having obtained the most new wind farms by 2035 (yellow grid cells). As another example, Figure 31a (*Default* scenario) projects the southernmost of the clusters expanding westward in 2045 (blue grid cells) and 2050 (purple grid cells), which the other two scenarios do not. Most interesting is Figure 31c (*Custom* scenario) from 2040 (green grid cells) onward, with the two clusters furthest east merging to become one larger cluster along the border with Idaho, and the westernmost cluster expanding northward earlier than in the other two maps. These differences in the timing of projected wind energy development are comparable to the effects of changing grid cell size (Figure 28), though those of modifying scenario setups are more localized to the present wind farm clusters. However, it should be acknowledged that this method of modifying WiFSS-LRCA's equation-based transition rule inherently alters the LR equation's goodness-of-fit [415].

Compared to modifying the other parameters discussed here, constructing scenarios to represent changes in WiFSS-LRCA's predictors affects the timing more so than the location of future wind energy development. That being said, the dynamic land-use change simulated by WiFSS-LRCA's Cellular Automata component is still evident, specifically considering how the grid cells that gain wind farms in prior iterations affect those of subsequent iterations. For instance, while all maps in Figure 29 show that wind farms in later iterations tend to appear on the periphery of those of earlier iterations, this tendency strengthens with larger neighborhood size (Figure 29c),

with the Southeast Maine cluster's youngest wind farms by 2050 being consistently further away from the center of the cluster. This dynamic of later land-use change being influenced by earlier land-use change is also evident from the sensitivity to scenario setup. While Figure 31c depicts a well-defined merging of Washington's two easternmost wind farm clusters after 2040, Figures 31a and 31b project similar behavior but without the clusters becoming consistently closer to each other with each iteration. The presented sensitivity analysis does not obscure the overall pattern of future wind farms appearing concentrically around those already present or that are projected in WiFSS-LRCA's first few iterations, attesting to the continuation of cluster-like wind farm installation that has historically occurred across the CONUS [73].

The results presented in this chapter show that WiFSS-LRCA is indeed able to project wind farm siting futures that that reflect geographical features of the CONUS, with the predictors most associated with these projections changing based on the study area and spatial scale considered (Section 5.1.1). In addition to geographical context, the definition of WiFSS-LRCA's constraint and neighborhood effect transition rules also have important consequences for the model's projections, with there being demonstrable sensitivity to both rules. Altering the equation-based transition rule, i.e., the predictors used in the LR equation, has less impact on WiFSS-LRCA's projections, limited mostly to the timing of when grid cells acquire future wind farms (Section 5.2.4). By contrast, changes to neighborhood size and the restrictiveness of constraints greatly alter how tightly future wind energy development is clustered at both state-level and nationwide scales (Sections 5.2.2 and 5.2.3). Most importantly, the ongoing pattern of commercial wind farms being installed in regions of the CONUS with commonly suitable locations [73] is expected to continue in future decades, given the number of grid cells initially classified as false positive that gain wind farms in WiFSS-LRCA's subsequent iterations (Section 5.1.2).

Chapter 6: Discussion – Summary, Limitations, and Future Work.

The objective of this dissertation is to present a new model that is capable of projecting suitable locations for future commercial wind energy development across the Conterminous United States. With the ongoing pressure to address climate change by decarbonizing the nation's electricity production [9], as well as to improve domestic energy security by reducing reliance on foreign energy imports [15-17], Wind Farm Site Suitability models have the role of informing decision-makers of where said development should be focused. As with other Socio-Environmental Systems (SES) models, a key advantage of WiFSS models is their systematic approach and simplification of the wind farm siting process to produce a better system understanding [28] by combining a set of predictors (e.g., wind speed, transmission line proximity, political climate, etc.) in some manner to compute a suitability score [114]. Whereas existing WiFSS models identify locations for potential wind farms based on present-day conditions, whether by constructing a suitability surface [44,47] or ranking candidate sites [51,52], there previously existed no WiFSS models before this study that explicitly project which of the most suitable locations could gain wind farms first in the years to come. The addition of such a temporal component allows for projection of a range of possible wind farm siting futures, taking cues from Cellular Automata-based modeling studies that simulate future land-use change scenarios, especially of urban growth [66,98]. Unlike most CA-based approaches to land-use change modeling, however, WiFSS-LRCA (the model built for this dissertation) has been developed for application to larger spatial domains, while still assuming that the constraint and neighborhood effect transition rules [434] of CA models are applicable. This dissertation focuses on demonstrating WiFSS-LRCA's ability to make accurate predictions of present-day commercial wind farm locations at both state- and CONUS-wide spatial scales, and that its

subsequent projections are geographically sound, thereby accrediting its temporally explicit CA-based modeling approach. This chapter summarizes how each research question laid out in Section 1.2 has been answered, critical comments about the model's limitations, and future work that could follow from this dissertation.

6.1. Answering the Research Questions.

6.1.1. Question 1: Where is Logistic Regression-Cellular Automata situated within the broader scope of Wind Farm Site Suitability modeling approaches?

The decision to develop a WiFSS model that combines a Logistic Regression equation and Cellular Automata decision rules came from evaluation of the abilities of existing WiFSS models. This evaluation took the form of a review of common modeling approaches (Section 2.1) and of predictor selection and representation [40] (Section 2.2), conveying WiFSS-LRCA's differences from models with similar objectives as well as their shared abilities. Table 12 summarizes the similarities and differences in WiFSS-LRCA's construction and results presentation compared to those of the four WiFSS modeling approaches covered in Section 2.1. Between Table 12 and the discussion below, a consensus is derived of where WiFSS-LRCA fits among other modeling approaches and when WiFSS-LRCA would be the preferred model to use.

1. GIS-Based Multicriteria Decision Analysis.

Limiting land use change to feasible areas is achieved by CA-based models using constraint transition rules [434] that designate certain grid cells as unsuitable for development, based on violation of Boolean conditions or quantitative values for each predictor. Incorporating data layers that represent undevelopable tracts of land is also important to GIS-MCDA modeling approaches [109], with the same objective of preventing undesirable predictions of land-use change. Restricting wind energy development to feasible areas e.g., absence of vulnerable bird

Model Approach	Similarities with WiFSS-LRCA	Differences from WiFSS-LRCA
GIS-Based Multicriteria Decision Analysis (GIS-MCDA)	1. Representation of predictors as Boolean constraints and/or as changing continuously in space. 2. Computation of a suitability score for each individual spatial unit. 3. Flexible model design that accommodates spatial scale choices and scenario building. 4. Model output vulnerable to inaccurate or overly strict Boolean constraints.	1. Predictors are retained as layers of data superimposed on top of each other. 2. Weighting schemes that involve human decision-making compute the influence of each predictor's data layer. 3. Including a discrete predictor requires explicit spatialization or a proxy continuous predictor. 4. Temporal explicitness requires data preparation at multiple points in time.
Non-GIS-Based Multicriteria Decision Analysis (Non-GIS-MCDA)	1. Predictor selection is not limited by whether each one's data are discrete or continuous. 2. Identification of specific locations most suited for wind energy development first.	1. Weighting schemes that involve human decision-making compute the influence of each predictor's data layer. 2. Suitability surfaces are not constructed, instead presenting wind farm site comparisons in tabular form. 3. Greater reliance on primary, non-spatial data to inform selected predictors. 4. Temporal explicitness requires data preparation at multiple points in time.
Bayesian Network (BN)	1. Wind farm site suitability is inherently represented as a probability between 0 and 1. 2. Accommodation of scenario building by modifying the state of predictors. 3. Capable of handling predictor datasets that are natively discrete or continuous.	1. Depiction of causal relationships between predictors through conceptualizing "child" and "parent" nodes. 2. Scenario building is more useful for present-day risk assessment than projecting wind farm siting futures. 3. Representation of all predictors as discrete data, regardless of original data type. 4. Lack of spatiotemporal explicitness without integration of a GIS or other extra model components.
Logistic Regression (LR)	1. Wind farm site suitability is inherently represented as a probability between 0 and 1. 2. Used to define WiFSS-LRCA's equation-based transition rule. 3. Applicable to a GIS for producing spatial depictions of wind farm site suitability. 4. Capable of handling predictor datasets that are natively discrete or continuous.	1. Temporal explicitness requires data preparation at multiple points in time. 2. No mechanism for constraining spatial units as unsuitable for wind energy development.

Table 12: Summary of the similarities and differences between WiFSS-LRCA and models that are commonly enlisted for wind farm site suitability analysis, based on the literature presented in Sections 2.1 and 2.2.

habitats [244], high average wind speed [233], and proximity to road networks [224], is thus common practice in WiFSS modeling studies, which Cellular Automata inherently incorporate. Both WiFSS-LRCA and GIS-MCDA are sensitive to how their enlisted constraints are defined. Section 5.2.3 showed that WiFSS-LRCA's default constraints (Table 8) can be restrictive enough that the model cannot complete its iterations over certain states (due to no remaining suitable grid cells), while also characterizing some grid cells that contain wind farms in the first iteration as unsuitable. According to Li and Yeh [435, p.135], constraints in CA-based models "are used to make more reliable and reproducible predictions of actual land-use patterns", which WiFSS-LRCA showed to be occasionally imprecise. Constraints used in GIS-MCDA models are similarly vulnerable to missing data or imprecise definition [120], as shown by Rodríguez-Rodríguez et al.'s [305] study of GIS-MCDA sensitivity to protected bird habitat definitions. Beyond enlisting some predictors as constraints, the final product of both WiFSS-LRCA and GIS-MCDA is a continuous suitability surface that assigns every spatial unit a score. Whereas GIS-MCDA approaches typically produce aggregate scores derived from compositing all predictors' (weighted) dataset layers [219], WiFSS-LRCA condenses this information into a single probability value between zero and one (e.g., Figure 21). Both approaches are equally vulnerable to a loss of information compared to the original predictors [120] while also spatially conveying suitable locations for wind energy development.

There are some differences between WiFSS-LRCA and GIS-MCDA's approach to wind farm siting decisions, notably how discrete predictor data are handled. The "Tabular" datasets in Table 5a and 5b are inherently discrete because they aggregate a single value (whether of Green Lobbies (*Gree_Lobbs*), Percent Female Population (*Fem_15_19*), Unemployment Rate (*Unem_15_19*), etc.) to each state/county across the CONUS (see Section 3.2), as opposed to the

predictors represented by raster and vector datasets that are continuous in space. Aggregation onto each hexagonal grid cell allows WiFSS-LRCA's Logistic Regression equation to handle predictors depicted by either discrete or continuous data, "so long as interactions among the explanatory variables do not affect the response" [60, p.199]. Including discrete predictor data is hard in GIS-MCDA models because each dataset layer must be explicitly spatial and combinable with other predictors to produce a composite suitability surface [61]. As such, predictors in Table 5b are rarely used in GIS-MCDA approaches without constructing raster layers from one or more proxy predictors, such as Harper et al. [201] constructing a social acceptability data layer based on predictor associations derived from an LR equation (e.g., distance to national parks, mean age, political representation). Weighting each predictor's importance also works differently in these two modeling approaches. Literature examples of GIS-MCDA applied to WiFSS assessment frequently use an AHP method that computes predictor weights that enlist ranked expert opinions of each predictor's importance [44,53], with alternative weighting methods including BWM [55,134], PROMETHEE [131,139], and TOPSIS [230,263] (see Section 2.1.1 and 2.1.2 for summaries of each). WiFSS-LRCA circumvents expert involvement by preparing scenarios that iteratively modify the LR equation's predictor coefficients with each model iteration, taking cues from Harper et al.'s [62] LR-based study of wind farm project acceptance and Yang et al.'s [98] method of scenario construction within an LRCA modeling approach.

2. Non-GIS-Based Multicriteria Decision Analysis.

Non-GIS-MCDA benefits from a similarly broad range of predictors as those enlisted by WiFSS-LRCA. Since the former relies on primary data to represent the predictors that are of interest to either the researchers or experts [51,135], the limitation that GIS-MCDA faces of requiring all predictors of WiFSS to be representable by secondary data does not apply. Non-GIS-MCDA

approaches can thus seek data about any of the predictors listed in Tables 5a and 5b, as well as predictors that WiFSS-LRCA does not include due to a lack of secondary data availability at state- or CONUS-wide spatial scales, such as noise pollution [143], visual impact [52], and life cycle wind farm costs [135]. Another similarity between Non-GIS-MCDA approaches and WiFSS-LRCA is that both models share a common objective of identifying an order of development for potential future wind farm sites, albeit approached in different ways. In WiFSS-LRCA, grid cells with the greatest probability of gaining a wind farm based on running Equation 4 are identified in each iteration (every five years), explicitly timing when each grid cell would be most suited for development (e.g., Figures 23 and 24). Conversely, the typical approach of Non-GIS-MCDA is to select a small number of possible wind farm locations and rank them from most to least suitable for development, with the ranking typically based on the linguistic [51,143] or ordinal [131,134] judgment of a set of predictors by outsider experts. These predictors are frequently weighted using the same methods as those used in GIS-MCDA approaches to WiFSS, with Rouyendegh et al.'s [132] Non-GIS-MCDA study of wind farm site selection in Turkey using a TOPSIS approach to "reflect the judgments of decision makers and deal with the complexity of the decision process". The approaches that Non-GIS-MCDA and WiFSS-LRCA take to explicitly identifying the first and last locations for wind energy development are thus quite different, with the former considering individual locations and collecting primary data for a single point in time, though the motivation is essentially the same.

Although Non-GIS-MCDA and WiFSS-LRCA both aggregate their predictors to obtain dimensionless numbers, their methods of doing so are different, resulting in different representations of the same predictors. As described in Section 3.2, despite all data being secondary, the employed aggregation methods depend on each predictor's native datatype (raster,

vector, tabular) and whether the data are Boolean or quantitative, in line with the aggregation methods laid out by Plassin et al. [77] and LR-based environmental modeling studies [399-401]. Following aggregation, data values for all relevant predictors are assigned to each hexagonal grid cell; Boolean data are assigned as a value of 0 (“N”, no) or 1 (“Y”, yes) but the quantitative data’s magnitude varies, meaning all predictors are not aggregated to a common scale by WiFSS-LRCA. As for Non-GIS-MCDA, primary data for all predictors are collected and aggregated in a uniform linguistic or ordinal manner. For example, Wu et al.’s [135] offshore wind farm siting study in China’s Shandong Province asked experts to rate importance of six siting criteria (wind resources, construction, onshore conditions, environmental impact, economy, societal benefit) on a linguistic scale from “extremely high” to “very very low”, later converted into a prescribed numerical format. Regardless of how predictors may be classified or the phenomena they each represent, Non-GIS-MCDA approaches to WiFSS treat all predictors equally. Another crucial difference between these two modeling approaches is that the output of Non-GIS-MCDA is discrete by design. Figure 2 exemplifies the tabular output of most Non-GIS-MCDA studies [136], in which each site’s candidacy for wind energy development is represented by a collection of numbers and metrics. Map construction in Non-GIS-MCDA studies is typically used only to convey locations of candidate wind farms, as in Kaya and Kahraman’s [51] study of onshore WiFSS in Turkey. Maps that display the suitability for wind farm installation and projected wind farm locations are, by contrast, the core model output of WiFSS-LRCA.

3. Bayesian Networks.

The strongest similarity between WiFSS-LRCA and BN models is that both represent likelihood of land-use change (whether wind energy development or otherwise) in a probabilistic way to derive a binary outcome. In the case of BN, this outcome is based on statistical relationships

constructed between the acyclic connections of the model's predictors [150], whereas WiFSS-LRCA takes the log-odds of the sum of the predictors to also obtain a value between zero and one (Equation 1, see Section 3.3). Another similarity between these two modeling approaches is their ability to handle both discrete and continuous datasets, in contrast to GIS-MCDA models' ability to only use the latter. WiFSS-LRCA's methods of aggregation and its LR equation's ability to limit errors when combining different datatypes [60] allow for predictor datatypes to be retained as either discrete or continuous. The approach of most BN models is to represent each predictor as discrete potential states (e.g., strong versus weak winds, far or close to transmission lines), in order to "increase the [BN's] level of generalization" [161, p.398] for different study contexts. As such, any predictor whose dataset is natively continuous is discretized for inclusion in the network [157], such as constructing probability distributions from wind speed data split into quantiles. While BNs do not maintain predictors as continuous, and a consistent method for their discretization is lacking [165,166], these models are theoretically capable of enlisting as wide a range of predictors as WiFSS-LRCA (and of Non-GIS-MCDA models). Another key feature of BNs is that they are specifically designed for performing scenario-based modeling, like WiFSS-LRCA. Prescribing a condition to a BN's parent node(s) (e.g., setting "wave height" in Figure 3 to "Level1" or "Level2" [65]) changes the likelihood of the possible outcomes of the network's child node, which BN studies like Pinarbaşi et al. [156] have done to assess scenarios of offshore wind farm siting potential. This approach of devising scenarios is different though analogous to the predictor coefficient modifications performed using WiFSS-LRCA.

The distinctive feature of a BN is its illustration of the acyclic connections between predictors, thereby depicting causal relationships. Borunda et al. [64] use this depiction to their advantage by using a BN to illustrate how a selected city connects both to local energy consumption and

regional differences in wind speed. Arora et al. [162, p.440] note specifically that LR-based models are not suited to depicting causal relationships because “they operate under restrictive assumptions about the relationships among variables”, which in WiFSS-LRCA means a linear sum (Equation 1) cannot convey how predictors may vary together or influence each other. Some LR-based models enlist multilevel equation structures that account for how singular predictors can influence groups of predictors that operate at different spatial or temporal scales [93], such as Shu et al. [76] enlisting “grid cell level” and “town level” predictors in their LRCA simulations of urban growth. Such modifications effectively introduce a depiction of causality between the predictors in LR-based models. Beyond differences in representing causality, BN and WiFSS-LRCA also differ in terms of their treatment of space and location. Much like Non-GIS-MCDA models, spatial explicitness is difficult to represent in BN models, particularly when interested in representing spatial interactions between different locations, given the need to construct a new acyclic network for each location of interest [171]. Uusitalo et al. [157] also note that the acyclic structure of BN models prevents them from representing feedback loops, which WiFSS-LRCA can represent thanks to its iterative adjustment of predictor coefficients and the consequent effects on its equation-based transition rule (see Section 5.2.4). Most notably, BNs cannot construct suitability surfaces on their own like WiFSS-LRCA does, because each constructed acyclic network typically represents the conditions of a point location or small region in space. New studies are appearing, however, that combine BNs and GIS techniques to construct maps of the probabilities of event occurrence derived from the former [172].

4. Logistic Regression.

Since an LR equation is what defines one of WiFSS-LRCA's transition rules, and is therefore used to train and test the model to identify the locations of present commercial wind farms, similarities do exist between the two of them. LR-based approaches to WiFSS typically possess two core outputs: a map that illustrates spatial differences in the probability (between zero and one) of an event's occurrence [62,179], and computed coefficients/ORs showing the association between each predictor and said event [180,189]. LR approaches that do not construct suitability surfaces illustrating these probabilities are typically those that rely on primary data, often in the form of survey responses to opinions about predictors relevant to the wind farm siting process [181]. LR models that rely on secondary data for predictors are, by contrast, frequently used to construct suitability surfaces, Figure 4 being a quintessential example that illustrates probabilities for every raster cell of aggregated data [58]. WiFSS-LRCA's first iteration relies on secondary data in the same manner, constructing hexagonal surfaces of the probability of a commercial wind farm's existence (e.g., Figure 21) to convey suitability for wind energy development. The LR-based studies referenced in this dissertation produce model outputs analogous to that of WiFSS-LRCA's first iteration, i.e., the results presented in Chapter 4. It is the lack of temporal explicitness of using an LR model alone that distinguishes these referenced studies from WiFSS-LRCA. Unless provided with data for each predictor for a future point in time, an LR equation is capable only of assessing present land-use change potential, hence the multitude of urban growth studies that combine LR equations with CA to iterate all grid cell states using the latter's constraint and neighborhood effect transition rules [71,75,99]. Moreover, LR models do not

possess anything similar to a constraint transition rule, meaning that grid cells of a suitability surface are not immediately assigned a probability of zero if, for instance, any of the default constraints in Table 8 are violated.

Each of the WiFSS modeling approaches discussed here has at least one unique feature that equips said model to address specific needs regarding decision-making and system understanding [28]. GIS-MCDA is equipped to manage high-resolution secondary datasets that also allow locations unsuitable for wind energy development to be constrained and mapped. Non-GIS-MCDA is best for performing in-depth assessment of specific locations for potential wind energy development, subsequently ranking said locations from highest to lowest development priority. BNs are most useful when model users wish to understand how predictors relevant to the wind farm siting process influence each other, as well as how the influence of each individual predictor can be described probabilistically. LR combines GIS-MCDA's ability to produce maps of the most and least suitable locations for wind energy development with BNs' ability to aggregate both discrete and continuous datasets, making for a (theoretically) more comprehensive WiFSS assessment. The limitation that all of these WiFSS modeling approaches have in common is their lack of a temporal component, only being able to communicate the likelihood, acceptance, or suitability of installing wind farms at a single point in time. WiFSS-LRCA's niche among these other approaches is its integration of a Logistic Regression equation with Cellular Automata's transition rules [67]. Their integration allows the initial probabilities computed by an LR equation to be contextualized by violation of where wind farms cannot be installed (i.e., the constraint transition rule) and where wind energy development has already occurred (i.e., the neighborhood effect transition rule). Iterating these two rules, along with a

prescribed demand of gained wind farm capacity to drive land-use change [66,75], are ultimately what allow a Logistic Regression-Cellular Automata approach to project future wind farm locations, being beyond the capacity of existing approaches that predict present locations.

6.1.2. Question 2: What are currently the most suitable locations for present wind energy development across the Conterminous United States?

The first iteration of WiFSS-LRCA provides an assessment of present wind farm site suitability across the domain of interest, whether of the CONUS or an individual U.S. state, thus revealing the locations most and least suitable for wind energy development based on present conditions. Performing this assessment first required determining that WiFSS-LRCA could use its predictors (Tables 5a and 5b) to compute probabilities for each grid cell that correspond to the presence or absence of true commercial wind farms [73], through repeated calibration of the model's coefficients and validation of its overall performance (see Section 3.3). Furthermore, verifying that WiFSS-LRCA can construct suitability surfaces that are consistent with current wind farm siting practices makes the model's projections of future wind farm siting potential easier to trust in the hands of decision-makers [450,451]. One would also expect, however, that many high-probability locations for wind farm construction across the CONUS have not yet been developed, which WiFSS-LRCA identifies in its first iteration using its classification scheme (i.e.., the false positive grid cells in Figure 20). The results in Chapter 4 thus provide important context to the projections of future wind energy development potential presented in Chapter 5. This section explains the intended interpretation of WiFSS-LRCA's calibration and validation outputs, with the final two paragraphs concluding with how these interpretations connect to the suitability maps that convey the most and least suitable locations for present wind energy development.

WiFSS-LRCA contributes to a broader effort of SES model development that prioritizes combining quantitative and categorical datatypes and model runs at multiple spatial scales, among other priorities listed by Iwanaga et al. [471]. It was with these priorities in mind that WiFSS-LRCA was conceived to allow end-users to customize the model's spatial scale (CONUS or U.S. state), project capacity from 20 different grid cell sizes (see Table 6), and predictor configuration (see Section 3.5.2) of interest. WiFSS-LRCA thus accommodates experimentation of how suitable locations for wind energy development are affected by user priorities, while also meeting this dissertation's purpose (see Section 1.2) of preparedness for immediate end-use and serving as an educational tool about how end-user choices affect the modeled system's behavior [41]. Chapter 4 uses the customizability built into WiFSS-LRCA to demonstrate how its ability to construct robust suitability surfaces and identify grid cells that presently contain commercial wind farms are affected by these choices. The presented analysis shows that, when compared to a null model (intercept-only version of Equation 1), this model's LR equation performed statistically significantly well at classifying grid cells as (not) possessing commercial wind farms across multiple U.S. states (Table 10), based on assessment of WiFSS-LRCA's calibration (see Section 4.1.1). This assessment used a log-likelihood ratio statistic (λ, Equation 2) to assess how well the LR equation's coefficients fit to whether a study area's training grid cells contain wind farms [417], with the statistic maximizing under a *Reduced* configuration that refined the enlisted set of predictors (Section 3.5.2 explains the configuration's construction). This result held when running WiFSS-LRCA over most U.S. states and the CONUS, agreeing with Smith and McKenna's [452] expectation that well-fitted LR equations are those that statistically significantly improve over a null version of the same equation.

A greater log-likelihood ratio should mean that WiFSS-LRCA makes more accurate predictions of where the CONUS' present wind farms are located, but this dissertation also shows that this accuracy varied with both grid cell size and the number of present wind farms. Sensitivity to the latter was evident when validating this model's LR equation using ROC curves (see Section 4.2.1), with Figure 14 showing that there was a greater tendency for WiFSS-LRCA to incorrectly classify testing grid cells when fewer of them contained a wind farm, resulting in more frequent underperformance versus an AUC threshold for random chance of 0.5 [454]. This underperformance became less frequent when selecting a smaller grid cell size over the same study area, due to the same wind farm locations being split among smaller hexagons, suggesting that Linden's [457] observation of ROC's sensitivity to the low prevalence of a binary outcome also applies to WiFSS-LRCA. Conversely, in states containing a large number of present wind farms (such as Texas, Figure 13), the ROC curves constructed by WiFSS-LRCA portrayed a much greater rate of correct classification, showing that ROC curves are a more reliable means of validating WiFSS-LRCA's LR equation when more of the testing grid cells contain wind farms. Furthermore, based on the AUC statistics that accompany the ROC curves in Figure 13, the proportion of correctly classified grid cells maximized under the *Full* and *Reduced* predictor configurations. This result represents a reduction in the number of Type 1 errors (i.e., fewer false positive grid cells) compared to enlisting the *No_Wind* or *Wind_Only* configurations, and more importantly the value of constructing a predictor set that includes not only wind speed (a key predictor from the perspective of prior WiFSS studies [44,124,218]) but other predictors as well.

Whereas the ROC curves constructed by WiFSS-LRCA illustrate the proportion of (in)correctly classified testing grid cells at various probability thresholds, the Confusion Matrices explicitly quantify these rates of (in)correct classification and crucially the number of Type 1 and Type 2

errors (see Section 4.2.2). The Confusion Matrices presented in this dissertation again confirm the smaller proportion of false positive grid cell classifications that occur when using a *Full* or *Reduced* predictor configuration (Figure 15), with study areas that contain few commercial wind farms being vulnerable to a lower percentage of correct classifications. More interestingly, however, is how Confusion Matrices convey the LR equation's sensitivity to grid cell size, with Figure 16 showing a greater correct classification percentage at higher resolution due to fewer grid cells being classified as false positive. A higher resolution meant that more testing grid cells were likely to exist in both classes of the dependent variable (i.e., whether or not a grid cell contains a wind farm), meaning that the LR equation's assumption of linearity of the continuous predictor datasets was more likely to hold (see Section 3.3) [412]. Furthermore, more grid cells containing wind farms made it is less likely that the predictor datasets would all be associated with one outcome (e.g., almost no testing grid cells containing a wind farm), thus reducing the risk of quasi-complete separation and thus allowing for computation of finite log-likelihood ratios [413]. The validation of WiFSS-LRCA's performance in its first iteration is thus more trustworthy under the following conditions: a more complete (*Full*) or more refined (*Reduced*) predictor dataset is used, a smaller grid cell size is selected, and the chosen study area presently contains more commercial wind farms. As such, the calibrated fit produced by WiFSS-LRCA to its predictor datasets, and thus its recommended sites for wind energy development, are best for smaller wind farm projects. Published LR-based WiFSS assessments do not rely on classifying spatial units to validate their models' performance [58,59,179], though WiFSS-LRCA encapsulates the utility of ROC curves and Confusion Matrices for this very purpose.

In cases where these two validation metrics suggest WiFSS-LRCA's classification accuracy being lower, the ORs obtained from calibrating the model's LR equation should not be

overinterpreted. Although ORs strictly represent non-causal associations between the equation's predictors and whether a grid cell contains a wind farm [422], geographical influences on the final suitability surface can be inferred from their magnitudes. As exemplified by Figures 11 and 12 (see Section 4.1.2), certain predictors are consistently among those most strongly associated with grid cells being assigned high probabilities, namely Average Wind Speed (*Avg_Wind*), Distance to the Nearest Transmission Line (*Near_Trans*), Proportion of Undeveloped Land (*Undev_Land*), Average Elevation (*Avg_Elevat*), Interconnection *(Interconn*), and Total Number of Statewide Legislative Pieces (*Numb_Pols*). There are three caveats to using these ORs to suggest predictors' general importance to wind farm siting decisions. Firstly, the magnitude of the ORs obtained from training and testing WiFSS-LRCA's LR equation do not always agree with expectations. According to Table 7a and 7b, *Avg_Elevat* is expected to have a positive association because of wind speed increasing with height [348], as would *Interconn* due to the assumed positive influence on the political will to install renewable energy systems [384]. Figure 11 subverts these expectations, showing both predictors to have strong negative associations with grid cell outcomes in a model run over the CONUS, highlighting the imperfection of ORs as geographical indicators of WiFSS. Secondly, the importance of the predictors changes with the scale at which WiFSS-LRCA runs, such as *Avg_Wind* being the most strongly associated predictor in model runs over the CONUS (Figure 11) but not over Indiana alone (Figure 12), suggesting that there exist regional-scale differences in the wind farm siting roles played by predictors. Finally, some predictors are important (and thus have different ORs) in one U.S. state but not another, evidenced by comparing predictors in the *Reduced* configurations for California (Table 11) with those for Indiana (Figure 11, bottom-right), with *Undev_Land* and Public Support for Renewable Portfolio Standards (*supp_2018*) seldom retained in California's *Reduced*

predictor set. These differences in ORs across spatial scale and study area show that WiFSS-LRCA can differentially parameterize for spatial contexts of interest to the user. In other words, WiFSS-LRCA accounts for the varying importance of predictors depending on the study area of interest, hence its utility to decision-makers, emphasizing importance of spatially explicit modeling approaches that capture regional heterogeneities in the represented predictors.

These interpretations of the calibrated predictors and validation metrics provide key context to the suitability surfaces constructed in WiFSS-LRCA's first iteration. The examination of Figures 18 and 19 (see Section 4.3.2) showed that the high-probability grid cells across South and East Wisconsin were frequently part of statistically significant hotspots according to Getis-Ord statistics. These hotspots were explicable based on the strongly positive ORs of *Avg_Wind* and Military Installations (*Military*) and validated by around 85% of testing grid cell states being correctly classified by the model's LR equation. Since true positive grid cells frequently have the highest probability of containing wind farms across most study areas and grid cell sizes (e.g., Figure 17, see Section 4.3.1), the true positive grid cells within these hotspots suggest that the constructed WiFSS surface is reflective of current wind farm siting practices. If WiFSS-LRCA does not generate strongly associated ORs for each predictor that can be rationalized, and also does not perform well during its validation step, such interpretation of the final suitability surface should be more cautious. Furthermore, it was noted in Section 4.3.2 that false positive grid cells within hotspots may be candidates for future wind energy development, should true positive grid cells also exist among them. As such, given the United States' priorities of energy security [472] and decarbonization [473], a WiFSS surface with verifiable accuracy could convey future wind farm project sites. However, this deduction assumes that present wind farms are already in suitable locations, since these locations were used to both train and test WiFSS-LRCA (see

Section 3.3). This assumption may be inappropriate in U.S. states with previous wind farm siting controversies, such as bird mortalities in California's Altamont Pass [32] and social opposition to the abandoned Cape Wind project in Massachusetts [35]. These controversies may reflect temporal changes in important wind farm siting predictors [25], hence the value of projecting future wind farms in WiFSS-LRCA's subsequent iterations.

When examining Figure 21, the highest-probability locations for wind energy development (based on WiFSS-LRCA's first iteration) can be grouped into five broad regions: the Central Plains, Southern California, the Pacific Northwest, the Great Lakes, and the Northeastern United States. All five of these regions are currently populated by commercial wind farms [73], with runs of WiFSS-LRCA across the CONUS thus implying that these regions are indeed those most suitable for wind farm construction. As covered in Section 4.3.3, Figure 22 (top-left) shows these computed probabilities to be strongly associated with high wind speeds (*Avg_Wind*), closeness to transmission lines (*Near_Trans*), an abundance of undeveloped land (*Undev_Land*), legislation that supports wind energy (*Numb_Pols*), and ambitious RPS (*Renew_Targ*), all possessing ORs consistent with expectations (Table 7a and 7b). Some strongly associated predictors, however, have ORs that cannot be rationalized, particularly those for *Avg_Elevat* and *Gree_Lobbs*, though the five aforementioned predictor properties are generally shared among the five broad regions (except for *Avg_Wind* being relatively low in Southern California and the Pacific Northwest [321]). Beyond ORs that can mostly be rationalized, the constructed WiFSS surface for the CONUS is validated by over 80% of testing grid cells being correctly classified (Figure 22, bottom-left), and the correctly classified grid cells (true positive and true negative) having a statistically significantly different ($p < 0.05$) median probability than those that are incorrectly classified based on a Mann-Whitney U-test (Figure 22, bottom-right). Furthermore, Figure 20

shows that most grid cells classified as false positive, particulary those in the Central Plains and Great Lakes, are clustered around true positive cells, which Getis-Ord statistics often found to be statistically significant ($p < 0.05$). The verifiable accuracy of the suitability surface constructed by WiFSS-LRCA for the CONUS in Figure 21 means that its high-probability grid cells serve as indicators for wind energy development potential, particularly in the five broad regions.

These conclusions of suitable wind energy development regions across the CONUS should be made recognizing the limits to the interpretation of WiFSS surfaces. Figures 20 and 21 show that WiFSS-LRCA tends to predict greater probabilities and classify more grid cells as false positive in CONUS-level runs than in runs over individual U.S. states, particularly in regions greatly populated by wind farms (i.e., the Central Plains and Great Lakes). For instance, the CONUS-level run predicts high probabilities over almost all of Iowa, but the state-level run over Iowa confines high probabilities to the state's north, center, and west. Furthermore, CONUS-level model runs generate more false negative classifications outside of these populated regions than state-level runs do, suggesting less capability at capturing isolated commercial wind farm locations. A few reasons exist for WiFSS-LRCA's lower performance in CONUS-level model runs versus runs over individual U.S. states. Firstly, the predictors most pertinent to wind farm siting differ between regions, e.g., *Avg_Wind* was a less important predictor in Indiana (Figure 12) and political attitudes toward wind energy are more positive in some states [87,88]. Coefficients for each predictor in Equation 1 therefore represent more than a single U.S. state in CONUS-level model runs, changing the computed P_i values. Secondly, CONUS-level runs use up to 15 more predictors than state-level runs (Table 5a/5b), raising the risk of overfitting WiFSS-LRCA to the aggregated data [192] and potentially biasing computed probabilities. Finally, since WiFSS-LRCA is trained and tested based on presence of commercial wind farms

(see Section 3.3.1), this process is biased by most present wind farms being in two regions: the Central Plains and Great Lakes. Maximizing WiFSS-LRCA's goodness-of-fit thus reflects conditions in these regions moreso than regions further away with fewer wind farms [462], hence there being more true positives and fewer false negatives across the Central Plains and Great Lakes in CONUS-level model runs (Figure 20). For these reasons, although WiFSS-LRCA's performance in CONUS-level model runs has verifiable accuracy, state-level model runs generally possess greater accuracy since they can capture local differences in the predictors most pertinent to wind farm siting (see Section 4.1.2). It is thus recommended that CONUS-model runs be used to identify broad regions for wind energy development potential that cover multiple states, and that state-level runs of WiFSS-LRCA identify the highest-probability areas within these broader regions, as exemplified by comparing CONUS and state-level runs over Iowa and West Virginia in Figure 21.

6.1.3. Question 3: Which regions of the CONUS (at nationwide and state-level scales) are projected to acquire wind farms out to the year 2050, and what geographical features may explain these projections?

The first iteration of WiFSS-LRCA uses its Logistic Regression equation to predict present suitability for wind energy development based on the states of its enlisted predictors. Whereas, the subsequent iterations project future locations for wind energy development out to the year 2050, using both customizable scenarios that modify the LR equation's predictors and the constraint and neighborhood effect transition rules of Cellular Automata to replicate and continue observed patterns and limitations on this development. The decision to combine an LR equation and CA transition rules for these projections is predicated on two core assumptions. Firstly, that neighborhood effects indeed influence preferred installation sites of commercial wind farms, and

that a CA's neighborhood effect transition rule is applicable on state- to CONUS-wide spatial scales. Previous wind farm installations have clustered in common locations across the CONUS, such as the Texas Panhandle and south of Lake Michigan [73], suggesting commonly suitable locations for wind energy development, hence WiFSS-LRCA incorporates neighborhood effects (alongside constraints and scenario setups) to drive model projections. Secondly, that the idea of using neighborhood effects in CA models to modify probability of a grid cell's land-use change is equally applicable to commercial wind farms installed kilometers apart as it is to previous smaller-scale applications to urban growth [71,76], local deforestation [474] and landslides [475]. Verifying this application required a comprehensive sensitivity analysis (see Section 5.2), demonstrating the influence of modifying WiFSS-LRCA's neighborhood effects compared to the model's other parameters and determining whether the neighborhood effect parameter influences cluster-like land-use change patterns in a manner similar to other CA applications [72]. This section thus uses the results from Chapter 5 to summarize how this dissertation met its purpose (see Section 1.2) of showing that LRCA models are applicable in a WiFSS context, while also showing that geographical reasoning can be provided for the generated model outputs.

The maps presented throughout Chapter 5 accredit the application of WiFSS-LRCA to informing planning efforts for future wind farm projects. The high-probability regions identified in WiFSS-LRCA's first iteration (Figure 21), particularly the Central Plains and Great Lakes, were consistently projected to experience the most wind farm expansion across the CONUS (Figure 23a), areas that are already highly populated by wind farms. In CONUS-level model runs, this projection within broad regions held true whether or not the neighborhood effect transition rule was activated (Figure 23b), the only difference being that switching this rule off enabled projected locations to appear further away from present-day clusters. Neighborhood effects are

crucial, however, when running WiFSS-LRCA over individual U.S. states. Section 4.3.2 discussed how Getis-Ord statistics revealed that grid cells classified as false positive in WiFSS-LRCA's first iteration were frequently clustered around those that are true positive (Figure 19), thus perhaps representing candidate locations for future wind energy development. Figure 25 showed that these clustered false positive grid cells were frequently those projected to gain commercial wind farms in WiFSS-LRCA's subsequent iterations, a result that can be captured thanks to how neighborhood effects simulate the influence area of grid cells (i.e., the influence of common conditions for wind energy development in the same location) [437]. Micrositing concerns such as downwind turbulence and reduced energy generation necessitate building wind turbines in the same farm up to kilometers apart [100], hence WiFSS-LRCA has a larger spatial unit for defining a neighborhood than other LRCA applications, which often define tracts of land or groups of buildings tens [99] to hundreds [71] of meters across. This difference in scale and density of land-use change does not, however, seem to effect WiFSS-LRCA's ability to make projections in regions that are both 1) of high probability according to the model's LR equation; and 2) consistent with previous patterns of wind energy development across the CONUS, accrediting the use of neighborhood effects in this model.

Trust in the projections made by WiFSS-LRCA further comes from the model's response to different predictor configurations, specifically how the predictors incorporated into the model's LR equation affected these projections. Much like previous WiFSS studies that have noted Wind Speed to be an important (if not the most important) predictor for wind farm siting decisions [44,124,218], WiFSS-LRCA evidenced the same conclusion, with Figures 26 and 27 both showing that excluding Wind Speed affected the locations of future wind energy development clusters. Using ORs as evidence for the association between model predictors and model outputs,

the roles of predictors besides Wind Speed also became apparent. Much like in WiFSS-LRCA's first iteration, Figure 23 suggests that *Avg_Wind*, *Numb_Pols, Renew_Targ.* and *Undev_Land* are all strongly associated with future wind farm locations in CONUS model runs, regardless of applied scenarios modifying the relationships that these predictors possess (see Table 9). However, the importance of other predictors based on their associations changed when running WiFSS-LRCA over individual U.S. states. Figure 24a suggested that projected wind energy development in Florida frequently clustered around existing infrastructure and potentially in areas with younger demographics. Additionally, projections over Illinois in Figure 24b pointed to projected wind farm locations often being those not overlapping vulnerable wildlife areas, again not predictors with the strongest association in CONUS-level model runs. The spatial scale of WiFSS-LRCA's application therefore affects the geographical interpretation of its projections, which decision-makers could use to their advantage. Specifically, end-users of WiFSS-LRCA that may be concerned with wind energy development's response to these wildlife areas [341, 351] and demographics [339,398] could use WiFSS-LRCA to evaluate these responses, thus identifying localized influences on future wind farm siting decisions.

Geographical explanations for trends in the projections constructed by WiFSS-LRCA were also apparent when conducting its sensitivity analysis, while simultaneously illustrating the model's response to its four parameters: grid cell size, neighborhood size, constraint definitions, and scenario setup (see Section 5.2). Adjusting the prescribed neighborhood size (i.e., increasing the value of n in Equation 5 to allow grid cells further away to influence the state of grid cell i [68]) had arguably one of the greatest influences on WiFSS-LRCA's projections. Projections performed over Maine (Figure 29, see Section 5.2.2) showed future wind farms to appear earlier and in denser clusters to the southeast as neighborhood size was increased, an area of Maine with

high WiFSS likely due to relatively flat terrain [79] and dense transmission line infrastructure [326]. Modifying grid cell size (Figure 28, see Section 5.2.1) similarly influenced a clustering of future wind farm locations, since smaller grid cells in CA models have been shown to project land-use change over shorter distances [443]. Modifying these two parameters reaffirms WiFSS-LRCA's ability to project continued cluster-like patterns of wind energy development that have previously occurred across the CONUS [73]. It was also possible to infer geographical explanations from altering WiFSS-LRCA's constraints. As shown in Figure 30, many grid cells over Central Montana were only projected to gain commercial wind farms once the Wildlife Refuges constraint had been disabled (Figure 30b), suggesting a prevalence of protected land across the state that limits wind energy development [342]. Based on computed QADI statistics for these three parameters [445], adjustments to neighborhood size and grid cell size had the stronger influence on disagreement between separate runs of WiFSS-LRCA, but only adjustments to the constraint transition rule could reduce this statistic's quantity disagreement component (rather than just allocation disagreement).

As for the scenario setup parameter, regardless of the scenario that was enlisted when running WiFSS-LRCA over Washington (Figure 31), future wind energy development remained concentrated to the state's east side. Possible geographical interpretations for these projections include, based on computed ORs, greater Average Wind Speed [321] and few competing land uses (Wildlife Refuges, National Parks, Military Operations) compared to elsewhere in Washington. However, WiFSS-LRCA's projections were demonstrably less sensitive to scenario setup than the other three parameters discussed here. According to Figure 31, scenario modifications altered the timing more so than the location of when grid cells over Washington gained wind farms in subsequent iterations, with the other three parameters impacting both

timing and location of these projections. This smaller influence of modifying scenarios on WiFSS-LRCA's projected wind farm locations could be due to other parameters superseding the scenarios' effects. For instance, increasing neighborhood size and loosening the default constraints (Table 8) were both observed to cause tighter clustering of future wind farms around existing ones (Figures 28 and 30), resulting in the model's projections more closely resembling the outputs of a null model (i.e., a lower QADI score occurred). In other words, grid cells having a high P_i value becomes less important if WiFSS-LRCA is able to cluster future wind farms around present ones (i.e., looser constraints and larger neighborhoods are set as the model's parameters), making scenario modifications to predictor coefficients less influential. Much like in existing WiFSS studies and other studies that enlisted LRCA approaches, WiFSS-LRCA is demonstrably sensitive to how its neighborhood effects [437,444], constraints [226,435], and grid cell size [443] are defined. Moreover, the repeated training and testing of this model's LR equation (see Section 3.3.1) mitigates, but cannot negate, the effect of randomness on its final projections [421]. This randomness and sensitivity are two reasons why deduction of the geographical processes that influence these projections should not be overstated, because in addition to ORs not being a causal depiction of relationships between predictors and the model outcome [422], the model's development choices also influence their interpretation.

When running WiFSS-LRCA at the CONUS-level (as in Figure 23), the locations in which wind farm construction is projected to occur follow the most suitable regions identified in the model's first iteration (Figure 21). Two regions in particular are the Central Plains and Great Lakes, regions characterized by conditions suited for wind energy development (e.g., strong winds, undeveloped land, transmission line proximity, renewable energy legislation), substantiated by the OR associations obtained from initially calibrating the model's LR equation. As discussed in

Section 6.1.2, however, training and testing of this equation is biased by the many commercial wind farms that exist in these two regions, meaning grid cells that are in fact suitable for development in other parts of the CONUS may not gain a wind farm during WiFSS-LRCA's iterations. Furthermore, CONUS-level model runs only account for nationwide wind energy capacity targets (see Section 3.4) [92], hence the value of also running WiFSS-LRCA over individual U.S. states, both to identify local geographical influences on projections and to represent state-level capacity targets. The model requirement of at least two grid cells containing a wind farm to train and test WiFSS-LRCA at the state level (see Section 3.3.2) means that U.S. states currently lacking wind farms gain them based on ORs identified from CONUS model runs, hence states like Florida (Figure 24a) and Kentucky (Figure 27) are subject to the same equation-based transition rules. The constraint and neighborhood effect transition rules of the model's CA component result in projections of future wind farms that cluster in proximity to present wind farms (Figure 23, 29, and 30), thereby continuing the CONUS's ongoing pattern of clustered wind energy development. WiFSS-LRCA is, however, sensitive to how these transition rules, other parameters (grid cell size, scenarios), and predictor configurations are defined, with this sensitivity occurring in similar ways to LRCA models developed for other land-use change contexts. This dissertation ultimately represents a proof of concept in integration of Cellular Automata transition rules and a Logistic Regression equation on a spatial scale larger than that of previous work, within a model that incorporates a robust sensitivity analysis that both verifies its own predictive accuracy and enables experimentation in the hands of end-users.

6.2. Limitations in WiFSS-LRCA's Construction and Application.

In answering these three research questions, WiFSS-LRCA also fulfills the purposes set out at the beginning of this dissertation (see Section 1.2). The future wind farm locations projected by WiFSS-LRCA have been shown to change considerably depending on the model setup, e.g., grid cell size, predictor configuration, study area, constraints, etc., with these projections explicable based on results of its calibration, meeting the model's first purpose. WiFSS-LRCA's incorporation of constraint and neighborhood effect transition rules to project future wind farm sites addresses the second purpose, since these projections successfully continue ongoing patterns of commercial wind farm construction across the CONUS while also illustrating development opportunities in states presently lacking wind farms. The third purpose has been met thanks to WiFSS-LRCA being constructed to allow end-users to customize its parameters and other features (e.g., neighborhood size, gained wind power capacity, scenario setup) to serve their own interests. WiFSS-LRCA can thus function as a decision-making tool to inform future wind farm construction strategy over the next 30 years. Furthermore, the dataset aggregation, calibration, validation, and grid cell iteration techniques presented in this dissertation could be used to inform LRCA-based WiFSS analyses in other countries. However, there are a few limitations to the transferability of WiFSS-LRCA into other geographical contexts moving forward. Some of these limitations have been discussed throughout this dissertation, specifically the non-causal nature of ORs warning against their over-interpretation as geographical indicators (see Section 6.1.2), and the inability to train and test the model's LR equation over individual U.S. states that presently lack commercial wind farms (see Section 3.3.2). Other important limitations in WiFSS-LRCA's construction and application are addressed in this section.

Since an LRCA model had not previously been developed for performing WiFSS assessment, there existed few common standards to inform WiFSS-LRCA's development, with there being three prominent examples. Firstly, while LR-based WiFSS studies like Harper et al. [62] have proposed a hierarchical combination of predictors to construct predictor configurations, WiFSS-LRCA's *Reduced* configuration (see Section 3.5.2) lacked precedent and sought basis from wireless signal quantization [447]. Secondly, Variance Inflation Factor (VIF) cutoffs in LR-based WiFSS studies lack consistency. While WiFSS-LRCA used a value of 10 as a recommended upper limit [411], other studies have used a more conservative VIF of 2.5 [63] or did not specify their cutoff [62]. Thirdly, although this dissertation surveyed literature to define WiFSS-LRCA's default constraints (Table 8), these constraints frequently lacked consensus in the literature, such as prescribed maximum distance to transmission lines ranging from 2,000 meters [218] to 40,000 meters [49]. Each of these components of WiFSS-LRCA impacts its modeled outputs; predictor configurations and the VIF cutoff impact the predictors retained by Equation 1, and the default constraints impact which grid cells are unsuitable for future wind energy development. This lack of standards for model development extends to modeling approaches beyond WiFSS-LRCA, as discussed within the systematic review in Section 2.2 [40], such as the lack of common classification vocabulary, the varying application of predictors for constraint or evaluation, and inconsistent dataset citation methods. This dissertation serves in part to highlight the benefit of standardizing aspects of WiFSS modeling approaches, particularly predictor selection and representation, to ultimately synergize the discipline into one with consistent goals.

As mentioned in Section 3.3.1, the effects of randomness on WiFSS-LRCA's projected wind farm locations cannot be eliminated. During WiFSS-LRCA's first iteration, 75% of grid cells are used to train the LR equation (i.e., calibrate its coefficients to maximize the predictors' goodness-

of-fit to the locations of wind farms [415]) and the other 25% are used to test the equation (i.e., validate the fitted equation's ability to correctly classify grid cells as containing a wind farm or not). Random sampling of grid cells is accounted for by repeating the training and testing process 30 times [421] to derive median fitted coefficients and to obtain a median validated performance. Random effects on WiFSS-LRCA's outputs can persist regardless, especially when enlisting the *Reduced* predictor configuration because of how the sampled grid cells affect its refined combination of predictors (Table 11, see Section 4.1.3). Each run of WiFSS-LRCA produces slightly different results from a previous run because of this random grid cell sampling, regardless of using the same parameters, predictor configuration, and gained wind farm capacity. A possible solution is to bring spatial stratification into the random sampling of training and testing grid cells, such that balanced proportions of grid cells containing wind farms are extracted from strata covering a study area [476]. The benefit of spatial stratification would be evident when running WiFSS-LRCA over large domains, such as the CONUS or other countries. Doing so ensures that training and testing grid cell samples better represent a study area's geography, preventing random chance from sampling grid cells where most wind farms exist.

While experimenting with geographic stratification of predictor data is beyond this dissertation's scope, not doing so is partly why CONUS-level runs of WiFSS-LRCA are biased to project future wind energy development in areas populated by wind farms, regardless of whether neighborhood effects are active (Figure 23a and 23b). WiFSS-LRCA's training and testing grid cell samples must contain a proportional number of grid cells that possess a wind farm [416], ensuring that fitted coefficients reflect both presence and absence of wind farms (see Section 3.3.1). However, lack of spatial stratification means that grid cells with wind farms are sampled where most are found: the Central Plains and Great Lakes. Moreover, as discussed in Section

6.1.2, sampling grid cells from the same region means that WiFSS-LRCA's predictor coefficients are fitted to the predictor values associated with grid cell states in these regions, hence the large number of false positive classifications across the Central Plains and Great Lakes in Figure 20 (and the large number of false negatives outside of these two regions). Spatial stratification would address this problem by ensuring that grid cells containing wind farms are extracted proportionately across the CONUS, meaning that WiFSS-LRCA is trained and tested to recognize conditions suited for wind energy development outside of these populated regions. Another solution would be using state-level wind energy capacity targets [92] in CONUS-level model runs, to place regional limits on how many grid cells gain wind farms per iteration.

The interpretation that a grid cell containing a wind farm is related to geographical conditions depends on whether there is a statistically significant ($p < 0.05$) difference in the probabilities of grid cells classified as true/false positive and true/false negative. Section 4.3.1 summarized WiFSS-LRCA's use of Mann-Whitney U-tests to determine whether a statistically significant difference in ranked probabilities of grid cells exists between these pairs of classifications [430], which was the case for the CONUS (Figure 17) and indeed most study areas and grid cell sizes. A lack of statistical significance would caution against, for instance, asserting that true positive grid cells are those that possess higher wind speeds or are closer to transmission lines. More importantly, given the consensus that statistical tests involving differences in two groups' central tendencies require at least 30 sample members in each group [477], the results of WiFSS-LRCA's Mann-Whitney U-test may not always be robust. It was shown in Section 4.2 that false positive grid cell classifications (Type 2 errors) are produced infrequently by WiFSS-LRCA, hence the interpretation of differences in the probabilities of grid cells classified as true negative versus false negative should be made cautiously.

The method of defining scenarios for WiFSS-LRCA had little effect on the model's outputs. Section 5.2.4 showed that WiFSS-LRCA's projections are less sensitive to the scenario setup parameter than to neighborhood size, grid cell size, or the specified constraints, more so affecting the timing than location of future wind energy development. Modifying the constraints and neighborhood effects superseded modifying the coefficients of the model's LR equation, since a grid cell that violates constraints, or is too far away from grid cells that currently contain wind farms, are instantly assigned a $Prob_i^t$ value of zero (Equation 4), regardless of the closeness of the P_i value to one (Equation 1). WiFSS-LRCA's scenario approach of modifying these coefficients assumes that grid cells projected to gain wind farms are partially in response to changes in associations (Odds Ratios) between predictors and the binary dependent variable. It is established practice to use ORs to express these associations [62], and studies such as Yang et al. [98] also constructed scenarios that modified groups of predictors in an LRCA model (Tables 9a and 9b). However, this scenario definition compromises the maximized goodness-of-fit from calibrating WiFSS-LRCA's LR equation, and understanding the scenario outputs is subject to confounding predictors that can obscure OR interpretations [96]. Alternative scenario definitions in other LRCA studies represent potential to refine the definition used by WiFSS-LRCA. For instance, Gomes et al. [478] prescribed groups of drivers deemed important by local farmers (e.g., population density, road networks, land slope) to construct bespoke scenarios for their model's equation-based transition rule, and Mirbagheri and Alimohammadi [479] used Geographically Weighted Logistic Regression to weight global and local probabilities of land-use change differently to depict two respective scenarios of urban development.

A final key limitation is that WiFSS-LRCA has been optimized for the projection of onshore future wind energy development, rather than offshore development. Given the investment tax

credits [298], development plans [480], and capacity targets [89] recently released by the United States' federal government to support offshore wind energy, offshore WiFSS models are in demand. Indeed, non-LR-based offshore WiFSS models have already been developed for Egypt [262], Brazil [265], and South Korea [272]. However, with only two wind farms currently installed on the Atlantic Seaboard [73], running WiFSS-LRCA in the CONUS' surrounding waters would validly represent the predictor conditions of a relatively small region. Furthermore, assessing offshore wind energy development would require changes to WiFSS-LRCA's enlisted predictors. Proportion of Rugged Land (*Prop_Rugg*) and Proportion of Undevelopable Land (*Undev_Land*) would be obsolete in an offshore WiFSS assessment, and datasets representing ocean depth and offshore bird/fish habitat ranges would likely need to be included among WiFSS-LRCA's aggregated data. Under WiFSS-LRCA's current method of training and testing its LR equation (see Section 3.3), extending the model to project offshore wind farm locations would require separate model runs for onshore and offshore WiFSS assessment due to the differences in the necessary predictors.

6.3. Possible Directions for Future Work.

Acknowledging the above limitations does not discredit the research contributions made by WiFSS-LRCA, namely the development of a SES model that produces temporally explicit, geographically sound projections of future locations for wind energy development for a given study area, and the verified application of LRCA models on scales considerably larger than their more common city-to-county scale assessments of land-use change. As such, there are three directions of interest for expanding WiFSS-LRCA's use in future work, the first of which being the projection of offshore wind energy development. The United States' ongoing interest in increasing its offshore wind energy capacity means that WiFSS models that can inform where

this development should be focused would be of use. Amending WiFSS-LRCA to meet this demand for knowledge of offshore WiFSS should be straightforward for two reasons. Firstly, despite the lack of present offshore wind farms in the United States' surrounding waters [73], the Bureau of Ocean Energy Management maintains activity records for the nation's planned and completed offshore wind projects [481]. The locations of these projects could be used to define dummy offshore wind farm locations to train and test WiFSS-LRCA in the manner described in Section 3.3, allowing the model to project suitable offshore sites that have not yet been planned. Secondly, predictors that have been included in published offshore WiFSS studies, such as Ocean Depth [267], Distance to Shipping Lanes [268], and Distance to Commercial Fishing Areas [262], are available for the United States' surrounding waters as public domain data [482,483,484]. These datasets are therefore readily integrated into WiFSS-LRCA's dataset aggregation (see Section 3.2), allowing the model to project offshore WiFSS using the same hexagonal grid cell setup. Integration of offshore WiFSS assessment into WiFSS-LRCA could also expand the devised scenarios in Table 9 to include one for changes to offshore-specific predictors, along with new default constraints (Table 8), e.g., a maximum Ocean Depth of 60 meters for fixed-bottom wind turbine construction [261,480].

Another intended direction for future work is a more detailed analysis of how the spatial scale of WiFSS-LRCA impacts both projected wind farm locations and the geographical interpretation of said projections. This dissertation showed that running WiFSS-LRCA at a state-level or CONUS-level scale impacted the classified grid cell states in WiFSS-LRCA's first iteration (e.g., Figure 20) and the final projected wind farm locations (e.g., Figure 23 and 24). Additionally, state-level model runs generally allowed for greater predictive accuracy and identification of predictors that are less strongly associated in CONUS-level model runs (e.g., *Critical* and *Bat_Count* in Figure

24b). However, since individual U.S. states vary greatly in size and have borders based on a combination of river courses and latitude-longitude lines [485], alternative spatial scale definitions may reveal previously undetected patterns in the influence of predictors on wind farm siting decisions. For example, running WiFSS-LRCA over multiple states simultaneously may reveal which predictors have strong associations with model outputs across larger regions. Another option may be to construct spatial scales by using the predictors to define geographic regimes, e.g., predominant regional land types using the National Land Cover Database [80] or regional wind speed classes based on the National Renewable Energy Laboratory's wind resource maps [321]. Constructing study areas based on overarching geographic conditions may also provide more context to the ORs computed by WiFSS-LRCA, in this case for *Undev_Land* and *Avg_Wind*, respectively. Such definition of spatial scale as broad regions encompassing multiple states would also allow WiFSS-LRCA to be trained and tested over individual U.S. states that do not presently contain commercial wind farms. Section 3.3.2 stated that training and testing the model's LR equation requires both classes of the dependent variable to possess at least one grid cell that contains a wind farm [416], meaning 14 U.S. states rely on training WiFSS-LRCA over the CONUS to produce model projections. Running WiFSS-LRCA over multiple neighboring states, or basing study areas on broad geographic regimes, would allow the model's predictive accuracy in its first iteration to be verified over these 14 states.

Finally, the potential applications of WiFSS-LRCA go beyond projecting wind energy development. Aggregation of WiFSS-LRCA's dataset across multiple grid cell sizes and study areas means that it is readily applicable to suitability analyses of other types of decentralized energy, particularly solar energy. Indeed, many predictors relevant to WiFSS are also relevant to commercial solar farm site suitability, such as transmission line proximity, population density,

and ruggedness of terrain, predictors that have been used in SES modeling studies of solar farm siting potential [45,127]. The only essential additions to WiFSS-LRCA's aggregated dataset would be locations of present commercial solar farms and solar irradiance, with datasets for both existing in the United States' public domain [486,487]. Furthermore, any form of decentralized land-use change over the CONUS of interest to this model's end-users could be projected with WiFSS-LRCA, given the necessary substitution of the binary dependent variable. Despite the limitations previously discussed, WiFSS-LRCA's verified predictive accuracy of both present and future wind farm locations across the CONUS mean it is prepared for application in other contexts, expanding beyond the more common land-use change applications of LRCA models.

References

[1] IRENA. Renewable Capacity Statistics 2023, https://www.irena.org/Publications/2023/Mar/Renewable-capacity-statistics-2023; 2023 [accessed 4 May 2023].

[2] Ma L, Xu D. Toward Renewable Energy in China: Revisiting Driving Factors of Chinese Wind Power Generation Development and Spatial Distribution. Sustainability 2021;13:13pp. https://doi.org/10.3390/su13169117.

[3] Wind Europe. Wind Energy in Europe, 2020 Statistics and the Outlook for 2021-2025, https://s1.eestatic.com/2021/02/24/actualidad/210224_windeurope_combined_2020_stats.pdf; 2021 [accessed 4 May 2023].

[4] American Clean Power. Wind power facts. https://cleanpower.org/facts/wind-power/; 2023 [accessed 4 May 2023].

[5] Abbas Q, Khan AR, Bashir A, Alemzero DA, Sun H, Iram R, et al. Scaling up renewable energy in Africa: measuring wind energy through econometrics approach. Environ Sci Pollut Res 2020;27:36282- 94. https://doi.org/10.1007%2Fs11356-020-09596-1.

[6] New Zealand Ministry of Business, Innovation & Employment. New Zealand Energy Strategy – Terms of Reference, https://www.mbie.govt.nz/dmsdocument/25373-terms-of-reference-new-zealand-energy-strategy; 2022 [accessed 4 May 2023].

[7] BP. Reimagining energy for people and our planet – bp sustainability report 2022, https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/sustainability/group-reports/bp-sustainability-report-2022.pdf; 2022 [accessed 4 May 2023]

[8] Repsol. Repsol Buys Renewables Assets in Chile, https://www.repsol.com/content/dam/repsol-corporate/en_gb/sala-de-prensa/documentos-sala-de-prensa/2020/PR23072020_repsol_buys_renewables_assets_in_chile_tcm14-198321.pdf; 2020 [accessed 4 May 2023].

[9] Barthelmie RJ, Pryor SC. Climate Change Mitigation Potential of Wind Energy. Climate 2021;9:22pp. https://doi.org/10.3390/cli9090136.

[10] Nugent D, Sovacool BK. Assessing the lifecycle greenhouse gas emissions from solar PV and wind energy: A critical meta-survey. Energ Pol 2014;65:229-44. https://doi.org/10.1016/j.enpol.2013.10.048.

[11] Wang S, Wang S, Liu J. Life-cycle green-house gas emissions of onshore and offshore wind turbines. J Clean Prod 2019;210:804-10. https://doi.org/10.1016/j.jclepro.2018.11.031.

[12] Saidur R, Rahim NA, Islam MR, Solangi KH. Environmental impact of wind energy. Renew Sust Energ Rev 2011;15:2423-30. https://doi.org/10.1016/j.rser.2011.02.024.

[13] Panwar NL, Kaushik SC, Kothari S. Role of renewable energy sources in environmental protection: A review. Renew Sust Energ Rev 2011;15:1513-24. https://doi.org/10.1016/j.rser.2010.11.037.

[14] Owusu PA, Asumadu-Sarkodie S. A review of renewable energy sources, sustainable issues and climate change mitigation. Cogent Eng 2016;3:15pp. https://doi.org/10.1080/23311916.2016.1167990.

[15] Zafirakis D, Chalvatzis KJ. Wind energy and natural gas-based energy storage to promote energy security and lower emissions in island regions. Fuel 2014;115:203-19. https://doi.org/10.1016/j.fuel.2013.06.032.

[16] Hamed TA, Bressler L. Energy security in Israel and Jordan: The role of renewable energy sources. Renew Energ 2019;135:378-89. https://doi.org/10.1016/j.renene.2018.12.036.

[17] Alsharif MH, Kim J, Kim JH. Opportunities and Challenges of Solar and Wind Energy in South Korea: A Review. Sustainability 2018;10:23pp. https://doi.org/10.3390/su10061822.

[18] Duffy A, Hand M, Wiser R, Lantz E, Riva AD, Berkhout V, et al. Land-based wind energy cost trends in Germany, Denmark, Ireland, Norway, Sweden and the United States. Appl Energ 2020;277:14pp. https://doi.org/10.1016/j.apenergy.2020.114777.

[19] Jacobsen HK, Hevia-Koch P, Wolter C. Nearshore and offshore wind development: Costs and competitive advantage exemplified by nearshore wind in Denmark. Energ Sustain Dev 2019;50:91-100. https://doi.org/10.1016/j.esd.2019.03.006.

[20] Bosch J, Staffell I, Hawkes AD. Global levelised cost of electricity from offshore wind. Energ 2019;189:13pp. https://doi.org/10.1016/j.energy.2019.116357.

[21] Smirnova E, Kot S, Kolpak E, Shestak V. Governmental support and renewable energy production: A cross-country review. Energ 2021;230:11pp. https://doi.org/10.1016/j.energy.2021.120903.

[22] Eitan A, Herman L, Fischhendler I, Rosen G. Community-private sector partnerships in renewable energy. Renew Sust Energ Rev 2019;105:95-104. https://doi.org/10.1016/j.rser.2018.12.058.

[23] Hamilton LC, Bell E, Hartter J, Salerno JD. A change in the wind? US public views on renewable energy and climate compared. Energ Sustain Soc 2018;8:13pp. https://doi.org/10.1186/s13705-018-0152-5.

[24] Latinopoulos D, Kechagia K. A GIS-based multi-criteria evaluation for wind farm site selection. A regional scale application in Greece. Renew Energ 2015;78:550-60. https://doi.org/10.1016/j.renene.2015.01.041.

[25] Pavlowsky C, Koch J, Gliedt T. Place attachment and social barriers to large scale renewable energy development: a social-ecological systems analysis of a failed wind energy

project in the south-central United States. Socio-Ecol Pract Res 2023. https://doi.org/10.1007/s42532-023-00142-0.

[26] Hassan MM. Arsenic in Groundwater: Poisoning and Risk Assessment. 1st ed. CRC Press Taylor & Francis Group; 2018. https://doi.org/10.1201/9781315117034.

[27] Jakeman AJ, Lechter RA, Norton JP. Ten iterative steps in development and evaluation of environmental models. Environ Modell Softw 2006;21:602-14. https://doi.org/10.1016/j.envsoft.2006.01.004.

[28] Kelly RA, Jakeman AJ, Barreateau O, Borsuk ME, Elsawah S, Hamilton SH, et al. Selecting among five common modeling approaches for integrated environmental assessment and management. Environ Modell Softw 2013;47:159-81. https://doi.org/10.1016/j.envsoft.2013.05.005.

[29] Vasileiou M, Loukogeorgaki E, Vagiona DG. GIS-based multi-criteria decision analysis for site selection of hybrid offshore wind and wave energy systems in Greece. Renew Sust Energ Rev 2017;73:745-57. https://doi.org/10.1016/j.rser.2017.01.161.

[30] Vavatsikos AP, Arvanitidou A, Petsas D. Wind farm investments portfolio information using GIS-based suitability analysis and simulation procedures. J Environ Manage 2019;252:12pp. https://doi.org/10.1016/j.jenvman.2019.109670.

[31] Jørgensen ML, Anker HT, Lassen J. Distributive fairness and local acceptance of wind turbines: The role of compensation schemes. Energ Pol 2020;138:12pp. https://doi.org/10.1016/j.enpol.2020.111294.

[32] Smallwood KS, and Thelander C. Bird Mortality in the Altamont Pass Wind Resource Area, California. J Wildl Manag 2008;72:215-23. https://doi.org/10.2193/2007-032.

[33] Morinha F, Travassos P, Seixas F, Martins A, Bastos R, Carvalho D, et al. Differential mortality of birds killed at wind farms in Northern Portugal. Bird Study 2014;61:255-9. https://doi.org/10.1080/00063657.2014.883357.

[34] Kaldellis JK. Social attitude toward wind energy applications in Greece. Energ Pol 2005;33:595-602. https://doi.org/10.1016/j.enpol.2003.09.003.

[35] Phadke R. Steel forests or smoke stacks: the politics of visualisation in the Cape Wind controversy. Environ Polit 2010;19:1-20. https://doi.org/10.1080/09644010903396051.

[36] Sen S, Ganguly S. Opportunities, barriers and issues with renewable energy development – A discussion. Renew Sust Energ Rev 2017;69:1170-81. https://doi.org/10.1016/j.rser.2016.09.137.

[37] Ramírez-Rosado IJ, García-Garrido E, Fernández-Jiménez LA, Zorzano-Santamaría PJ, Monteiro C, Miranda V. Promotion of new wind farms based on a decision support system. Renew Energ 2008;33:558-66. https://doi.org/10.1016/j.renene.2007.03.028.

[38] Kabak M, Taşkinöz G. Determination of the Installation Sites of Wind Power Plants with Spatial Analysis: A Model Proposal. Sigma J Eng Nat Sci 2020;38:441-57. https://dergipark.org.tr/en/pub/sigma/issue/65119/1003377#article_cite.

[39] Wu Y, Zhang T, Xu C, Zhang B, Li L, Ke Y, et al. Optimal location selection for offshore wind-PV-seawater pumped storage power plant using a hybrid MCDM approach: A two-stage framework. Energ Convers Manage 2019;199:18pp. https://doi.org/10.1016/j.enconman.2019.112066.

[40] Wimhurst JJ, Nsude CC, Greene JS. Standardizing the factors used in wind farm site suitability models: A review. Heliyon 2023;9:21pp. https://doi.org/10.1016/j.heliyon.2023.e15903.

[41] Elsawah S, Filatova T, Jakeman AJ, Kettner AJ, Zellner ML, Athanasiadis IN, et al. Eight grand challenges in socio-environmental systems modeling. Socio Environ Syst Model 2020;2:34pp. https://doi.org/10.18174/sesmo.2020a16226.

[42] Mekonnen AD, Gorsevski PV. A web-based participatory GIS (PGIS) for offshore wind farm suitability within Lake Erie, Ohio. Renew Sust Energ Rev 2015;41:162-77. https://doi.org/10.1016/j.rser.2014.08.030.

[43] Spyridonidou S, Vagiona DG. Systematic Review of Site-Selection Processes in Onshore and Offshore Wind Energy Research. Energies 2020;13:26pp. https://doi.org/10.3390/en13225906.

[44] Baseer MA, Rehman S, Meyer JP, Mahbub Alam Md. GIS-based site suitability analysis for wind farm development in Saudi Arabia. Energ 2017;141:1166-76. https://doi.org/10.1016/j.energy.2017.10.016.

[45] Janke JR. Multicriteria GIS modeling of wind and solar farms in Colorado. Renew Energ 2010;35:2228-34. https://doi.org/10.1016/j.renene.2010.03.014.

[46] Sánchez-Lozano JM, García-Cascales MS, Lamata MT. GIS-based onshore wind farm site selection using Fuzzy Multi-Criteria Decision Making methods. Evaluating the case of Southeastern Spain. Appl Energ 2016;171:86-102. https://doi.org/10.1016/j.apenergy.2016.03.030.

[47] Villacreses G, Gaona G, Martínez-Gómez J, Jijón DJ. Wind farms suitability location using geographical information system (GIS), based on multi-criteria decision making (MCDM) methods: The case of continental Ecuador. Renew Energ 2017;109:275-86. https://doi.org/10.1016/j.renene.2017.03.041.

[48] Elkadeem MR, Younes A, Sharshir SW, Campana PE, Wang S. Sustainable siting and design optimization of hybrid renewable energy system: A geospatial multi-criteria analysis. Appl Energ 2021;295:35pp. https://doi.org/10.1016/j.apenergy.2021.117071.

[49] Saraswat SK, Digalwar AK, Yadav SS, Kumar G. MCDM and GIS based modeling technique for assessment of solar and wind farm locations in India. Renew Energ 2021;169:865-84. https://doi.org/10.1016/j.renene.2021.01.056.

[50] Bobeck M. A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales from a Sustainable Development Perspective [dissertation on the Internet]. Lund (SW): Lund University; 2017 [cited 2023 May 5]. Available from: https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8903253&fileOId=8903255 .

[51] Kaya T, Kahraman C. Multicriteria renewable energy planning using an integrated fuzzy VIKOR & AHP methodology: The case of Istanbul. Energ 2010;35:2517-27. https://doi.org/10.1016/j.energy.2010.02.051.

[52] Wang C-N, Huang Y-F, Chai Y-C, Nguyen VT. A Multi-Criteria Decision Making (MCDM) for Renewable Energy Plants Location Selection in Vietnam under a Fuzzy Environment. Appl Sci 2018;8:33pp. https://doi.org/10.3390/app8112069.

[53] Saaty TL. What is the Analytic Hierarchy Process? In Mitra G, Greenberg HJ, Lootsma FA, Rijkaert MJ, Zimmermann HJ, editors. Mathematical Models for Decision Support, NATO ASI Series (Series F: Computer and Systems Sciences) vol 48; 1988. Berlin: Springer; 1988. p. 109-121. https://doi.org/10.1007/978-3-642-83555-1_5.

[54] Watson JJW, Hudson MD. Regional Scale wind farm and solar farm suitability assessment using GIS-assisted multi-criteria evaluation. Landscape Urban Plan 2015;138:20-31. https://doi.org/10.1016/j.landurbplan.2015.02.001.

[55] Pamučar D, Gigović L, Bajić Z, Janošević M. Location Selection for Wind Farms Using GIS Multi-Criteria Hybrid Model: An Approach Based on Fuzzy and Rough Numbers. Sustainability 2017;9:23pp. https://doi.org/10.3390/su9081315.

[56] Mokarram M, Pourghasemi HR, Mokarram MJ. A multi-criteria GIS-based model for wind farm site selection with the least impact on environmental pollution using the OWA-ANP method. Environ Sci Pollut Res 2022:22pp. https://doi.org/10.1007/s11356-022-18839-2.

[57] Petrov AN, Wessling JM. Utilization of machine-learning algorithms for wind turbine site suitability modeling in Iowa, USA. Wind Energ 2014;18:713-27. https://doi.org/10.1002/we.1723.

[58] Mann D, Lant C, Schoof J. Using map algebra to explain and project spatial patterns of wind energy development in Iowa. Appl Geogr 2012;34:219-29. https://doi.org/10.1016/j.apgeog.2011.11.008.

[59] Foley B. Wind Turbine Siting in Maine. Atlas Maine 2018;2018:5pp. https://digitalcommons.colby.edu/atlas_docs/vol2018/iss2/4.

[60] Knoke JD. Discriminant Analysis with Discrete and Continuous Variables. Biometrics 1982;38:191-200. https://doi.org/10.2307/2530302.

[61] Malczewski J. GIS-based multicriteria decision analysis: a survey of the literature. Int J Geogr Inf Sci 2006;20;703-26. https://doi.org/10.1080/13658810600661508.

[62] Harper M, Anderson B, James PAB, Bahaj AS. Onshore wind and the likelihood of planning acceptance: Learning from a Great Britain context. Energ Pol 2019;128:954-66. https://doi.org/10.1016/j.enpol.2019.01.002.

[63] Roddis P, Carver S, Dallimer M, Norman P, Ziv G. The role of community acceptance in planning outcomes for onshore wind and solar farms: An energy justice analysis. Appl Energ 2018;226:353-64. https://doi.org/10.1016/j.apenergy.2018.05.087.

[64] Borunda M, de la Cruz J, Garduno-Ramirez R, Nicholson A. Technical assessment of small-scale wind power use in Mexico: A Bayesian intelligence approach. Plos One 2020;15:26pp. https://doi.org/10.1371/journal.pone.0230122.

[65] Li M, Liu K, Zhang R, Hong M, Pan Q. Using the Cloud-Bayesian Network in Environmental Assessment of Offshore Wind-Farm Siting. Math Probl Eng 2019;2019:16pp. https://doi.org/10.1155/2019/9710839.

[66] Barredo JI, Kasanko M, McCormick N, Lavalle C. Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape Urban Plan 2003;64:145-60. https://doi.org/10.1016/S0169-2046(02)00218-9.

[67] Douvinet J, Van De Wiel MJ, Delahaye D, Cossart E. A flash flood hazard assessment in dry valleys (northern France) by cellular automata modelling. Nat Hazards 2015;75:2905-29. https://doi.org/10.1007/s11069-014-1470-3.

[68] White R, Engelen G. Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. Environ Plann B 1997;24:235-46. https://doi.org/10.1068/b240235.

[69] Santé I, García AM, Miranda D, Crecente R. Cellular automata models for the simulation of real-world urban processes: A review and analysis. Landscape Urban Plan 2010;96:108-22. https://doi.org/10.1016/j.landurbplan.2010.03.001.

[70] González PB, Aguilera-Benavente F, Gómez-Delgado M. Partial validation of cellular automata based model simulations of urban growth: An approach to assessing factor influence using spatial methods. Environ Modell Soft 2015;69:77-89. https://doi.org/10.1016/j.envsoft.2015.03.008.

[71] Mustafa A, Heppenstall A, Omrani H, Saadi I, Cools M, Teller J. Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic 31 algorithm. Comput Environ Urban 2018;67:147-56. https://doi.org/10.1016/j.compenvurbsys.2017.09.009.

[72] He C, Zhao Y, Tian J, Shi P. Modeling the urban landscape dynamics in a megalopolitan cluster area by incorporating a gravitational field model with cellular automata. Landscape Urban Plan 2013;113:78-89. https://doi.org/10.1016/j.landurbplan.2013.01.004.

[73] Hoen BD, Diffendorfer JE, Rand JT, Kramer LA, Garrity CP, Hunt HE. United States Wind Turbine Database. United States Geological Survey, American Clean Power Association, and Lawrence Berkeley National Laboratory, vOct2022. https://eerscmap.usgs.gov/uswtdb.

[74] Liao J, Tang L, Shao G, Su X, Chen D, Xu T. Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. Environ Modell Soft 2016;75:163-75. https://doi.org/10.1016/j.envsoft.2015.10.014.

[75] Shu B, Bakker MM, Zhang H, Li Y, Qin Q, Carsjens GJ. Modeling urban expansion by using variable weights logistic cellular automata: a case study of Nanjing, China. Int J Geogr Inf Sci 2017;31:1314-33. https://doi.org/10.1080/13658816.2017.1283505.

[76] Shu B, Zhu S, Qu Y, Zhang H, Li X, Carsjens GJ. Modelling multi-regional urban growth with multilevel logistic cellular automata. Comput Environ Urban Syst 2020;80:12pp. https://doi.org/10.1016/j.compenvurbsys.2019.101457.

[77] Plassin S, Koch J, Paladino S, Friedman JR, Spencer K, Vaché KB. A socio-environmental geodatabase for integrative research in the transboundary Rio Grande/Río Bravo basin. Nature Sci Data 2020;7:14pp. https://doi.org/10.1038/s41597-020-0410-1.

[78] EIA. Electric Sales, Revenue, and Average Price – Average Annual Price. US Energy Information Administration. 2021. https://www.eia.gov/electricity/data/state/avgprice_annual.xlsx.

[79] United States Geological Survey. 3DEP Product Metadata – 1 arc-second resolution. USGS National Map Viewer Downloader. 2022. https://apps.nationalmap.gov/downloader/.

[80] Dewitz J. National Land Cover Database (NLCD) 2019 Products. United States Geological Survey Data Release, v2.0. 2021. https://doi.org/10.5066/P9KZCM54.

[81] United States Census Bureau. County Population by Characteristics: 2010-2019. United States Census Bureau. 2019. https://www.census.gov/data/tables/timeseries/demo/popest/2010s-counties-detail.html.

[82] United States Census Bureau Vintage 2019. County Population Totals: 2010-2019 – United States. United States Census Bureau. 2019. https://www.census.gov/data/tables/timeseries/demo/popest/2010s-counties-total.html

[83] Klink K. Trends in mean monthly maximum and minimum surface wind speeds in the coterminous United States, 1961 to 1990. Clim Res 1999;13:193-205. https://doi.org/10.3354/cr013193.

[84] Pryor SC, Barthelmie RJ, Young DT, Takle ES, Arritt RW, Flory D, et al. Wind speed trends over the contiguous United States. J Geophys Res-Atmos 2009;114:18pp. https://doi.org/10.1029/2008JD011416.

[85] Righter RW. Pioneering in wind energy: The California experience. Renew Energ 1996;9:781-4. https://doi.org/10.1016/0960-1481(96)88399-6.

[86] Haces-Fernandez F. Higher Wind: Highlighted Expansion Opportunities to Repower Wind Energy. Energies 2021;14:19pp. https://doi.org/10.3390/en14227716.

[87] Stokes LC, Warshaw C. Renewable energy policy design and framing influence public support in the United States. Nature Energ 2017;2:15pp. https://doi.org/10.1038/nenergy.2017.107.

[88] Olson-Hazboun SK, Howe PD, Leiserowitz A. The influence of extractive activities on public support for renewable energy policy. Energ Pol 2018;123:117-26. https://doi.org/10.1016/j.enpol.2018.08.044.

[89] The White House. FACT SHEET: Biden-Harris Administration Announces New Actions to Expand U.S. Offshore Wind Energy, https://www.whitehouse.gov/briefing-room/statements-releases/2022/09/15/fact-sheet-biden-harris-administration-announces-new-actions-to-expand-u-s-offshore-wind-energy/; 2022 [accessed 9 May 2023].

[90] United States Department of Energy. Guide to the Federal Investment Tax Credit for Commercial Solar Photovoltaics, https://www.energy.gov/sites/prod/files/2020/01/f70/Guide%20to%20the%20Federal%20Investment%20Tax%20Credit%20for%20Commercial%20Solar%20PV.pdf; 2020 [accessed 9 May 2023].

[91] Congressional Research Service. Residential Energy Tax Credits: Changes in 2023, https://crsreports.congress.gov/product/pdf/IN/IN12051; 2022 [accessed 9 May 2023].

[92] United States Department of Energy. Wind Vision: A New Era for Wind Power in the United States, https://www.energy.gov/sites/prod/files/WindVision_Report_final.pdf; 2015 [accessed 9 May 2023].

[93] Sommet N, Morselli D. Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. Int Rev Soc Psychol 2017;30:203-18. http://doi.org/10.5334/irsp.90.

[94] Hailpern SM, Visintainer PF. Odds Ratios and Logistic Regression: Further examples of their use and Interpretation. Stata J 2003;3:213-25. https://doi.org/10.1177/1536867X0300300301.

[95] Chen H, Cohen P, Chen S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. Commun Stat Simulat 2010;39:860-4. https://doi.org/10.1080/03610911003650383.

[96] Zhang J, Yu KF. What's the Relative Risk? A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. JAMA 1998;280:1690-1. https://doi.org/10.1001/jama.280.19.1690.

[97] Sohl TL, Sayler KL, Drummond MA, Loveland TR. The FORE-SCE model: a practical approach for projecting land cover change using scenario-based modeling. J Land Use Sci 2007;2:103-26. https://doi.org/10.1080/17474230701218202.

[98] Yang J, Liu W, Li Y, Li X, Ge Q. Simulating Intraurban Land Use Dynamics under Multiple Scenarios Based on Fuzzy Cellular Automata: A Case Study of Jinzhou District, Dalian. Complexity 2018;17pp. https://doi.org/10.1155/2018/7202985.

[99] Shahbazian Z, Faramarzi M, Rostami N, Mahdizadeh H. Integrating logistic regression and cellular automata-Markov models with the experts' perceptions for detecting and simulating land-use changes and their driving forces. Environ Monit Assess 2019;191:17pp. https://doi.org/10.1007/s10661-019-7555-4.

[100] Moorthy CB, Deshmukh MK, Mukherejee D. New Approach for Placing Wind Turbines in a Wind Farm Using Genetic Algorithm. Wind Eng 2014;38:633-42. https://doi.org/10.1260/0309-524X.38.6.633.

[101] Liao J, Tang L, Shao G, Qiu Q, Wang C, Zheng S, et al. A neighbor decay cellular automata approach for simulating urban expansion based on particle swarm intelligence. Int J Geogr Inf Sci 2014;28:720-38. https://doi.org/10.1080/13658816.2013.869820.

[102] Egelhofer JL, Lecheler S. Fake news as a two-dimensional phenomenon: a framework and research agenda. Ann Int Comm Assoc 2019;43:97-116. https://doi.org/10.1080/23808985.2019.1602782.

[103] Grimm V, Augusiak J, Focks A, Frank BM, Gabsi F, Johnston ASA. Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. Eco Model 2014;280:129-39. https://doi.org/10.1016/j.ecolmodel.2014.01.018.

[104] Scacchi W. Free/Open Source Software Development: Recent Research Results and Methods. Adv Comput 2007;69:243-95. https://doi.org/10.1016/S0065-2458(06)69005-0.

[105] Mergel I. Open collaboration in the public sector: The case of social coding on GitHub. Gov Inf Quarter 2015;32:464-72. https://doi.org/10.1016/j.giq.2015.09.004.

[106] Cowen DJ. GIS versus CAD versus DBMS: What Are the Differences? In: Peuquet DJ, Marble DF, editors. Introductory Readings in Geographic Information Systems; 1988. London: Taylor & Francis, 1988. p. 1551-55.

[107] Rybarczyk G, Wu C. Bicycle facility planning using GIS and multi-criteria decision analysis. Appl Geogr 2010;30:282-93. https://doi.org/10.1016/j.apgeog.2009.08.005.

[108] Rahmati O, Zeinivand H, Besharat M. Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis. Geomat Haz Risk 2016;7:1000-17. https://doi.org/10.1080/19475705.2015.1045043.

[109] Eastman JR. Multi-criteria evaluation and GIS. In Goodchild PA, Maguire MF, Rhind DW, editors. Geographical Information Systems; 1999. New York: John Wiley and Sons, 1999. p. 493-502.

[110] Ajanaku BA, Strager MP, Collins AR. GIS-based multi-criteria decision analysis of utility-scale wind farm site suitability in West Virginia. Geojournal 2021:23pp. https://doi.org/10.1007/s10708-021-10453-y.

[111] Sotiropoulou KF, Vavatsikos AP. Onshore wind farms GIS-Assisted suitability analysis using PROMETHEE II. Energ Pol 2021;158:14pp. https://doi.org/10.1016/j.enpol.2021.112531.

[112] Taoufik M, Fekri A. GIS-based multi-criteria analysis of offshore wind farm development in Morocco. Energ Convers Manage X 2021;11:11pp. https://doi.org/10.1016/j.ecmx.2021.100103.

[113] Rediske G, Burin HP, Rigo PD, Rosa CB, Michels L, Siluk JCM. Wind power plant site selection: A systematic review. Renew Sust Energ Rev 2021;148:13pp. https://doi.org/10.1016/j.rser.2021.111293.

[114] Shao M, Han Z, Sun J, Xiao C, Zhang S, Zhao Y. A review of multi-criteria decision making applications for renewable energy site selection. Renew Energ 2020;157:377-403. https://doi.org/10.1016/j.renene.2020.04.137.

[115] Baban SMJ, Parry T. Developing and applying a GIS-assisted approach to locating wind farms in the UK. Renew Energ 2001;24:59-71. https://doi.org/10.1016/S0960-1481(00)00169-5.

[116] Liang F, Brunelli M, Rezaei J. Consistency issues in the best worst method: Measurements and thresholds. Omega 2020;96:11pp. https://doi.org/10.1016/j.omega.2019.102175.

[117] Pavić Z, Novoselac V. Notes on TOPSIS Method. Int J Res Eng Sci 2013;1:5-12. https://www.ijres.org/papers/v1-i2/B120512.pdf.

[118] Saaty TL. What is the Analytic Hierarchy Process? In Mitra G, Greenberg HJ, Lootsma FA, Rijkaert MJ, Zimmermann HJ, editors. Mathematical Models for Decision Support, NATO ASI Series (Series F: Computer and Systems Sciences) vol 48; 1988. Berlin: Springer; 1988. p. 109-121. https://doi.org/10.1007/978-3-642-83555-1_5.

[119] Li M, Xu Y, Guo J, Li Y, Li W. Application of a GIS-Based Fuzzy Multi-Criteria Evaluation Approach for Wind Farm Site Selection in China. Energies 2020;13:19pp. https://doi.org/10.3390/en13102426.

[120] Gonzalez A, Enríquez-de-Salamanca Á. Spatial Multi-Criteria Analysis in Environmental Assessment: A Review and Reflection on Benefits and Limitations. J Environ Assess Pol Manage 2018;20:24pp. https://doi.org/10.1142/S146433321840001X.

[121] González-Ramiro A, Gonçalves G, Sánchez-Rios A, Su Jeong J. Using a VGI and GIS-Based Multicriteria Approach for Assessing the Potential of Rural Tourism in Extremadura (Spain). Sustainability 2016;8:15pp. https://doi.org/10.3390/su8111144.

[122] Wang F. Why public health needs GIS: a methodological overview. Annal GIS 2020;26:1-12. https://doi.org/10.1080/19475683.2019.1702099.

[123] Giamalaki M, Tsoutsos T. Sustainable siting of solar power installations in Mediterranean using a GIS/AHP approach. Renew Energ 2019;141:64-75. https://doi.org/10.1016/j.renene.2019.03.100.

[124] Xu Y, Li Y, Zheng L, Cui L, Li S, Li W, et al. Site selection of wind farms using GIS and multi-criteria decision-making method in Wafangdian, China. Energ 2020;207,12pp. https://doi.org/10.1016/j.energy.2020.118222.

[125] Rodman LC, Meentemeyer RK. A geographic analysis of wind turbine placement in Northern California. Energ Pol 2006;34:2137-49. https://doi.org/10.1016/j.enpol.2005.03.004.

[126] Ruiz HS, Sunarso A, Ibrahim-Bathis K, Murti SA, Budiarto I. GIS-AHP Multi Criteria Decision Analysis for the optimal location of solar energy plants at Indonesia. Energ Report 2020;6:3249-63. https://doi.org/10.1016/j.egyr.2020.11.198.

[127] Brewer J, Ames DP, Solan D, Lee R, Carlisle J. Using GIS analytics and social preference data to evaluate utility-scale solar power site suitability. Renew Energ 2015;81:825-36. https://doi.org/10.1016/j.renene.2015.04.017.

[128] Miller ME, Hui SL. Validation techniques for logistic regression models. Stat Med 1991;10:1213-26. https://doi.org/10.1002/sim.4780100805.

[129] Huang IB, Keisler J, Linkov I. Multi-criteria decision analysis in environmental sciences: Ten years of application and trends. Sci Total Environ 2011;409:3578-94. https://doi.org/10.1016/j.scitotenv.2011.06.022.

[130] Lee AHI, Chen HH, Kang H-Y. Multi-criteria decision making on strategic selection of wind farms. Renew Energ 2009;34:120-6. https://doi.org/10.1016/j.renene.2008.04.013.

[131] Rehman AU, Abidi MH, Umer U, Usmani YS. Multi-Criteria Decision-Making Approach for Selecting Wind Energy Power Plant Locations. Sustainability 2019;11:20pp. https://doi.org/10.3390/su11216112.

[132] Rouyendegh BD, Yildizbasi A, Arikan ÜZB. Using Intuitionistic Fuzzy TOPSIS in Site Selection of Wind Power Plants in Turkey. Adv Fuzz Sys 2018;2018:14pp. https://doi.org/10.1155/2018/6703798.

[133] Wu B, Yip TL, Xie L, Wang Y. A fuzzy-MADM based approach for site selection of offshore wind farm in busy waterways in China. Ocean Eng 2018;168:121-32. https://doi.org/10.1016/j.oceaneng.2018.08.065.

[134] Deveci M, Özcan E, John R, Pamučar D, Karaman H. Offshore wind farm site selection using interval rough numbers based Best-Worst Method and MARCOS. Appl Soft Comput 2021;109:28pp. https://doi.org/10.1016/j.asoc.2021.107532.

[135] Wu Y, Jinying Z, Jianping Y, Shuai G, Haobo Z. Study of decision framework of offshore wind power station site selection based on ELECTRE-III under intuitionistic fuzzy environment:

A case of China. Energ Convers Manage 2016;113:66-81. https://doi.org/10.1016/j.enconman.2016.01.020.

[136] Toklu MC, Uygun Ö. Location Selection for Wind Plant using AHP and Axiomatic Design in Fuzzy Environment. Period Eng Nat Sci 2018;6:120-8. https://doi.org/10.21533/pen.v6i2.198.

[137] Aras H, Erdoğmuş Ş, Koç E. Multi-criteria selection for a wind observation station location using analytic hierarchy process. Renew Energ 2004;29:1383-92. https://doi.org/10.1016/j.renene.2003.12.020.

[138] Gamboa G, Munda G. The problem of windfarm location: A social multi-criteria evaluation framework. Energ Pol 2007;35:1564-83. https://doi.org/10.1016/j.enpol.2006.04.021.

[139] Wu Y, Tao Y, Zhang B, Wang S, Xu C, Zhou J. A decision framework of offshore wind power station site selection using a PROMETHEE method under intuitionistic fuzzy environment: A case in China. Ocean Coast Manage 2020;184:16pp. https://doi.org/10.1016/j.ocecoaman.2019.105016.

[140] Malczewski J, Rinner C. Desktop GIS-MCDA. In: Multicriteria Decision Analysis in Geographic Information Science. Springer, Berlin, Heidelberg. pp 269-92. https://doi.org/10.1007/978-3-540-74757-4_10.

[141] Al-Shabeeb AR, Al-Adamat R, Mashagbah A. AHP with GIS for a Preliminary Site Selection of Wind Turbines in the North West of Jordan. Int J Geosci 2016;7:1208-1221. http://dx.doi.org/10.4236/ijg.2016.710090.

[142] Cai L, Zhu Y. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. Data Sci J 2015;14:1-10. http://dx.doi.org/10.5334/dsj-2015-002.

[143] Fetanat A, Khorasaninejad E. A novel hybrid MCDM approach for offshore wind farm site selection: A case study of Iran. Ocean Coast Manage 2015;109:17-28. https://doi.org/10.1016/j.ocecoaman.2015.02.005.

[144] Osborne ZM, van de Gevel SL, Eck MA, Sugg M. An Assessment of Geospatial Technology Integration in K-12 Education. J Geogr 2020;119:12-21. https://doi.org/10.1080/00221341.2019.1640271.

[145] Jha MK, Chowdary VM. Challenges of using remote sensing and GIS in developing nations. Hydrogeol J 2007;15:197-200. https://doi.org/10.1007/s10040-006-0117-1.

[146] Jun D, Tian-tian F, Yi-sheng Y, Yu M. Macro-site selection of wind/solar hybrid power station based on ELECTRE-II. Renew Sust Energ Rev 2014;35:194-204. https://doi.org/10.1016/j.rser.2014.04.005.

[147] Ricker BA, Rickles PR, Fagg GA, Haklay ME. Tool, toolmaker, and scientist: case study experiences using GIS in interdisciplinary research. Cartog Geogr Inf Sci 2020;47:350-66. https://doi.org/10.1080/15230406.2020.1748113.

[148] Jayarathna L, Rajapaska D, Managi S, Athukorala W, Torgler B, Garcia-Valiñas MA, et al. A GIS based spatial decision support system for analysing residential water demand: A case study in Australia. Sustain Cit Soc 2017;32:67-77. https://doi.org/10.1016/j.scs.2017.03.012.

[149] Smelcer JB, Carmel E. The Effectiveness of Different Representations for Managerial Problem Solving: Comparing Tables and Maps. Decision Sci 1997;28:391-420. https://doi.org/10.1111/j.1540-5915.1997.tb01316.x.

[150] Ben-Gal I. Bayesian Networks. In Ruggeri F, Kenett RS, Faltin FW, editors In: Encyclopedia of Statistics in Quality and Reliability; 2008. John Wiley & Sons, Pennsylvania, USA. 6pp. https://doi.org/10.1002/9780470061572.eqr089.

[151] Aguilera PA, Fernández A, Fernández R, Rumí R, Salmerón A. Bayesian networks in environmental modelling. Environ Model Softw 2011;26:1376-88. https://doi.org/10.1016/j.envsoft.2011.06.004.

[152] Moe SJ, Madsen AL, Connors KA, Rawlings JM, Belanger SE, Landis WG. Development of a hybrid Bayesian network model for predicting acute fish toxicity using multiple lines of evidence. Environ Model Softw 2020;126:17pp. https://doi.org/10.1016/j.envsoft.2020.104655.

[153] Sevinc V, Kucuk O, Goltas M. A Bayesian network model for prediction and analysis of possible forest fire causes. Forest Ecol Manage 2020;457:11pp. https://doi.org/10.1016/j.foreco.2019.117723.

[154] Carta JA, Velázquez S, Matías JM. Use of Bayesian networks classifiers for long-term mean wind turbine energy output estimation at a potential wind energy conversion site. Energ Convers Manage 2011;52:1137-49. https://doi.org/10.1016/j.enconman.2010.09.008.

[155] Cheng M-Y, Wu Y-F, Wu Y-W, Ndure S. Fuzzy Bayesian schedule risk network for offshore wind turbine installation. Ocean Eng 2019;188:19pp. https://doi.org/10.1016/j.oceaneng.2019.106238.

[156] Pınarbaşı K, Galparsoro I, Depellegrin D, Bald J, Pérez-Morán G, Borja Á. A modelling approach for offshore wind farm feasibility with respect to ecosystem-based marine spatial planning. Sci Total Environ 2019;667:306-17. https://doi.org/10.1016/j.scitotenv.2019.02.268.

[157] Uusitalo L. Advantages and challenges of Bayesian networks in environmental modelling. Ecol Model 2007;203:312-8. https://doi.org/10.1016/j.ecolmodel.2006.11.033.

[158] Kaikkonen L, Parviainen T, Rahikainen M, Uusitalo L, Lehikoinen A. Bayesian Networks in Environmental Risk Assessment: A Review. Integr Environ Assess Manage 2021;17:62-78. https://doi.org/10.1002/ieam.4332.

[159] Adedipe T, Shafiee M, Zio E. Bayesian Network Modelling for the Wind Energy Industry: An Overview. Reliab Eng Syst Safe 2020;202:23pp. https://doi.org/10.1016/j.ress.2020.107053.

[160] Marcot BG, Holthausen RS, Raphael MG, Rowland MM, Wisdom MJ. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management

alternative from an environmental impact statement. Forest Ecol Manage 2001;153:29-42. https://doi.org/10.1016/S0378-1127(01)00452-2.

[161] Johansson F, Falkman G. Detection of vessel anomalies – a Bayesian network approach. 2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information. 2007 Dec 3-6; Melbourne, Australia. pp. 395-400. https://doi.org/10.1109/ISSNIP.2007.4496876.

[162] Arora P, Boyne D, Slater JJ, Gupta A, Brenner DR, Druzdel MJ. Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine. Value Health 2019;22:439-45. https://doi.org/10.1016/j.jval.2019.01.006.

[163] George PC, Renjith VR. Evolution of Safety and Security Risk Assessment methodologies towards the use of Bayesian Networks in Process Industries. Process Safe Environ Protec 2021;149:758-75. https://doi.org/10.1016/j.psep.2021.03.031.

[164] Nojavan FA, Qian SS, Stow CA. Comparative analysis of discretization methods in Bayesian networks. Environ Model Softw 2017;87:64-71. https://doi.org/10.1016/j.envsoft.2016.10.007.

[165] Flores MJ, Gámez JA, Martínez AM, Puerta JM. Handling numeric attributes when comparing Bayesian network classifiers: does the discretization method matter? Appl Intell 2011;34:372-85. https://doi.org/10.1007/s10489-011-0286-z.

[166] Beuzen T, Marshall L, Splinter KD. A comparison of methods for discretizing continuous variables in Bayesian Networks. Environ Model Softw 2018;108:61-6. https://doi.org/10.1016/j.envsoft.2018.07.007.

[167] Sebe N, Cohen I, Huang TS, Gevers T. Skin detection: a Bayesian network approach. Proceedings of the 17th International Conference on Pattern Recognition, 2004. 2004 Aug 26-26; Cambridge UK. pp. 903-906. https://doi.org/10.1109/ICPR.2004.1334405.

[168] Niedermeyer D. An Introduction to Bayesian Networks and Their Contemporary Applications. In: Holmes DE, Jain LC, editors. Innovations in Bayesian Networks – Theory and Applications; 2008. Berlin: Springer; 2008. p. 117-130. https://doi.org/10.1007/978-3-540-85066-3_5.

[169] Sperotto A, Molina J-L, Torresan S, Critto A, Marcomini A. Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective. J Environ Manage 2017;202:320-31. https://doi.org/10.1016/j.jenvman.2017.07.044.

[170] Pitchforth J, Mengersen K. A proposed validation framework for expert elicited Bayesian Networks. Exp Syst Applic 2013;40:162-7. https://doi.org/10.1016/j.eswa.2012.07.026.

[171] Stritih A, Rabe S-E, Robaina O, Grêt-Regamey A, Celio E. An online platform for spatial and iterative modelling with Bayesian Networks. Environ Model Softw 2020;127:18pp. https://doi.org/10.1016/j.envsoft.2020.104658.

[172] Marcot BG, Penman TD. Advances in Bayesian network modelling: Integration of modelling techniques. Environ Model Softw 2019;111:386-93. https://doi.org/10.1016/j.envsoft.2018.09.016.

[173] Yang J, Tang W, Gong J, Shi R, Zheng M, Dai Y. Simulating urban expansion using cellular automata model with spatiotemporally explicit representation of urban demand. Landsc Urb Plann 2023;231:15pp. https://doi.org/10.1016/j.landurbplan.2022.104640.

[174] Mood C. Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It. Euro Soc Rev 2009;26:67-82. https://doi.org/10.1093/esr/jcp006.

[175] Wang N, Brown DG, An L, Yang S, Ligmann-Zielinska A. Comparative performance of logistic regression and survival analysis for detecting spatial predictors of land-use change. Int J Geogr Inf Sci 2013;27:1960-82. https://doi.org/10.1080/13658816.2013.779377.

[176] Shu B, Zhang H, Li Y, Qu Y, Chen L. Spatiotemporal variation analysis of driving forces of urban land spatial expansion using logistic regression: A case study of port towns in Taicang City, China. Habitat Int 2014;43:181-90. https://doi.org/10.1016/j.habitatint.2014.02.004.

[177] Buya S, Tongkumchum P, Owusu BE. Modelling of land-use change in Thailand using binary logistic regression and multinomial logistic regression. Arab J Geosci 2020;13:12pp. https://doi.org/10.1007/s12517-020-05451-2

[178] Pedersen E. Health aspects associated with wind turbine noise – Results from three field studies. Noise Control Eng J 2011;59:47-53. https://doi.org/10.3397/1.3533898.

[179] Stevens TK, Hale AM, Karsten KB, Bennett VJ. An analysis of displacement from wind turbines in a wintering grassland bird community. Biodivers Conserv 2013;22:1755-67. https://doi.org/10.1007/s10531-013-0510-8.

[180] Caporale D, De Lucia C. Social acceptance of on-shore wind energy in Apulia Region (Southern Italy). Renew Sust Energ Rev 2015;52:1378-90. https://doi.org/10.1016/j.rser.2015.07.183.

[181] Chen, J-L, Liu, H-H, Chuang C-T, Lu H-J. The factors affecting stakeholders' acceptance of offshore wind farms along the western coast of Taiwan: Evidence from stakeholders' perceptions. Ocean Coast Manage 2015;109:40-50. https://doi.org/10.1016/j.ocecoaman.2015.02.012.

[182] Spatial Prediction of Landslide Hazard Using Logistic Regression and ROC Analysis. Trans GIS 2006;10:395-415. https://doi.org/10.1111/j.1467-9671.2006.01004.x.

[183] Rana S, Midi H, Sarkar SK. Validation and Performance Analysis of Binary Logistic Regression Model. Proceedings of the WSEAS International Conference on Environment, Medicine and Health Sciences, 2010. 2010 Mar 23-25; Penang, Malaysia. pp. 51-55. https://www.wseas.org/multimedia/books/2010/Penang/EMEH.pdf.

[184] Raja NB, Çiçek I, Türkoğlu N, Aydin O, Kawasaki A. Landslide susceptibility mapping of the Sera River Basin using logistic regression model. Nat Hazard 2017;85:1323-46. https://doi.org/10.1007/s11069-016-2591-7.

[185] Abdel-Kader FH. Digital soil mapping at pilot sites in the northwest coast of Egypt: A multinomial logistic regression approach. Egypt J Remot Sens Space Sci 2011;14:29-40. https://doi.org/10.1016/j.ejrs.2011.04.001.

[186] Witteveen A, Nane GF, Vliegen IMH, Siesling S, Ijzerman MJ. Comparison of Logistic Regression and Bayesian Networks for Risk Prediction of Breast Cancer Recurrence. Medic Decis Making 2018;38:822-33. https://doi.org/10.1177/0272989x18790963.

[187] Goyes-Peñafiel P, Hernandez-Rojas A. Landslide susceptibility index based on the integration of logistic regression and weights of evidence: A case study in Popayan, Colombia. Eng Geol 2021;280:9pp. https://doi.org/10.1016/j.enggeo.2020.105958.

[188] Bai S-B, Wang J, Lü G-N, Zhou P-G, Hou S-S, Xu S-N. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. Geomorph 2010;115:23-31. https://doi.org/10.1016/j.geomorph.2009.09.025.

[189] Mueller JT, Brooks MM. Burdened by renewable energy? A multi-scalar analysis of distributional justice and wind energy in the United States. Energ Res Soc Sci 2020;63:13pp. https://doi.org/10.1016/j.erss.2019.101406.

[190] Zucca A, Sharifi AM, Fabbri AG. Application of spatial multi-criteria analysis to site selection for a local park: A case study in the Bergamo Province, Italy. J Environ Manage 2008;88:752-69. https://doi.org/10.1016/j.jenvman.2007.04.026.

[191] Gkeka-Serpetsidaki P, Tsoutsos T. A methodological framework for optimal siting of offshore wind farms: A case study on the island of Crete. Energ 2022;239:19pp. https://doi.org/10.1016/j.energy.2021.122296.

[192] Merckx B, Steyaert M, Vanreusel A, Vincx M, Vanaverbeke J. Null models reveal preferential sampling, spatial autocorrelation and overfitting in habitat suitability modelling. Ecol Model 2011;222:588-97. https://doi.org/10.1016/j.ecolmodel.2010.11.016.

[193] Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49:1225-31. https://doi.org/10.1016/S0895-4356(96)00002-9.

[194] Subramanian J, Simon R. Overfitting in prediction models – Is it a problem only in high dimensions. Contemp Clin Trial 2013;36:636-41. https://doi.org/10.1016/j.cct.2013.06.011.

[195] Khan HR, Shaw JEH. Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data. J Data Sci 2011;9:93-110. https://ssrn.com/abstract=2019344.

[196] Kong SG, Jin D, Li S, Kim H. Fast fire flame detection in surveillance video using logistic regression and temporal smoothing. Fire Safe J 2016;79:37-43. https://doi.org/10.1016/j.firesaf.2015.11.015.

[197] Eroğlu H. Multi-criteria decision analysis for wind power plant location selection based on fuzzy AHP and geographic information systems. Environ Dev Sust 2021;23:18278-310. https://doi.org/10.1007/s10668-021-01438-5.

[198] Shaheen M, Khan MZ. A method of data mining for selection of site for wind turbines. Renew Sust Energ Rev 2016;55:1225-33. https://doi.org/10.1016/j.rser.2015.04.015.

[199] Kim C-K, Jang S, Kim TY. Site selection for offshore wind farms in the southwest coast of South Korea. Renew Energ 2018;120:151-62. https://doi.org/10.1016/j.renene.2017.12.081.

[200] Solangi YA, Tan Q, Khan MWA, Mirjat NH, Ahmed, I. The Selection of Wind Power Project Location in the Southeastern Corridor of Pakistan: A Factor Analysis, AHP, and Fuzzy-TOPSIS Application. Energies 2018;11:26pp. https://doi.org/10.3390/en11081940.

[201] Harper M, Anderson B, James PAB, Bahaj AS. Assessing socially acceptable locations for onshore wind energy using a GIS-MCDA approach. Int J Low Carb Technol 2019;14:160-9. https://doi.org/10.1093/ijlct/ctz006.

[202] Höfer T, Sunak Y, Siddique H, Madlener R. Wind farm siting using a spatial Analytic Hierarchy Process approach: A case study of the Städteregion Aachen. Appl Energ 2016;163:222-43. https://doi.org/10.1016/j.apenergy.2015.10.138.

[203] Höltinger S, Salak B, Schauppenlehner T, Scherhaufer P, Schmidt J. Austria's wind energy potential – A participatory modeling approach to assess socio-political and market acceptance. Energ Pol 2016;98:49-61. https://doi.org/10.1016/j.enpol.2016.08.010.

[204] Lotfi R, Mostafaeipour A, Mardani N, Mardani S. Investigation of wind farm location planning by considering budget constraints. Int J Sust Energ 2018;37:799-817. https://doi.org/10.1080/14786451.2018.1437160.

[205] Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 2016;9:1937-58. https://doi.org/10.5194/gmd-9-1937-2016.

[206] Petrova MA. NIMBYism revisited: public acceptance of wind energy in the United States. Wires Clim Change 2013;4:575-601. https://doi.org/10.1002/wcc.250.

[207] Sheikh NJ, Kocaoglu DF, Lutzenhiser L. Social and political impacts of renewable energy: Literature review. Technol Forecast Soc 2016;108:102-10. https://doi.org/10.1016/j.techfore.2016.04.022.

[208] Dai K, Bergot A, Liang C, Xiang W-N, Huang Z. Environment issues associated with wind energy – A review. Renew Energ 2015;75:911-21. https://doi.org/10.1016/j.renene.2014.10.074.

[209] Xiao Y, Watson M. Guidance on Conducting a Systematic Literature Review. J Plan Educ Res 2017;39:93-112. https://doi.org/10.1177/0739456X17723971.

[210] Siddaway AP, Wood AM, Hedges LV. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. Annu Rev Psychol 2019;70:747-70. https://doi.org/10.1146/annurev-psych-010418-102803.

[211] Badampudi D, Wohlin C, Petersen K. Experiences from using snowballing and database searches in systematic literature studies. Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering; 2015 April 27-29; Nanjing, China. New York: Association for Computing Machinery; 2015. p. 1-10. https://doi.org/10.1145/2745802.2745818.

[212] Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering; 2014 May 13-14; London, UK. New York: Association for Computing Machinery; 2014. p. 1-10. https://doi.org/10.1145/2601248.2601268.

[213] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Ann Intern Med 2009;151:264-9. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.

[214] Esri. “Light Gray Canvas” [basemap] 1:143,432,406 scale “Light Gray Canvas Map”, https://www.arcgis.com/home/item.html?id=979c6cc89af9449cbeb5342a439c6a76; 2022 [accessed 2 June 2023].

[215] Cunden TSM, Doorga J, Lollchund MR, Rughooputh SDDV. Multi-level constraints wind farms siting for a complex terrain in a tropical system using MCDM approach coupled with GIS. Energ 2020;211:17pp. https://doi.org/10.1016/j.energy.2020.118533.

[216] Ayodele TR, Ogunjuyigbe ASO, Odigie O, Munda JL. A multi-criteria GIS based model for wind farm site selection using interval type-2 fuzzy analytic hierarchy process: The case study of Nigeria. Appl Energ 2018;228:1853-69. https://doi.org/10.1016/j.apenergy.2018.07.051.

[217] Ali S, Lee S-M, Jang C-M. Determination of the Most Optimal On-Shore Wind Farm Site Location Using a GIS-MCDM Methodology: Evaluating the Case of South Korea. Energies 2017;10:22pp. https://doi.org/10.3390/en10122072.

[218] Tegou L-I, Polatidis H, Haralambopoulos DA. Environmental management framework for wind farm siting: Methodology and case study. J Environ Manage 2010;91:2134-47. https://doi.org/10.1016/j.jenvman.2010.05.010.

[219] Al-Yahyai S, Charabi Y, Gastli A, Al-Badi A. Wind farm land suitability indexing using multi-criteria analysis. Renew Energ 2012;44:80-87. https://doi.org/10.1016/j.renene.2012.01.004.

[220] Jangid J, Bera AK, Joseph M, Singh V, Singh TP, Pradhan BK, et al. Potential zones identification for harvesting wind energy resources in desert region of India – A multi criteria

evaluation approach using remote sensing and GIS. Renew Sust Energ Rev 2016;65:1-10. https://doi.org/10.1016/j.rser.2016.06.078.

[221] Ali Y, Butt M, Sabir M, Mumtaz U, Salman A. Selection of suitable site in Pakistan for wind power plant installation using analytic hierarchy process. J Control Decis 2017;5:117-128. https://doi.org/10.1080/23307706.2017.1346490.

[222] Ali S, Taweekun J, Techato K, Waewsak J, Gyawali S. GIS based site suitability assessment for wind and solar farms in Songkhla, Thailand. Renew Energ 2018;132:1360-72. https://doi.org/10.1016/j.renene.2018.09.035.

[223] Ayodele TR, Ogunjuyigbe ASO, Odigie O, Jimoh AA. On the most suitable sites for wind farm development in Nigeria. Data Br 2018;19:29-41. https://doi.org/10.1016/j.dib.2018.04.144.

[224] Bili A, Vagiona DG. Use of multicriteria analysis and GIS for selecting sites for onshore wind farms: the case of Andros Island (Greece). Eur J Environ Sci 2018;8:5-13. https://doi.org/10.14712/23361964.2018.2.

[225] Değirmenci S, Bingöl F, Sofuoglu SC. MCDM analysis of wind energy in Turkey: decision making based on environmental impact. Environ Sci Pollut Res 2018;25:19753-66. https://doi.org/10.1007/s11356-018-2004-4.

[226] Díaz-Cuevas P. GIS-Based Methodology for Evaluating the Wind-Energy Potential of Territories: A Case Study from Andalusia (Spain). Energies 2018;11:16pp. https://doi.org/10.3390/en11102789.

[227] Hannsen F, May R, van Dijk J. Spatial Multi-Criteria Decision Analysis Tool Suite for Consensus-Based Siting of Renewable Energy Structures. J Environ Assess Pol Manage 2018;20:28pp. https://www.jstor.org/stable/90025515.

[228] Li Z. Study of site suitability assessment of regional wind resources development based on multi-criteria decision. Clean Technol Environ Pol 2018;20:1147-66. https://doi.org/10.1007/s10098-018-1538-y.

[229] Koc A, Turk S, Şahin G. Multi-criteria of wind-solar site selection problem using a GIS-AHP-based approach with an application in Igdir Province/Turkey. Environ Sci Pollut Res 2019;26:32298-310. https://doi.org/10.1007/s11356-019-06260-1.

[230] Konstantinos I, Georgios T, Garyfalos A. A Decision Support System methodology for selecting wind farm installation locations using AHP and TOPSIS: Case study in Eastern Macedonia and Thrace region, Greece. Energ Pol 2019;132:242-46. https://doi.org/10.1016/j.enpol.2019.05.020.

[231] Mukhamediev RI, Mustakayev R, Yakunin K, Kiseleva S, Gopejenko V. Multi-Criteria Spatial Decision Making Supportsystem for Renewable Energy Development in Kazakhstan. IEEE Access 2019;7:122275-88. https://doi.org/10.1109/ACCESS.2019.2937627.

[232] Ari ES, Gencer C. The use and comparison of a deterministic, a stochastic, and a hybrid multiple-criteria decision-making method for site selection of power plants: An application in Turkey. Wind Eng 2020;44:60-74. https://doi.org/10.1177/0309524X19849831.

[233] Adedeji PA, Akinlabi SA, Madushele N, Olatunji OO. Hybrid neurofuzzy investigation of short-term variability of wind resource in site suitability analysis: A case study in South Africa. Neural Comput Appl 2021;33:13049-74. https://doi.org/10.1007/s00521-021-06001-x.

[234] Amjad F, Agyekum EB, Shah LA, Abbas A. Site location and allocation decision for onshore wind farms, using spatial multi-criteria analysis and density-based clustering. A techno-economic environmental assessment, Ghana. Sustain Energ Technol Assess 2021;47:18pp. https://doi.org/10.1016/j.seta.2021.101503.

[235] Barzekhar M, Parnell KE, Dinan NM, Brodie G. Decision support tools for wind and solar farm site selection in Isfahan Province, Iran. Clean Technol Environ Pol 2021;23:1179-95. https://doi.org/10.1007/s10098-020-01978-w.

[236] Díaz-Cuevas P. Haddad B, Fernandez-Nunez M. Energy for the future: Planning and mapping renewable energy. The case of Algeria. Sustain Energ Technol Assess 2021;47:13pp. https://doi.org/10.1016/j.seta.2021.101445.

[237] Gharaibeh AA, Al-Shboul DA, Al-Rawabdeh AM, Jaradat RA. Establishing Regional Power Sustainability and Feasibility Using Wind Farm Land-Use Optimization. Land 2021;10:32pp. https://doi.org/10.3390/land10050442.

[238] Spyridonidou S, Sismani G, Loukogeorgaki E, Vagiona DG, Ulanovsky H, Madar D. Sustainable Spatial Energy Planning of Large-Scale Wind and PV Farms in Israel: A Collaborative and Participatory Planning Approach. Energies 2021;14:23pp. https://doi.org/10.3390/en14030551.

[239] Xing L, Wang Y. A practical wind farm siting framework integrating ecosystem services – A case study of coastal China. Environ Impact Assess Rev 2021;90:14pp. https://doi.org/10.1016/j.eiar.2021.106636.

[240] Zahid F, Tahir A, Khan HU, Naeem MA. Wind farms selection using geospatial technologies and energy generation capacity in Gwadar. Energ Rep 2021;7:5857-70. https://doi.org/10.1016/j.egyr.2021.08.165.

[241] Zalhaf AS, Elboshy B, Kotb KM, Han Y, Almaliki AH, Aly RMH, et al. A High-Resolution Wind Farms Suitability Mapping Using GIS and Fuzzy AHP Approach: A National-Level Case Study in Sudan. Sustainability 2021;14:21pp. https://doi.org/10.3390/su14010358.

[242] Azizi A, Malekmohammadi B, Jafari HR, Nasiri H, Parsa VA. Land suitability assessment for wind power plant site selection using ANP-DEMATEL in a GIS environment: case study of Ardabil province, Iran. Environ Monit Assess 2014;186:6695-709. https://doi.org/10.1007/s10661-014-3883-6.

[243] Shorabeh SN, Argany M, Rabiei J, Firozjaei K, Nematollahi O. Potential assessment of multi-renewable energy farms establishment using spatial multi-criteria decision analysis: A case study and mapping in Iran. J Clean Prod 2021;295:15pp. https://doi.org/10.1016/j.jclepro.2021.126318.

[244] Aydin NY, Kentel E, Duzgun S. GIS-based environmental assessment of wind energy systems for spatial planning: A case study from Western Turkey. Renew Sust Energ Rev 2010;14:364-73. https://doi.org/10.1016/j.rser.2009.07.023.

[245] Aydin NY, Kentel E, Duzgun S. GIS-based site selection methodology for hybrid renewable energy systems: A case study from Western Turkey. Energ Convers Manage 2013;70:90-106. https://doi.org/10.1016/j.enconman.2013.02.004.

[246] Tercan E. Land suitability assessment for wind farms through best-worst method and GIS in Balikesir province of Turkey. Sustain Energ Technol Assess 2021;47:14pp. https://doi.org/10.1016/j.seta.2021.101491.

[247] Noorollahi Y, Yousefi H, Mohammadi M. Multi-criteria decision support system for wind farm site selection using GIS. Sustain Energ Tech Assess 2016;13:38-50. https://doi.org/10.1016/j.seta.2015.11.007.

[248] Kazak J, van Hoof J, Szewranski S. Challenges in the wind turbines location process in Central Europe – The use of spatial decision support systems. Renew Sust Energ Rev 2017;76:425-33. https://doi.org/10.1016/j.rser.2017.03.039.

[249] Zheng C-W, Li C-Y, Xu, J-J. Micro-scale classification of offshore wind energy resource -- A case study of the New Zealand. J Clean Prod 2019;226:133-41. https://doi.org/10.1016/j.jclepro.2019.04.082.

[250] Sabil FDAQ, Damayanti A, Taqyuddin A. GIS Application for Determining Potential Locations for the Development of Wind Power Plants. Int J Geomate 2020;19:149-56. https://geomatejournal.com/geomate/article/view/1094.

[251] Sliz-Szkliniarz B, Vogt J. GIS-based approach for the evaluation of wind energy potential: A case study for the Kujawsko-Pomorskie Voivodeship. Renew Sust Energ Rev 2011;15:1696-707. https://doi.org/10.1016/j.rser.2010.11.045.

[252] Ouammi A, Ghigliotti V, Robba M, Mimet A, Sacile R. A decision support system for the optimal exploitation of wind energy on regional scale. Renew Energ 2012;37:299-309. https://doi.org/10.1016/j.renene.2011.06.027.

[253] Eichhorn M, Tafarte P, Thrän D. Towards energy landscapes - "Pathfinder for sustainable wind power locations." Energ 2017;134:611-21. https://doi.org/10.1016/j.energy.2017.05.053.

[254] Bina SM, Jalilinasrabady S, Fujii H, Farabi-Asl H. A comprehensive approach for wind power plant potential assessment, application to northwestern Iran. Energ 2018;164:344-58. https://doi.org/10.1016/j.energy.2018.08.211.

[255] Rehman S, Baseer MA, Alhems LM. GIS-Based Multi-Criteria Wind Farm Site Selection Methodology. FME Trans 2020;48:855-67. https://doi.org/10.5937/fme2004855R.

[256] Rehman S, Baseer MA, Alhems LM. A Heuristic Approach to Siting and Design Optimization of an Onshore Wind Farm Layout. Energies 2020;13:18pp. https://doi.org/10.3390/en13225946.

[257] Pojadas DJ, Abundo MLS. Spatio-temporal assessment and economic analysis of a grid-connected island province toward a 35% or greater domestic renewable energy portfolio: a case in Bohol, Philippines. Int J Energ Environ Eng 2021;12:251-80. https://doi.org/10.1007/s40095-020-00369-7.

[258] Gorsevski PV, Cathcart SC, Mirzaei G, Jamali MM, Ye X, Gomezdelcampo E. A group-based spatial decision support system for wind farm site selection in Northwest Ohio. Energ Pol 2013;55:374-85. https://doi.org/10.1016/j.enpol.2012.12.013.

[259] Sánchez-Lozano JM, García-Cascales MS, Lamata MT. Identification and selection of potential sites for onshore wind farms development in Region of Murcia, Spain. Energ 2014;73:311-24. https://doi.org/10.1016/j.energy.2014.06.024.

[260] Rezaei M, Mostafaeipour A, Qolipour M, Tavakkoli-Moghaddam R. Investigation of the optimal location design of a hybrid wind-solar plant: A case study. Int J Hydrogen Energ 2018;43:100-114. https://doi.org/10.1016/j.ijhydene.2017.10.147.

[261] Chaouachi A, Covrig CF, Ardelean M. Multi-criteria selection of offshore wind farms: Case study for the Baltic States. Energ Pol 2017;103:179-92. https://doi.org/10.1016/j.enpol.2017.01.018.

[262] Mahdy M, Bahaj AS. Multi criteria decision analysis for offshore wind energy potential in Egypt. Renew Energ 2018;118:278-89. https://doi.org/10.1016/j.renene.2017.11.021.

[263] Spyridonidou S, Vagiona DG. Spatial energy planning of offshore wind farms in Greece using GIS and a hybrid MCDM methodological approach. Eur Mediterr J Environ Integr 2020;5:13pp. https://doi.org/10.1007/s41207-020-00161-3.

[264] Tercan E, Tapkin S, Latinopoulos D, Dereli MA, Tsiropoulos A, Ak MF. A GIS-based multi-criteria model for offshore wind energy power plants site selection in both sides of the Aegean Sea. Environ Monit Assess 2020;192:20pp. https://doi.org/10.1007/s10661-020-08603-9.

[265] Vinhoza A, Schaeffer R. Brazil's offshore wind energy potential assessment based on a Spatial Multi-Criteria Decision Analysis. Renew Sust Energ Rev 2021;146:14pp. https://doi.org/10.1016/j.rser.2021.111185.

[266] Huang J, Huang X, Song N, Ma Y, Wei D. Evaluation of the Spatial Suitability of Offshore Wind Farm – A Case Study of the Sea Area of Liaoning Province. Sustainability 2022;14:17pp. https://doi.org/10.3390/su14010449.

[267] Cradden L, Kalogeri C, Barrios IM, Galanis G, Ingram D, Kallos G. Multi-criteria site selection for offshore renewable energy platforms. Renew Energ 2016;87:791-806. https://doi.org/10.1016/j.renene.2015.10.035.

[268] Genç MS, Karipoğlu F, Koca K, Azgın ŞT. Suitable site selection for offshore wind farms in Turkey's seas: GIS-MCDM based approach. Earth Sci Inform 2021;14:1213-25. https://doi.org/10.1007/s12145-021-00632-3.

[269] Wu Y, Geng S, Xu H, Zhang H. Study of decision framework of wind farm project plan selection under intuitionistic fuzzy set and fuzzy measure environment. Energ Convers Manage 2014;87:274-84. https://doi.org/10.1016/j.enconman.2014.07.001.

[270] Ouammi A, Dagdougui H, Sacile R. Optimal Planning With Technology Selection for Wind Power Plants in Power Distribution Networks. IEEE Sys J 2019;13:3059-69. https://doi.org/10.1109/JSYST.2019.2903555.

[271] Ouammi A, Sacile R, Zejli D, Mimet A, Benchrifa R. Sustainability of a wind power plant: Application to different Moroccan sites. Energ 2010;35:4226-36. https://doi.org/10.1016/j.energy.2010.07.010.

[272] Kim J-Y, Oh, K-Y, Kang K-S, Lee J-S. Site selection of offshore wind farms around the Korean Peninsula through economic evaluation. Renew Energ 2013;54:189-95. https://doi.org/10.1016/j.renene.2012.08.026.

[273] Azadeh A, Ghaderi SF, Nasrollahi MR. Location optimization of wind plants in Iran by an integrated hierarchical Data Envelopment Analysis. Renew Energ 2011;36:1621-31. https://doi.org/10.1016/j.renene.2010.11.004.

[274] Azadeh A, Rahimi-Golkhandan A, Moghaddam M. Location optimization of wind power generation-transmission systems under uncertainty using hierarchical fuzzy DEA: A case study. Renew Sust Energ Rev 2014;30:877-85. https://doi.org/10.1016/j.rser.2013.10.020.

[275] Qolipour M, Mostafaeipour A, Shamshirband S, Alavi S, Goudarzi H, Petković D. Evaluation of wind power generation potential using a three hybrid approach for households in Ardebil Province, Iran. Energ Convers Manage 2016;118:295-305. https://doi.org/10.1016/j.enconman.2016.04.007.

[276] Mostafaeipour A, Sadeghi S, Jahangiri M, Nematollahi O, Sabbagh AR. Investigation of accurate location planning for wind farm establishment: a case study. J Eng Design Technol 2019;18:821-45. https://doi.org/10.1108/JEDT-08-2019-0208.

[277] Rezaei-Shouroki M, Mostafaeipour A, Qolipour M. Prioritizing of wind farm locations for hydrogen production: A case study. Int J Hydrogen Energ 2017;42:9500-10. https://doi.org/10.1016/j.ijhydene.2017.02.072.

[278] Pambudi G, Nananukul N. A hierarchical fuzzy data envelopment analysis for wind turbine site selection in Indonesia. Energ Rep 2019;5:1041-7. https://doi.org/10.1016/j.egyr.2019.08.002.

[279] Jahangiri M, Ghaderi R, Haghani A, Nematollahi O. Finding the best locations for establishment of solar-wind power stations in Middle-East using GIS: A review. Renew Sust Energ Rev 2016;66:38-52. https://doi.org/10.1016/j.rser.2016.07.069.

[280] Passoni G, Rowcliffe JM, Whiteman A, Huber D, Kusak J. Framework for strategic wind farm site prioritisation based on modelled wolf reproduction habitat in Croatia. Eur J Wildl Res 2017;63:16pp. https://doi.org/10.1007/s10344-017-1092-7.

[281] Guo X, Zhang X, Du S, Li C, Siu YL, Rong Y, et al. The impact of onshore wind power projects on ecological corridors and landscape connectivity in Shanxi, China. J Clean Prod 2020;254:14pp. https://doi.org/10.1016/j.jclepro.2020.120075.

[282] Wu Y-n, Yang Y-s, Feng T-t, Kong L-n, Liu W, Fu Luo-j. Macro-site selection of wind/solar hybrid power station based on Ideal Matter-Element Model. Electr Power Energ Sys 2013;50:76-84. https://doi.org/10.1016/j.ijepes.2013.02.024.

[283] Tekin S, Guner ED, Cilek A, Cilek MU. Selection of renewable energy systems sites using the MaxEnt model in the Eastern Mediterranean region in Turkey. Environ Sci Pollut Res 2021;28:51405-24. https://doi.org/10.1007/s11356-021-13760-6.

[284] Ari ES, Gencer C. Proposal of a novel mixed integer linear programming model for site selection of a wind power plant based on power maximization with use of mixed type wind turbines. Energ Environ 2020;31:825-41. https://doi.org/10.1177/0958305X19882394.

[285] Zhang X-y, Wang X-k, Yu S-m, Wang J-q, Wang T-l. Location selection of offshore wind power station by consensus decision framework using picture fuzzy modelling. J Clean Prod 2018;202:980-92. https://doi.org/10.1016/j.jclepro.2018.08.172.

[286] Kamdar I, Ali S, Taweekun J, Ali HM. Wind Farm Site Selection Using WAsP Tool for Application in the Tropical Region. Sustainability 2021;13:25pp. https://doi.org/10.3390/su132413718.

[287] Rose S, Jaramillo P, Small MJ, Grossmann I, Apt J. Quantifying the hurricane risk to offshore wind turbines. P Natl Acad Sci USA 2012;109:3247-52. https://doi.org/10.1073/pnas.1111769109.

[288] Charlton TS, Rouainia M. Geotechnical fragility analysis of monopile foundations for offshore wind turbines in extreme storms. Renew Energ 2022;182:1126-40. https://doi.org/10.1016/j.renene.2021.10.092.

[289] Haggett C, ten Brink T, Russell A, Roach M, Firestone J, Dalton T, et al. Offshore Wind Projects and Fisheries: Conflict and Engagement in the United Kingdom and the United States. Oceanography 2020;33:38-47. https://doi.org/10.5670/oceanog.2020.404.

[290] Schupp MF, Kafas A, Buck BH, Krause G, Onyango V, Stelzenmüller V, et al. Fishing within offshore wind farms in the North Sea: Stakeholder perspectives for multi-use from Scotland and Germany. J Environ Manage 2021;279:10pp. https://doi.org/10.1016/j.jenvman.2020.111762.

[291] BirdLife International. Data Zone, http://datazone.birdlife.org/home; 2022 [accessed 30 May 2023].

[292] Kjørlaug RA, Kaynia AM. Vertical earthquake response of megawatt-sized wind turbine with soil-structure interaction effects. Earthquake Eng Struct Dyn 2015;44:2341-58. https://doi.org/10.1002/eqe2590.

[293] De Risi R, Bhattacharya S, Goda K. Seismic performance assessment of monopile-supported offshore wind turbines using unscaled natural earthquake records. Soil Dyn Earthquake Eng 2018;109:154-72. https://doi.org/10.1016/j.soildyn.2018.03.015.

[294] Luo Y. Wind Turbine Site Selection over Abandoned Mined Land. National Association of Abandoned Mine Land Programs 32nd Annual Conference; 2015 September 19-22; Scranton, PA, USA. 2015. 16pp. https://www.researchgate.net/publication/266740801_WIND_TURBINE_SITE_SELECTION_OVER_ABANDONED_MINED_LAND.

[295] Lerner M. Local power: Understanding the adoption and design of county wind energy regulation. Rev Pol Res 2021;39:120-42. https://doi.org/10.1111/ropr.12447.

[296] Scacchi W. Free/Open Source Software Development: Recent Research Results and Methods. Adv Comput 2007;69:243-95. https://doi.org/10.1016/S0065-2458(06)69005-0.

[297] Deveci M, Erdogan N, Cali U, Stekli J, Zhong S. Type-2 neutrosophic number based multi-attributive border approximation area comparison (MABAC) approach for offshore wind farm site selection in USA. Eng Appl Artif Intel 2021;103:14pp. https://doi.org/10.1016/j.engappai.2021.104311.

[298] Cooper H, Tingle P. COVID-19 Stimulus Bill Includes Key Renewable Energy Tax Credits, https://www.mwe.com/insights/covid-19-stimulus-bill-includes-key-renewable-energy-tax-credits; 2020 [accessed 31 May 2023].

[299] Lesser J. The Biden Administration's Offshore Wind Fantasy, https://www.manhattan-institute.org/lesser-biden-administrations-offshore-wind-fantasy; 2022 [accessed 31 May 2023].

[300] Zhang J, Fowai I, Sun K. A Glance at Offshore Wind Turbine Foundation Structures. Shipbuilding 2016;67:101-13. http://dx.doi.org/10.21278/brod67204.

[301] Refsgaard JC, Henriksen HJ. Modelling guidelines – terminology and guiding principles. Adv Water Resour 2004;27:71-82. https://doi.org/10.1016/j.advwatres.2003.08.006.

[302] Vanherle K, Werkman AM, Baete E, Barkmeijer A, Kolm A, Gast C et al. Proposed Standard model and consistent terminology for monitoring and outcome evaluation in different dietetic care settings: Results from the EU-sponsored IMPECD project. Clin Nutr 2018;37:2206-16. https://doi.org/10.1016/j.clnu/2018/08.040.

[303] Augusiak J, Van den Brink PJ, Grimm V. Merging validation and evaluation of ecological models to 'evaludation': A review of terminology and a practical approach. Ecol Model 2014;280:117-28. https://doi.org/10.1016/j.ecolmodel.2013.11.009.

[304] Raebel MA, Schmittdiel J, Karter AJ, Konieczny JL, Steiner JK. Standardizing Terminology and Definitiosn of Medication Adherence and Persistence in Research employing Electronic Databases. Med Care 2013;51:11-21. https://doi.org/10.1097%2FMLR.0b013e31829b1d2a.

[305] Rodríguez-Rodríguez D, Sebastiao J, Tierra AES, Martínez-Vega J. Effect of protected areas in reducing land development across geographic and climate conditions of a rapidly developing country, Spain. Land Degrad Dev 2019;30:991-1005. https://doi.org/10.1002/ldr.3286.

[306] Global Wind Atlas, https://globalwindatlas.info/en; 2023 [accessed 2 June 2023].

[307] Deveci M, Panucar DM, Cali U, Kantar E, Kölle K, Tande JO. Hybrid q-Rung Orthopair Fuzzy Sets Based CoCoSo Model for Floating Offshore Wind Farm Site Selection in Norway. CSEE J Pow Energ Syst 2022;8:1261-80. https://doi.org/10.17775/CSEEJPES.2021.07700.

[308] Kalliamvakou E, Gousios G, Blincoe K, Singer L, German DM, Damian D. An in-depth study of the promises and perils of mining GitHub. Empir Softw Eng 2016;21:2035-71. https://doi.org/10.1007/s10664-015-9393-5.

[309] Martinez-Moyano IJ. Documentation for model transparency. Syst Dynam Rev 2012;28:199-208. https://doi.org/10.1002/sdr.1471.

[310] Schumacher K, Yang Z. The determinants of wind energy growth in the United States: Drivers and barriers to state-level development. Renew Sust Energ Rev 2018;81:1-13. https://doi.org/10.1016/j.rser.2018.08.017.

[311] TIGER/Line Shapefiles. Current County and Equivalent National Shapefile. United States Census Bureau, Data Catalog. https://catalog.data.gov/dataset/tiger-line-shapefile-2019-nation-u-s-current-county-and-equivalent-national-shapefile; 2019 [accessed 1 June 2023].

[312] TIGER/Line Shapefiles. Current State and Equivalent National. United States Census Bureau, Data Catalog. https://catalog.data.gov/dataset/tiger-line-shapefile-2019-nation-u-s-current-state-and-equivalent-national; 2019 [accessed 1 June 2023].

[313] Ritter Z. USFWS Threatened & Endangered Species Active Critical Habitat Report. United States Fish & Wildlife Service, Environmental Conservation Online System. https://ecos.fws.gov/ecp/report/table/critical-habitat.html; 2023 [accessed 1 June 2023].

[314] Stutts M. National Register of Historic Places. National Park Services DataStore, Geospatial Dataset – (Code: 2210280). https://irma.nps.gov/DataStore/Reference/Profile/2210280; 2014 [accessed 1 June 2023].

[315] TIGER/Line Shapefiles. Military Installation National Shapefile. United States Census Bureau, Data Catalog. https://catalog.data.gov/dataset/tiger-line-shapefile-2019-nation-u-s-military-installation-national-shapefile#sec-dates; 2019 [accessed 1 June 2023].

[316] Horton JD, San Juan CA. Prospect- and mine-related features from U.S. Geological Survey 7.5- and 15-minute topographic quadrangle maps of the United States. United States Geological Survey, v8.0. https://doi.org/10.5066/F78W3CHG; 2022 [accessed 1 June 2023].

[317] National Park Services. Administrative Boundaries of National Park System Units – National Geospatial Data Asset (NGDA) NPS National Parks Dataset. National Park Services DataStore, Geospatial Dataset – (Code: 2225713). 2022. https://doi.org/10.57830/2225713.

[318] TIGER/Line Shapefiles. Current American Indian/Alaska Native/Native Hawaiian Areas National (AIANNH) National. United Census Bureau, Data Catalog. https://catalog.data.gov/dataset/tiger-line-shapefile-2018-nation-u-s-current-american-indian-alaska-native-native-hawaiian-area; 2018 [accessed 1 June 2023].

[319] Campbell J. FWS National Realty Tracts Simplified. United States Fish & Wildlife Service. https://gis-fws.opendata.arcgis.com/datasets/fws::fws-national-realty-tracts-simplified/about; 2022 [accessed 1 June 2023].

[320] University of Wyoming Department of Ecosystem Science. United States Annual Temperature Raster. United States Geological Survey – ScienceBase Catalog. https://www.sciencebase.gov/catalog/item/57a26dd6e4b006cb45553f7a; 2016 [accessed 1 June 2023].

[321] Draxl C, Clifton A, Hodge B-M, McCaa J. United States Wind Speed at 80-Meter above Surface Level. National Renewable Energy Laboratory – Wind Resource Maps and Data. https://www.nrel.gov/gis/wind-resource-maps.html; 2017 [accessed 1 June 2023].

[322] United States Geological Survey. GAP Analysis Project (GAP) – Download Species Range and Predicted Habitat Data. United States Geological Survey – ScienceBase Catalog. https://gapanalysis.usgs.gov/apps/species-data-download/; 2018 [accessed 1 June 2023].

[323] United States Department of Transportation. Aviation Facilities. Bureau of Transportation Statistics. https://data-usdot.opendata.arcgis.com/datasets/usdot::aviation-facilities/explore?location=11.590267%2C-1.633860%2C2.86; 2022 [accessed 1 June 2023].

[324] United States Geological Survey. USA Hospitals. Esri. https://www.arcgis.com/home/item.html?id=f114757725a24d8d9ce203f61eaf8f75; 2022 [accessed 1 June 2023].

[325] United States Geological Survey. USGS Small-scale Dataset – 1:1,000,000-Scale Major Roads of the United States 201403 Shapefile. United States Geological Survey – ScienceBase Catalog. https://www.sciencebase.gov/catalog/item/581d052be4b08da350d524ce; 2016 [accessed 1 June 2023].

[326] HIFLD. Electric Power Transmission Lines. Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-geoplatform.opendata.arcgis.com/datasets/electric-power-transmission-lines/explore; 2022 [accessed 1 June 2023].

[327] HIFLD. Public Schools. Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-geoplatform.opendata.arcgis.com/datasets/public-schools/explore; 2022 [accessed 1 June 2023].

[328] National Center for Education Statistics. Private School Locations – Current. EDGE Open Data. https://data-nces.opendata.arcgis.com/datasets/nces::private-school-locations-current/explore; 2021[78] [accessed 1 June 2023].

[329] KTH Royal Institute of Technology in Stockholm. Global Power Plant Database. World Resources Institute. https://datasets.wri.org/dataset/globalpowerplantdatabase; 2018 [accessed 1 June 2023].

[330] HIFLD. Independent System Operators. Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-geoplatform.opendata.arcgis.com/maps/geoplatform::independent-system-operators-1; 2022 [accessed 1 June 2023].

[331] United States Department of Agriculture. Agricultural Land Values – Annual Reports by the National Agricultural Statistics Service. https://usda.library.cornell.edu/concern/publications/pn89d6567?locale=es; 2023 [accessed 1 June 2023].

[332] Larson W, Shui J, Davis M, Oliner S. Working Paper 19-01: The Price of Residential Land for Counties, ZIP codes, and Census Tracts in the United States. Federal Housing Finance Agency. https://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/wp1901.aspx; 2020 [accessed 1 June 2023].

[333] NC Clean Energy Technology Center. Database of State Incentives for Renewables & Efficiency (DSIRE). https://www.dsireusa.org/; 2023 [accessed 1 June 2023].

[334] National Governors Association. Former Governors – NGA Archives. https://www.nga.org/former-governors/; 2023 [accessed 1 June 2023].

[335] MIT Election Data and Science Lab. County Presidential Election Returns 2000-2020. Harvard Dataverse, v11. 2023. https://doi.org/10.7910/DVN/VOQCHQ.

[336] OpenSecrets. Follow The Money – Lobbyist Link. https://www.followthemoney.org/tools/lobbyist-link; 2023 [accessed 1 June 2023]. Licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License by OpenSecrets.

[337] BEA. Employment by County, Metro, and Other Areas – Interactive Data. Bureau of Economic Analysis. https://www.bea.gov/data/employment/employment-county-metro-and-other-areas; 2023 [accessed 1 June 2023].

[338] BLS. Local Area Unemployment Statistics – Annual Average Labor Force Data by County. US Bureau of Labor Statistics. https://www.bls.gov/lau/tables.htm; 2023 [accessed 1 June 2023].

[339] Marlon J, Neyens, L, Jefferson M, Howe P, Mildenberger M, Leiserowitz A. Yale Climate Opinion Maps 2018. Yale Program on Climate Change Communication. https://climatecommunication.yale.edu/visualizations-data/ycom-us/; 2022 [accessed 1 June 2023].

[340] Ogle TM, Salazar K. Indiana Renewable Energy Community Planning Survey and Ordinance Inventory Summary. Purdue University Indiana (US): 2022. 37 pp. https://cdext.purdue.edu/wpcontent/uploads/2022/03/Renewable-Energy-Report.pdf.

[341] Kikuchi R. Adverse impacts of wind power generation on collision behaviour of birds and anti-predator behaviour of squirrels. J Nat Conserv 2008;16:44-55. https://doi.org/10.1016/j.jnc.2007.11.001.

[342] Saganeiti L, Pilogallo A, Faruolo G, Scorza F, Murgante B. Territorial Fragmentation and Renewable Energy Source Plants: Which Relationship? Sustainability 2020;12:14pp. https://doi.org/10.3390/su12051828.

[343] National Park Service. Management Policies 2006. United States Department of the Interior (US): 2006. 180 pp. https://www.nps.gov/subjects/policy/upload/MP_2006.pdf.

[344] Lewis DA. Identifying and Avoiding Conflicts Between Historic Preservation and the Development of Renewable Energy. NYU Environ Manag J 2015;22:274-360. https://www.nyuelj.org/wp-content/uploads/2015/02/Lewis_READY_FOR_WEBSITE.pdf.

[345] Auld T. McHenry MP, Whale J. US military, airspace, and meteorological radar system impacts from utility class wind turbines: Implications for renewable energy targets and the wind industry. Renew Energ 2013;55:24-30. https://doi.org/10.1016/j.renene.2012.12.008.

[346] Zimmerman MG, Reames TG. Where the wind blows: Exploring barriers and opportunities to renewable energy development on United States tribal lands. Energ Res Soc Sci 2021;72:12pp. https://doi.org/10.1016/j.erss.2020.101874.

[347] Wind Energy Technologies Office. U.S. Average Annual Wind Speed at 80 Meters. https://windexchange.energy.gov/maps-data/319; 2023 [accessed 1 June 2023].

[348] Davis MS, Madani MR. Geolocating Aeola for Use in Wind Turbine Siting. 2018 6th International Renewable and Sustainable Energy Conference (IRSEC); 2018 December 5-8; Rabat, Morocco. IEEE;2018. 6pp. https://doi.org/10.1109/IRSEC.2018.8702838.

[349] Grau L, Jung C, Schindler D. On the Annual Cycle of Meteorological and Geographical Potential of Wind Energy: A Case Study from Southwest Germany. Sustainability 2017;7:11pp. https://doi.org/10.3390/su9071169.

[350] Kraj AG, Bibeau EL. Phases of icing on wind turbine blades characterized by ice accumulation. Renew Energ 2010;35:966-72. https://doi.org/10.1016/j.renene.2009.09.013.

[351] Aschwanden J, Stark H, Peter D, Steuri T, Schmid B, Liechti F. Bird collisions at wind turbines in a mountainous area related to bird movement intensities measured by radar. Biol Conserv 2018;220:228-36. https://doi.org/10.1016/j.biocon.2018.01.005.

[352] Cryan PM, Barclay RMR. Causes of Bat Fatalities at Wind Turbines: Hypotheses and Predictions. J Mammal 2009;90:1330-40. https://doi.org/10.1644/09-MAMM-S-076R1.1.

[353] Grodsky SM, Behr MJ, Gendler A, Drake D, Dieterle BD, Rudd RJ, et al. Investigating the causes of death for wind turbine-associated bat fatalities. J Mammal 2011;92:917-25. https://doi.org/10.1644/10-MAMM-A-404.1.

[354] Enevoldsen P. Onshore wind energy in Northern European forests: Reviewing the risks. Renew Sust Energ Rev 2016;60:1251-62. https://doi.org/10.1016/j.rser.2016.02.027.

[355] Heintzman LJ, Auerbach ES, Kilborn DH, Starr SM, Mulligan KR, Barbato LS, et al. Identifying areas of wetland and wind turbine overlap in the south-central Great Plains of North America. Landscape Ecol 2020;35:1995-2011. https://doi.org/10.1007/s10980-020-01076-8.

[356] Mills A, Wiser R, Porter K. The cost of transmission for wind energy in the United States: A review of transmission planning studies. Renew Sust Energ Rev 2012;16:1-19. https://doi.org/10.1016/j.rser.2011.07.131.

[357] Bertsiou MM, Theochari AP, Baltas E. Multi-criteria analysis and Geographic Information Systems methods for wind turbine siting in a North Aegean island. Energ Sci Eng 2020;9:4-18. https://doi.org/10.1002/ese3.809.

[358] Peri E, Tal A. Is setback distance the best criteria for siting wind turbines under crowded conditions? An empirical analysis. Energ Pol 2021;155:10pp. https://doi.org/10.1016/j.enpol.2021.112346.

[359] Namowitz D. Oklahoma Bill Limits Wind Turbines' Encroachment. https://www.aopa.org/news-and-media/all-news/2015/april/28/oklahoma-bill-limits-wind-turbine-encroachment; 2015 [accessed 1 June 2023].

[360] Cuadra L, Ocampo-Estrella I, Alexandre E, Salcedo-Sanz S. A study on the impact of easements in the deployment of wind farms near airport facilities. Renew Energ 2019;135:566-88. https://doi.org/10.1016/j.renene.2018.12.038.

[361] Rode DC, Fischbeck PS, Páez AR. The retirement cliff: Power plant lives and their policy implications. Energ Pol 2017;106:222-32. https://doi.org/10.1016/j.enpol.2017.03.058.

[362] McGhee MJ, Galloway G, Catterson VM, Brown B, Harrison E. Prognostic Modeling of Valve Degradation within Power Stations. Annual Conference of the PHM Society 2014;6:6pp. https://doi.org/10.36001/phmconf.2014.v6i1.2356.

[363] Stevanovic VD, Petrovic MM, Wala T, Milivojevic S, Ilic M, Muszynski S. Efficiency and power upgrade at the aged lignite-fired power plant by flue gas waste heat utilization: High

pressure versus low pressure economizer installation. Energy 2019;187:12pp. https://doi.org/10.1016/j.energy.2019.115980.

[364] Sáenz de Miera G, del Río González P, Vizcaíno I. Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. Energ Pol 2008;36:3345-59. https://doi.org/10.1016/j.enpol.2008.04.022.

[365] Traber T, Kemfert C. Gone with the wind? – Electricity market prices and incentives to invest in thermal power plants under increasing wind energy supply. Energ Econo 2011;33:249-56. https://doi.org/10.1016/j.eneco.2010.07.002.

[366] Martinez-Anido CB, Brinkman G, Hodge B-M. The impact of wind power on electricity prices. Renew Energ 2016;94:474-87. https://doi.org/10.1016/j.renene.2016.03.053.

[367] Albuyeh F, Alaywan Z. Implementation of the California independent system operator. Proceedings of the 21st International Conference on Power Industry Computer Applications; 1999 May 21; Santa Clara, USA. IEEE;1999. 6pp. https://doi.org/10.1109/PICA.1999.779408.

[368] Kemabonta T, Kabalan M. Integration of renewable energy resources from the perspective of the Midcontinent Independent System Operator: A review. The Elec J 2018;31:22-33. https://doi.org/10.1016/j.tej.2018.10.013.

[369] Klass AB. Expanding the U.S. Electric Transmission and Distribution Grid to Meet Deep Carbonization Goals. Environ Law Rep 2017;47:18pp. https://dx.doi.org/10.2139/ssrn.3033829.

[370] Bird L, Lew D, Milligan M, Carlini EM, Estanquiero A, Flynn D, et al. Wind and solar energy curtailment: A review of international experience. Renew Sust Energ Rev 2016;65:577-86. https://doi.org/10.1016/j.rser.2016.06.082.

[371] Collins AR, Hansen E, Hendryx M. Wind versus coal: Comparing the local economic impacts of energy resource development in Appalachia. Energ Pol 2012;50:551-61. https://doi.org/10.1016/j.enpol.2012.08.001.

[372] Stafford BA, Wilson EJ. Winds of change in energy systems: Policy implementation, technology deployment, and regional transmission organizations. Energ Res Soc Sci 2016;21:222-36. https://doi.org/10.1016/j.erss.2016.08.001.

[373] Vyn RJ, McCullough RM. The Effects of Wind Turbines on Property Values in Ontario: Does Public Perception Match Empirical Evidence? Can J Agr Econ 2014;62:365-92. https://doi.org/10.1111/cjag.12030.

[374] Sampson GS, Perry ED, Tayler MR. The On-Farm and Near-Farm Effects of Wind Turbines on Agricultural Land Values. J Agr Resour Econ 2020;45:410-27. https://doi.org/10.22004/ag.econ.302463.

[375] Castleberry B, Greene JS. Wind power and real estate prices in Oklahoma. Int J Hous Mark Analysis 2018;11:808-27. https://doi.org/10.1108/IJHMA-02-2018-0010.

[376] Lantz E, Doris E. State Clean Energy Practices: Renewable Energy Rebates. Report No: NREL/TP-6A2-45039. [Internet] Golden, CO (US): National Renewable Energy Lab; 2008 [cited 2023 Jan 13]. Available from: https://doi.org/10.2172/950149.

[377] Abolhosseini S, Heshmati A. The main support mechanisms to finance renewable energy development. Renew Sust Energ Rev 2014;40:876-85. https://doi.org/10.1016/j.rser.2014.08.013.

[378] Ozcan M. Assessment of renewable energy incentive system from investors' perspective. Renew Energ 2014;71:425-32. https://doi.org/10.1016/j.renene.2014.05.053.

[379] Horne C, Kennedy EH. Explaining support for renewable energy: commitments to self-sufficiency and communion. Environ Pol 2019;28:929-49. https://doi.org/10.1080/09644016.2018.1517917.

[380] Huang M-Y, Alavalapati JRR, Carter DR, Langholtz MH. Is the choice of renewable portfolio standards random? Energ Pol 2007:35;5571-5. https://doi.org/10.1016/j.enpol.2007.06.010.

[381] Gustafson A, Goldberg MH, Kotcher JE, Rosenthal SA, Maibach EW, Ballew MT, et al. Republicans and Democrats differ in why they support renewable energy. Energ Pol 2020;141:12pp. https://doi.org/10.1016/j.enpol.2020.111448.

[382] Inês C, Guilherme PL, Esther M-G, Swantje G, Stephen H, Lars H. Regulatory challenges and opportunities for collective renewable energy prosumers in the EU. Energ Pol 2020;138:11pp. https://doi.org/10.1016/j.enpol.2019.111212.

[383] Yin H, Powers N. Do state renewable portfolio standards promote in-state renewable generation? Energ Pol 2010;38:1140-9. https://doi.org/10.1016/j.enpol.2009.10.067.

[384] Chernyakhovskiy I, Tian T, McLaren J, Miller M, Geller N. U.S. Laws and Regulations for Renewable Energy Grid Interconnections. Report No: NREL/TP-6A20-66724. [Internet] Golden, CO (US): National Renewable Energy Lab; 2016 [cited 2023 Jan 13]. Available from: https://doi.org/10.2172/1326721.

[385] Zhou S, Solomon BD. Do renewable portfolio standards in the United States stunt renewable electricity development beyond mandatory targets? Energ Pol 2020;140;11pp. https://doi.org/10.1016/j.enpol.2020.111377.

[386] Dumas M, Rising J, Urpelainen J. Political competition and renewable energy transitions over long time horizons: A dynamic approach. Ecol Econ 2016;124:175-184. https://doi.org/10.1016/j.ecolecon.2016.01.019.

[387] Yu F, Yu X. Corporate Lobbying and Fraud Detection. J Financ Quant Anal 2011;46:1865-91. https://doi.org/10.1017/S0022109011000457.

[388] Lu Y, Khan ZA, Alvarez-Alvarado MS, Zhang Y, Huang Z, Imran M. A Critical Review of Sustainable Energy Policies for the Promotion of Renewable Energy Sources. Sustainability 2020;12:30pp. https://doi.org/10.3390/su12125078.

[389] Slattery MC, Lantz E, Johnson BL. State and local economic impacts from wind energy projects: Texas case study. Energ Pol 2011;39:7930-40. https://doi.org/10.1016/j.enpol.2011.09.047.

[390] Lambert RJ, Silva PP. The challenges of determining the employment effects of renewable energy. Renew Sust Energ Rev 2012;16:4667-74. https://doi.org/10.1016/j.rser.2012.03.072.

[391] Louie EP, Pearce JM. Retraining investment for U.S. transition from coal to solar photovoltaic employment. Energ Econ 2016;57:295-302. https://doi.org/10.1016/j.eneco.2016.05.016.

[392] Dorrell J, Lee K. The Cost of Wind: Negative Economic Effects of Global Wind Energy Development. Energies 2020;13:24pp. https://doi.org/10.3390/en13143667.

[393] Huesca-Pérez ME, Sheinbaum-P C, Köppel J. Social implications of siting wind energy in a disadvantaged region – The case of the Isthmus of Tehuantepec, Mexico. Renew Sust Energ Rev 2016;58:952-65. https://doi.org/10.1016/j.rser.2015.12.310.

[394] Sokoloski R, Markowitz EM, Bidwell D. Public estimates of support for offshore wind energy: False consensus, pluralistic ignorance, and partisan effects. Energ Pol 2018;112:45-55. https://doi.org/10.1016/j.enpol.2017.10.005.

[395] Brannstrom C, Santos Leite N, Lavoie A, Gorayeb A. What explains the community acceptance of wind energy? Exploring benefits, consultation, and livelihoods in coastal Brazil. Energ Res Soc Sci 2022;83:13pp. https://doi.org/10.1016/j.erss.2021.102344.

[396] Jacquet JB, Stedman RC. Perceived Impacts from Wind Farm and Natural Gas Development in Northern Pennsylvania. Rural Sociol 2013;78:450-72. https://doi.org/10.1111/ruso.12022.

[397] Firestone J, Bates A, Knapp LA. See me, Feel me, Touch me, Heal me: Wind turbines, culture, landscapes, and sound impressions. Land Use Pol 2015;46:241-9. https://doi.org/10.1016/j.landusepol.2015.02.015.

[398] Crowe J. Explaining Popular Support for Wind Energy in the United States. J Rural Soc Sci 2020;35:32pp. https://egrove.olemiss.edu/jrss/vol35/iss2/2.

[399] Zhang ZX, Zhang HY, Zhou DW. Using GIS spatial analysis and logistic regression to predict the probabilities of human-caused grassland fires. J Arid Environ 2010;74:386-93. https://doi.org/10.1016/j.jaridenv.2009.09.024.

[400] Conoscenti C, Angileri S, Cappadonia C, Rotigliano E, Agnesi V, Märker M. Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). Geomorphology 2014;204:399-411. https://doi.org/10.1016/j.geomorph.2013.08.021.

[401] Kropat G, Bochud F, Murith C, Palacios M, Baechler S. Modeling of geologic radon in Switzerland based on ordered logistic regression. J Environ Radioactiv 2017;166:376-81. https://doi.org/10.1016/j.jenvrad.2016.06.007.

[402] Wang L, Ai T, Shen Y, Li J. The isotropic organization of DEM structure and extraction of valley lines using hexagonal grid. T GIS 2020;24;483-507. https://doi.org/10.1111/tgis.12611.

[403] Birch CPD, Oom SP, Beecham JA. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecol Model 2007;206:347-59. https://doi.org/10.1016/j.ecolmodel.2007.03.041.

[404] Frey HC, Patil SR. Identification and Review of Sensitivity Analysis Methods. Risk Anal 2002;22:553-78. https://doi.org/10.1111/0272-4332.00039.

[405] National Renewable Energy Laboratory. Land Use by System Technology. https://www.nrel.gov/analysis/tech-size.html; 2023 [accessed 1 June 2023].

[406] Center for Sustainable Systems. Wind Energy Factsheet. Report No: CSS07-09. [Internet] Ann Arbor, MI (US): University of Michigan; 2021 [cited 2023 Jan 15]. Available from: https://css.umich.edu/sites/default/files/2022-09/Wind_CSS07-09.pdf.

[407] Esri. "World Terrain Reference" [basemap] 1:6,335,223 scale "World Terrain Reference Map", https://www.arcgis.com/home/item.html?id=14fbc125ccc9488888b014db09f35f67; 2023 [accessed 1 June 2023].

[408] Feng J, Xu H, Mannor S, Yan S. Robust Logistic Regression and Classification. Advances in Neural Information Processing Systems 27 (NIPS 2014); 2014 December 8-13; Montréal, Canada. NIPS;2018. 9pp. https://proceedings.neurips.cc/paper/2014/file/6cdd60ea0045eb7a6ec44c54d29ed402-Paper.pdf.

[409] Martin N, Pardo L. On the asymptotic distribution of Cook's distance in logistic regression models. J Appl Stat 2009;36:1119-46. https://doi.org/10.1080/02664760802562498.

[410] Midi H, Sarkar SK, Rana S. Collinearity diagnostics of binary logistic regression model. J Interdiscipl Math 2010;13:253-67. https://doi.org/10.1080/09720502.2010.10700699.

[411] Craney TA, Surles JG. Model-Dependent Variance Inflation Factor Cutoff Values. Qual Eng 2002;14:391-403. https://doi.org/10.1081/QEN-120001878.

[412] Shrestha N. Application of Binary Logistic Regression Model to Assess the Likelihood of Overweight. Am J Theo Appl Stat 2019;8:18-25. https://doi.org/10.11648/j.ajtas.20190801.13.

[413] Šinkovec H, Geroldinger A, Heinze G. Bring More Data! – A Good Advice? Removing Separation in Logistic Regression by Increasing Sample Size. Int J Environ Res Public Health 2019;16:12pp. https://doi.org/10.3390%2Fijerph16234658.

[414] le Cessie S, van Houwelingen JC. Logistic Regression for Correlated Binary Data. J Roy Stat Soc C-App 1994;43:95-108. https://doi.org/10.2307/2986114.

[415] Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A Comparison of Goodness-of-Fit Tests for the Logistic Regression Model. Stat Med 1997;16:965-80.

https://doi.org/10.1002/(SICI)1097-0258(19970515)16:9%3C965::AID-SIM509%3E3.0.CO;2-O.

[416] Cano JR, Herrera F, Lozano M. On the combination of evolutionary algorithms and stratified strategies for training set selection in datamining. Appl Soft Comput 2006;6:323-32. https://doi.org/10.1016/j.asoc.2005.02.006.

[417] Wilks SS. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. Ann Math Stat 1938;9:60-2. https://doi.org/10.1214/aoms/1177732360.

[418] Fan J, Zhang C, Zhang J. Generalized Likelihood Ratio Statistics and Wilks Phenomenon. Ann Stat 2001;29:153-93. https://doi.org/10.1214/aos/996986505.

[419] Walker DA, Smith TJ. Nine Pseudo R2 Indices for Binary Logistic Regression Models. J Mod Appl Stat Methods 2016;15:848-54. https://doi.org/10.22237/jmasm/1462077720.

[420] Rutherford GN, Guisan A, Zimmermann NE. Evaluating sampling strategies and logistic regression methods for modelling complex land cover changes. J Appl Ecol 2007;44:414-24. https://doi.org/10.1111/j.1365-2664.2007.01281.x.

[421] Delice A. The Sampling Issues in Quantitative Research. Educ Sci-Theor Pract 2010;10:2001-18. https://eric.ed.gov/?id=EJ919871.

[422] Szumilas M. Explaining Odds Ratios. J Can Acad Child Adolesc Psychiatry 2010;19:227-9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/.

[423] Sperandei S. Understanding logistic regression analysis. Biochem Medica 2014;24:12-8. https://doi.org/10.11613/BM.2014.003.

[424] Hyandye C, Mandara CG, Safari J. GIS and Logit Regression Model Applications in Land Use/Land Cover Change and Distribution in Usangu Catchment. Am J Remote Sens 2015;3:6-16. https://doi.org/10.11648/j.ajrs.20150301.12.

[425] Pontius Jr RG, Schneider LC. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agr Ecosyst Environ 2001;85:239-48. https://doi.org/10.1016/S0167-8809(01)00187-6.

[426] Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery 2016;159:1638-45. https://doi.org/10.1016/j.surg.2015.12.029.

[427] Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Patt Recog 1997;30:1145-59. https://doi.org/10.1016/S0031-3203(96)00142-2.

[428] Indra ST, Wikarsa L, Turang R. Using logistic regression method to classify tweets into the selected topics. 2016 International Conference on Advanced Computer and Science and Information Systems; 2016 October 15-16; Malang, Indonesia. IEEE;2016. p. 385-90. https://doi.org/10.1109/ICACSIS.2016.7872727.

[429] Parker JR. Rank and response combination from confusion matrix data. Inform Fusion 2001;2:113-20. https://doi.org/10.1016/S1566-2535(01)00030-6.

[430] Wimhurst JJ, Greene JS. Oklahoma's future wind energy resources and their relationship with the Central Plains low-level jet. Renew Sust Energ Rev 2019;115:24pp. https://doi.org/10.1016/j.rser.2019.109374.

[431] Getis A, Ord JK. The Analysis of Spatial Association by Use of Distance Statistics. Geogr Anal 1992;24:189-206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x.

[432] Mennis J, Harris P. Contagion and repeat offending among urban juvenile delinquents. J Adolescence 2011;34:951-63. https://doi.org/10.1016/j.adolescence.2010.12.001.

[433] Mayfield HJ, Lowry JH, Watson CH, Kama M, Nilles EJ, Lau CL. Use of geographically weighted logistic regression to quantify spatial variation in the environmental and sociodemographic drivers of leptospirosis in Fiji: a modelling study. Lancet Planet Health 2018;2:223-32. https://doi.org/10.1016/S2542-5196(18)30066-4.

[434] Liu X, Li X, Liu L, He J, Ai B. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. Int J Geogr Inf Sci 2008;22:1247-69. https://doi.org/10.1080/13658810701757510.

[435] Li X, Yeh A G-O. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int J Geogr Inf Sci 2000;14:131-52. https://doi.org/10.1080/136588100240886.

[436] Maithani S. Cellular Automata Based Model of Urban Spatial Growth. J Indian Soc Remot Sens 2010;38:604-10. https://doi.org/10.1007/s12524-010-0053-3.

[437] Kocabas V, Dragicevic S. Assessing cellular automata model behaviour using a sensitivity analysis approach. Comput Environ Urban 2006;30:921-53. https://doi.org/10.1016/j.compenvurbsys.2006.01.001.

[438] Zaitsev DA. A generalized neighborhood for cellular automata. Theor Comput Sci 2017;666:21-35. https://doi.org/10.1016/j.tcs.2016.11.002.

[439] Shang X-C, Li X-G, Xie D-F, Jia B, Jiang R. Two-lane traffic flow model based on regular hexagonal cells with realistic lane changing behavior. Physica A 2020;560:13pp. https://doi.org/10.1016/j.physa.2020.125220.

[440] Douass S, Kbir MA. Flood zones detection using a runoff model built on Hexagonal shape based cellular automata. Int J Eng Trend Technol 2020;68:68-74. https://doi.org/10.14445/22315381/IJETT-V68I6P211S.

[441] Nugraha AT, Waterson BJ, Blainey SP, Nash FJ. On the consistency of urban cellular automata models based on hexagonal and square cells. Environ Urban Plann B 2020;48:845-60. https://doi.org/10.1177/2399808319898501.

[442] Ménard A, Marceau DJ. Exploration of spatial scale sensitivity in geographic cellular automata. Environ Plann B 2005;32:693-714. https://doi.org/10.1068/b31163.

[443] Pan Y, Roth A, Yu Z, Doluschitz R. The impact of variation in scale on the behavior of a cellular automata used for land use change modeling. Comput Environ Urban Syst 2010;34:400-8. https://doi.org/10.1016/j.compenvurbsys.2010.03.003.

[444] Wu H, Li Z, Clarke KC, Shi W, Fang L, Lin A, Zhou J. Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change. Int J Geogr Inf Sci 2019;33:1040-61. https://doi.org/10.1080/13658816.2019.1568441.

[445] Feizizadeh B, Darabi S, Blaschke T, Lakes T. QADI as a New Method and Alternative to Kappa for Accuracy Assessment of Remote Sensing-Based Image Classification. Sensors 2022;22:21pp. https://doi.org/10.3390/s22124506.

[446] Pontius Jr RG, Millones M. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int J Remot Sens 2011;32:4407-29. https://doi.org/10.1080/01431161.2011.552923.

[447] Liu X, Wang J, Gu F, Xiong J, Wei J. Median Based Adaptive Quantization of Log-Likelihood Ratios. 2018 IEEE 87th Vehicular Technology Conference; 2018 June 3-6; Porto, Portugal. IEEE;2018. 5pp. https://doi.org/10.1109/VTCSpring.2018.8417501.

[448] Mahmood Z, Khan S, Salzman P. Improved resampling technique for the choice of stopping criterion and model selection in stepwise logistic regression. Pak J Statist 2016;32:21-36. http://www.pakjs.com/wp-content/uploads/2019/10/3212.pdf.

[449] Hausfather Z, Drake HF, Abbott T, Schmidt GA. Evaluating the Performance of Past Climate Model Projections. Geophys Res Lett 2020;47:10pp. https://doi.org/10.1029/2019GL085378.

[450] Collins M, Chandler RE, Cox PM, Huthnance JM, Rougier J, Stephenson DB. Quantifying future climate change. Nature Clim Change 2012;2:403-9. https://doi.org/10.1038/nclimate1414.

[451] Räisänen J. How reliable are climate models? Tellus A Dyn Meteorol Oceanog 2007;59:2-29. https://doi.org/10.1111/j.1600-0870.2006.00211.x.

[452] Smith TJ, McKenna CM. A Comparison of Logistic Regression Pseudo R2 Indices. Mult Line Regres View 2013;39:17-26. https://www.glmj.org/archives/articles/Smith_v39n2.pdf.

[453] Hemmert GAJ, Schons LM, Wieseke J, Schimmelpfennig H. Log-likelihood-based Pseudo-R2 IN Logistic Regression: Deriving Sample-sensitive benchmarks. Sociol Method Res 2018;47:507-31. https://doi.org/10.1177/0049124116638107.

[454] Lo K, Wong M, Khalechelvam P, Tam W. Waist-to-height ratio, body mass index and waist circumference for screening paediatric cardio-metabolic risk factors: a meta-analysis. Obes Rev 2016;17:1258-75. https://doi.org/10.1111/obr.12456.

[455] Hernández-Orallo J, Flach P, Ferri C. ROC curves in cost space. Mach Learn 2013;93:71-91. https://doi.org/10.1007/s10994-013-5328-9.

[456] Mossman D. Evaluating Risk Assessments Using Receiver Operating Characteristic Analysis: Rationale, Advantages, Insights, and Limitations. Behav Sci Law 2013;31:23-39. https://doi.org/10.1002/bsl.2050.

[457] Linden A. Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. J Eval Clinic Practice 2006;12:132-9. https://doi.org/10.1111/j.1365-2753.2005.00598.x.

[458] Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom 2020;21:13pp. https://doi.org/10.1186/s12864-019-6413-7.

[459] Yun H. Prediction model of algal blooms using logistic regression and confusion matrix. Int J Electr Comput Eng 2021;11:2407-13. http://doi.org/10.11591/ijece.v11i3.pp2407-2413.

[460] Widen JC, Tholen M, Yim JJ, Bogyo. Methods for analysis of near-infrared (NIR) quenched-fluorescent contrast agents in mouse models of cancer. Method Enzymol 2020;639:141-66. https://doi.org/10.1016/bs.mie.2020.04.012.

[461] Doubrawa P, Barthelmie RJ, Pryor SC, Basager CB, Badger M, Karagali I. Satellite winds as a tools for offshore wind resource assessment: The Great Lakes Wind Atlas. Remot Sens Environ 2015;168:349-59. https://doi.org/10.1016/j.rse.2015.07.008.

[462] Ritter A, Muñoz-Carpena R. Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. J Hydrol 2013;480:33-45. https://doi.org/10.1016/j.jhydrol.2012.12.004.

[463] Lemos MC, Rood RB. Climate projections and their impact on policy and practice. WIREs Clim Change 2010;1:670-82. https://doi.org/10.1002/wcc.71.

[464] Bray D, von Storch H. “Prediction” or “Projection”? The Nomenclature of Climate Science. Sci Commun 2009;30:534-43. https://doi.org/10.1177/1075547009333698.

[465] Palmer TN, Shutts GJ, Hagedorn R, Doblas-Reyes FJ, Jung T, Leutbecher M. Representing model uncertainty in weather and climate prediction. Ann Rev Earth Planet Sci 2005;33:163-93. https://doi.org/10.1146/annurev.earth.33.092203.122552.

[466] United States Census Bureau. A State’s Median Age Does Not Tell The Whole Story: New Census Bureau Visualization Shows Broad Variations in Age Structure By State and County, https://www.census.gov/library/stories/2022/07/states-median-age-does-not-tell-whole-story.html; 2022 [accessed 14 June 2023].

[467] Songchitruska P, Zeng X. Getis-Ord Spatial Statistics to Identify Hot Spots by Using Incident Management Data. J Trans Res Record 2010;2165:42-51. https://doi.org/10.3141/2165-05.

[468] Li X, Zhou Y, Chen W. An improved urban cellular automata model by using the trend-adjusted neighborhood. Ecol Process 2020;9:13pp. https://doi.org/10.1186/s13717-020-00234-9.

[469] Long Y, Shen Z, Mao Q. Retrieving spatial policy parameters from an alternative plan using constrained cellular automata and regionalized sensitivity analysis. Environ Plann B: Plann Design 2012;39:586-605. https://doi.org/10.1068/b37122.

[470] Batty M, Xie Y, Sun S. Modeling urban dynamics through GIS-based cellular automata. Comput Environ Urban Sys 1999;23:205-33. https://doi.org/10.1016/S0198-9715(99)00015-0.

[471] Iwanaga T, Wang H-H, Hamilton SH, Grimm V, Koralewski TE, Salado A, et al. Socio-technical scales in socio-environmental modeling: Managing a system-of-systems modeling approach. Environ Modell Softw 2021;135. https://doi.org/10.1016/j.envsoft.2020.104885.

[472] Adua L, Zhang KX, Clark B. Seeking a handle on climate change: Examining the comparative effectiveness of energy efficiency improvement and renewable energy production in the United States. Global Environ Chang 2021;70:10pp. https://doi.org/10.1016/j.gloenvcha.2021.102351.

[473] Hache E. Do renewable energies improve energy security in the long run? Int Econ 2018;156:127-35. https://doi.org/10.1016/j.inteco.2018.01.005.

[474] Kamusoko C, Aniya M, Adi B, Manjoro M. Rural sustainability under threat in Zimbabwe – Simulation of future land use/cover changes in the Bindura district based on the Markov-cellular automata model. Appl Geogr 2009;29:435-47. https://doi.org/10.1016/j.apgeog.2008.10.002.

[475] Avolio MV, Di Gregorio S, Mantovani F, Pasuto A, Rongo R, Silvano S, et al. Simulation of the 1992 Tessina landslide by a cellular automata model and future hazard scenarios. Int J Appl Earth Obs 2000;2:41-50. https://doi.org/10.1016/S0303-2434(00)85025-4.

[476] Wang J-F, Jiang C-S, Hu M-G, Cao Z-D, Guo Y-S, Li L-F, et al. Design-based spatial sampling: Theory and implementation. Environ Modell Softw 2013;40:280-8. https://doi.org/10.1016/j.envsoft.2012.09.015.

[477] Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. Behav Ecol 2006;17:688-90. https://doi.org/10.1093/beheco/ark016.

[478] Gomes E, Abrantes P, Banos A, Rocha J. Modelling future land use scenarios based on farmers' intentions and a cellular automata approach. Land Use Pol 2019;85:142-54. https://doi.org/10.1016/j.landusepol.2019.03.027.

[479] Mirbagheri B, Alimohammadi A. Improving urban cellular automata performance by integrating global and geographically weighted logistic regression models. Trans GIS 2017;21:1280-97. https://doi.org/10.1111/tgis.12278.

[480] US Department of Energy. National Offshore Wind Strategy: Facilitating the Development of the Offshore Wind Industry in the United States,

https://www.energy.gov/sites/default/files/2016/09/f33/National-Offshore-Wind-Strategy-report-09082016.pdf; 2016 [accessed 27 June 2023].

[481] Bureau of Ocean Energy Management. Renewable Energy – State Activities. https://www.boem.gov/renewable-energy/state-activities; 2023 [accessed 28 June 2023].

[482] NOAA National Centers for Environmental Information. Bathymetric Data Viewer. https://www.ncei.noaa.gov/maps/bathymetry/; 2023 [accessed 28 June 2023].

[483] NOAA Fisheries and National Ocean Service – InPort. Shipping Fairways, Lanes, and Zones for US waters. https://www.fisheries.noaa.gov/inport/item/39986; 2015 [accessed 28 June 2023].

[484] NOAA Fisheries. Essential Fish Habitat – Data Inventory. https://www.habitat.noaa.gov/application/efhinventory/index.html; 2021 [accessed 28 June 2023].

[485] Kauffman GJ. What if… the United States of America were based on watersheds? Water Policy 2002;4:57-68. https://doi.org/10.1016/S1366-7017(02)00019-3.

[486] Sengupta M, Xie Y, Lopez A, Habte A, Maclaurin G, Shelby J. Global Horizontal Solar Irradiance – Americas. National Renewable Energy Laboratory – Solar Resource Maps and Data. https://www.nrel.gov/gis/solar-resource-maps.html; 2018 [accessed 28 June 2023].

[487] United States Energy Information Administration. Power Plants. https://atlas.eia.gov/datasets/power-plants/explore; 2023 [accessed 28 June 2023].

Appendix

A1: Link to the GitHub repository that contains WiFSS-LRCA's model code, model user instructions, and aggregated datasets for the CONUS and its 48 individual U.S. states: https://github.com/JoshuaW1994/Wind-Farm-Site-Suitability-CONUS.

A2: Link to the spreadsheet containing Supplementary Material for the systematic review (See Section 2.2): https://sooners-my.sharepoint.com/:x:/g/personal/joshua_j_wimhurst-1_ou_edu/EUSZO1GKaG5HjidzOaiXjq4BC6U0SPyixNfpnVsFDlQY4A?e=v46kmA.

A3: Example console output from running WiFSS-LRCA's Logistic Regression equation:

Console output

---------- DATASET SELECTION AND SETUP ----------

NOTE: The desired study region must be specified as 'CONUS' if one wishes to execute the logistic regression model over states that contain zero commercial wind farms (Louisiana, Mississippi Alabama, Georgia, South Carolina, Kentucky), states that possess wind farms in only one grid cell at all but the highest spatial resolutions (Arkansas, Florida, Virginia, Delaware, Connecticut, New Jersey, Tennessee), or states at low spatial resolutions at which too many predictors were removed due to collinearity (Rhode Island at the 100th or 80th percentile).

Specified study region: Indiana
Specified wind farm density: 65 acres/MW
Specified wind power capacity: 100th percentile (525 MW)
Predictor configurations specified by the user: ['Full']

Predictors removed from the model based on having a constant value in all grid cells: ['ISO_YN']

---------- TESTING ASSUMPTIONS ----------

Assumption #1: All continuous predictors have a linear relationship with the logit of the dependent variable, based on a Box-Tidwell test.

Bonferroni-corrected p-value: 0.001388888888888889

Results of the Box-Tidwell test:

Predictor	pval
Avg_Temp	0.000674
Avg_25	0.001022
Undev_Land	0.003864
Near_Sch	0.004940
Avg_Wind	0.006605
Unem_15_19	0.006907
Hisp_15_19	0.014735
Whit_15_19	0.038947
Near_Plant	0.145049
Prop_Rugg	0.150025
Type_15_19	0.236192
Near_Air	0.242192
Dens_15_19	0.276145
Near_Trans	0.412721
Near_Roads	0.545709
Near_Hosp	0.647832
Fem_15_19	0.733248
Avg_Elevat	0.783045

The Box-Tidwell test would have dropped ['Avg_Temp', 'Avg_25'] from the model, though the user chose to retain them.

Assumption #2: There is no multicollinearity, or pairwise collinearity, between the model's predictors, based on Variance Inflaction Factors (VIF).

Grouped Multicollinearity Test Results:

Predictor	VIF
Military	1.104673
Trib_Land	1.105936
Plant_Year	1.124445
Farm_Year	1.131381
Wild_Refug	1.152373
Near_Roads	1.188766
Near_Plant	1.278938
Near_Air	1.304042
Near_Trans	1.315547
Nat_Parks	1.325883
Fem_15_19	1.352213
Bat_Count	1.421241
Critical	1.433754
Near_Hosp	1.557812
Near_Sch	1.570322
Dens_15_19	1.570552
Avg_25	1.671992
Unem_15_19	2.025566
Bird_Count	2.153854
Hisp_15_19	2.418413
Dem_Wins	2.668159
Historical	2.718621
Mining	2.813543
Avg_Elevat	3.845557
Avg_Wind	3.961974
Prop_Rugg	3.981475
Undev_Land	4.366777
Type_15_19	4.909543
supp_2018	4.916688
Avg_Temp	5.002557
Whit_15_19	9.471814

Pairwise Multicollinearity Test Results:

Predictor1	Predictor2	VIF	
Near_Hosp	Plant_Year	1.000000	
Dens_15_19	Unem_15_19	1.000000	
Avg_Wind	Unem_15_19	1.000001	
Near_Roads	Unem_15_19	1.000002	
Type_15_19	Unem_15_19	1.000004	
...	...	...	...
Type_15_19	supp_2018	1.851897	
Avg_Temp	Avg_Wind	1.995120	
Prop_Rugg	Undev_Land	2.116583	
Whit_15_19	supp_2018	3.179561	
Type_15_19	Whit_15_19	3.819502	

Predictors to be removed based on multicollinearity: None

Assumption #3: None of the grid cells contain data that represent extreme outliers, based on a Cook's distance test.

Number of grid cells removed due to outlying observations according to a Cook's distance test: 0

Final list of predictors that did not pass the model's three assumptions:
['ISO_YN']

################## Full Configuration Output Begins ##################

---------- MODEL CALIBRATION (Training Data): Full Configuration ----------

Range of log-likelihood scores from 30 training runs of the Full model:
Maximum Score: 29.400950489623256
Median Score: 22.43672632550772
Minimum Score: 12.68783307947399

Range of log-likelihood scores of the Null model:
Maximum Score: -22.49203432219565
Median Score: -22.49203432219565
Minimum Score: -22.492034322195764

Number of times (out of 30) the Full model possesses a greater goodness-of-fit: 30
Number of times (out of 30) the Full model's outperformance of the Null model is statistically significant: 30

Median Log-Likelihood Ratio, Full model vs. Null model: 89.85752129540674
p-value of the Median Log-Likelihood Ratio: 1.22322214578621e-07

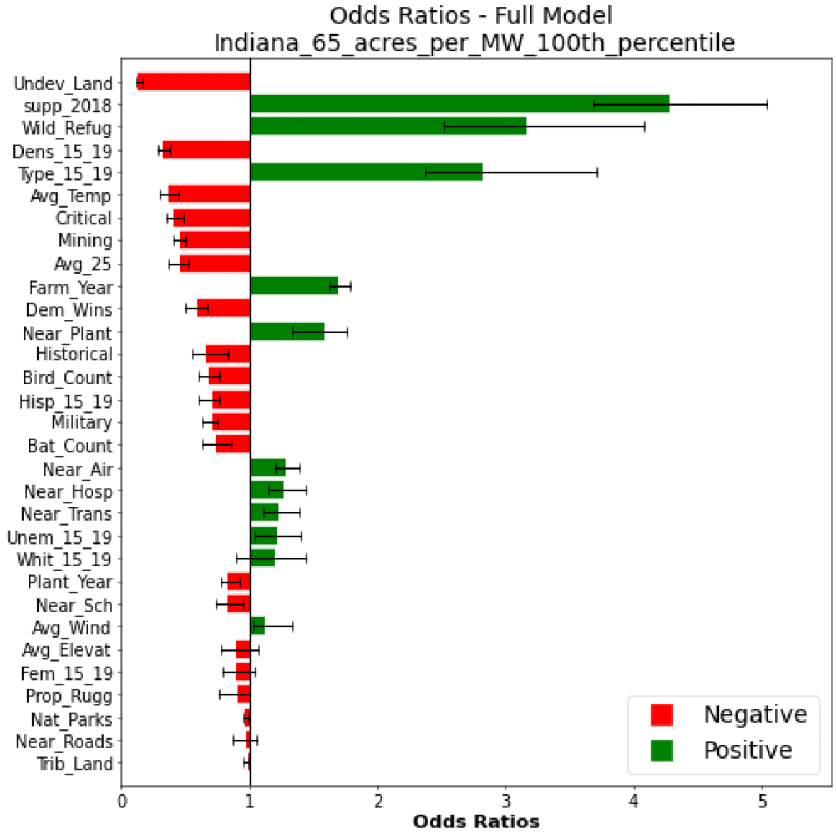
Range of McFadden Adjusted Psuedo R-Squared statistics for the Full model:
Minimum Pseudo R-Squared: 0.1413774919653077
Median Pseudo R-Squared: 0.5748150861989827
Maximum Pseudo R-Squared: 0.8844457787523501

The following dataframe summarizes the coefficients and odds ratios obtained from fitting the Full model to the aggregated dataset. Predictors are ranked by the magnitude of their coefficients to convey strength of association:

Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Undev_Land	0.117942	0.134264	0.167498	-2.007948	1
supp_2018	3.683232	4.277036	5.037322	1.453260	2
Wild_Refug	2.516396	3.162788	4.086971	1.151454	3
Dens_15_19	0.289170	0.325327	0.374990	-1.122926	4
Type_15_19	2.374538	2.826305	3.715410	1.038970	5
Avg_Temp	0.306526	0.373221	0.451746	-0.985585	6
Critical	0.347171	0.401288	0.480725	-0.913076	7
Mining	0.407635	0.461046	0.503785	-0.774257	8
Avg_25	0.368538	0.461121	0.526799	-0.774095	9
Farm_Year	1.628095	1.699741	1.785360	0.530476	10
Dem_Wins	0.499970	0.598039	0.671992	-0.514099	11
Near_Plant	1.333962	1.585337	1.762817	0.460797	12
Historical	0.553178	0.662843	0.824705	-0.411217	13
Bird_Count	0.602671	0.681959	0.770207	-0.382786	14

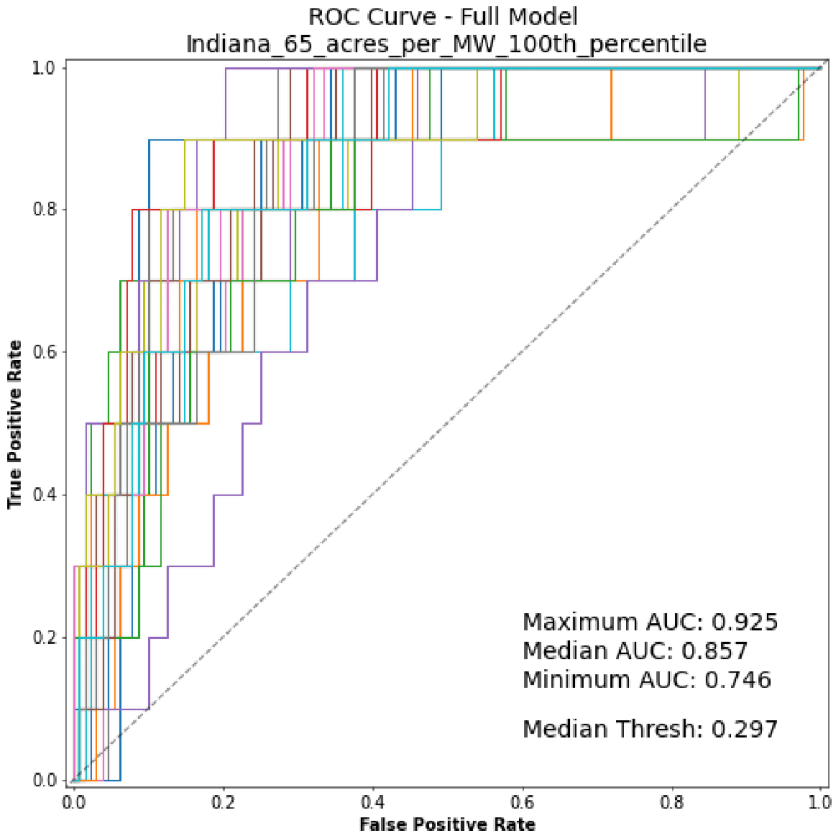
Hisp_15_19	0.604509	0.707557	0.759763	-0.345937	15
Military	0.628797	0.709099	0.753683	-0.343760	16
Bat_Count	0.637457	0.731975	0.853490	-0.312009	17
Near_Air	1.201431	1.285388	1.394982	0.251060	18
Near_Hosp	1.143740	1.264661	1.446436	0.234804	19
Near_Trans	1.107116	1.231330	1.389561	0.208095	20
Unem_15_19	1.046276	1.211820	1.401106	0.192124	21
Whit_15_19	0.901153	1.206441	1.445489	0.187674	22
Plant_Year	0.784104	0.837579	0.924114	-0.177240	23
Near_Sch	0.741542	0.837595	0.949223	-0.177220	24
Avg_Wind	1.026242	1.124632	1.336692	0.117456	25
Avg_Elevat	0.783076	0.894069	1.075098	-0.111973	26
Fem_15_19	0.787410	0.895429	1.040674	-0.110453	27
Prop_Rugg	0.766642	0.904545	0.997811	-0.100324	28
Nat_Parks	0.955519	0.967840	0.984540	-0.032689	29
Near_Roads	0.866377	0.979548	1.051215	-0.020664	30
Trib_Land	0.948605	0.990141	0.996208	-0.009908	31

Odds Ratio chart generated from the 30 Full model runs with the training data:



---------- MODEL Validation (Testing Data): Full Configuration ----------

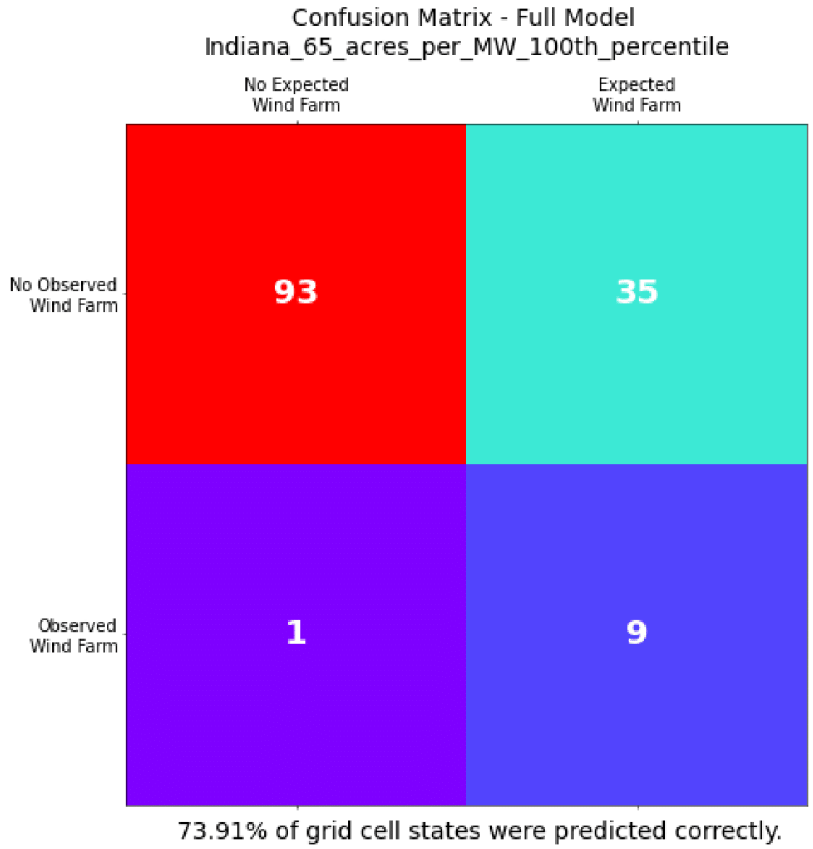
ROC curves generated from the 30 Full model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Full model:
Minimum AUC: 0.746875
Median AUC: 0.8570312499999999
Maximum AUC: 0.92578125

Range of optimal threshold classifications for the Full model:
Minimum Threshold: 0.05354667360507323
Median Threshold: 0.2976179369053258
Maximum Threshold: 0.7746442656427992

Median Confusion Matrix of the Full model's predictive accuracy:



Below are the range of confusion matrix results from the
30 Full model runs with the testing data:

Lower Quartile confusion matrix:
[[84 44]
 [2 8]]
Lower Quartile proportion of correctly predicted grid cell states by the Full model:
0.6666666666666666

Median confusion matrix:
[[93 35]
 [1 9]]
Median proportion of correctly predicted grid cell states by the Full model:
0.7391304347826086

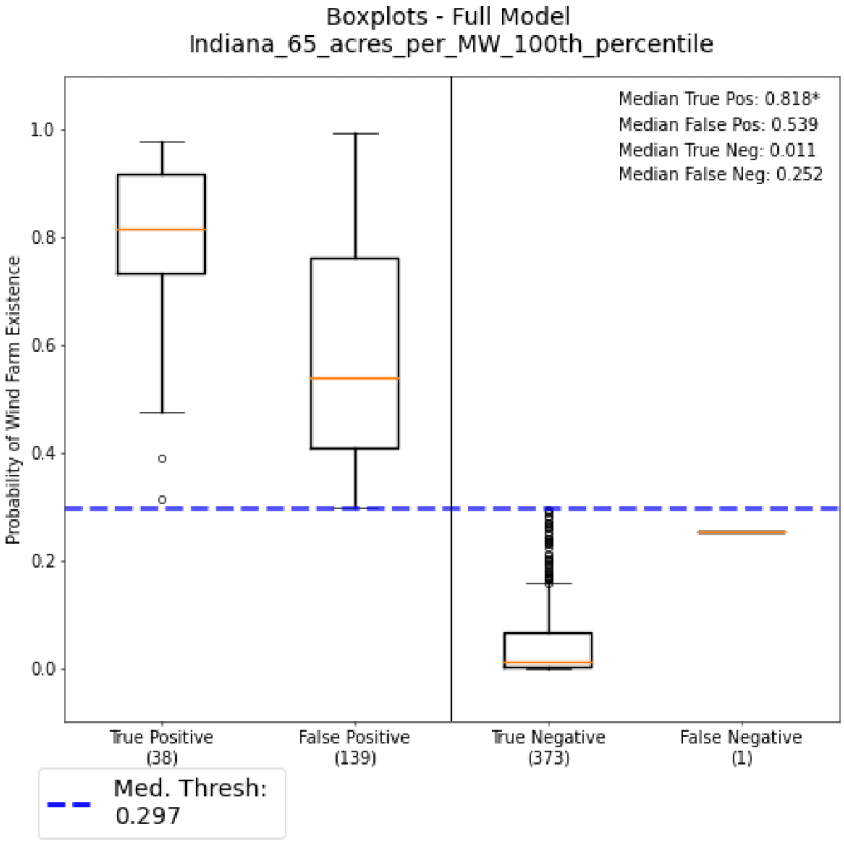
Upper Quartile confusion matrix:
[[107 21]
 [3   7]]
Upper Quartile proportion of correctly predicted grid cell states by the Full model:
0.8260869565217391

---------- BOXPLOT CONSTRUCTION (All Data): Full Configuration ----------

Grid cell classifications from executing the trained and tested Full model
over all grid cells in Indiana:

Number of True Positive Grid Cells: 38
Number of False Positive Grid Cells: 139
Number of True Negative Grid Cells: 373
Number of False Negative Grid Cells: 1

Boxplot of grid cell probabilities in each classification:



Median probabilities of wind farm existence for each grid cell classification.
An asterisk indicates a Mann-Whitney U-test result that is statistically significant (p<0.05):

Median False Pos: 0.539
Median True Pos: 0.818*
Median False Neg: 0.252
Median True Neg: 0.011

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:
U-statistic = 1188.0
p-value = 2.1123801686068578e-07

Mann-Whitney Statistic - True Negative vs False Negative:
U-statistic = 361.0
p-value = 0.10703885982450037

---------- MAP CONSTRUCTION: Full Configuration ----------

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\WiFSS_Surfaces/Hexagon_Grid_65_acres_per_MW_100th_percentile_Indiana_Full.gdb\Hexagon_Grid_65_acres_per_MW_100th_percentile_Indiana_Full_Map

Total (Percentage) of all grid cells over Indiana that exist in hotspots:
39 (7.08%)
Total (Percentage) True Positive grid cells over Indiana that exist in hotspots:
18 (47.37%)
Total (Percentage) False Positive grid cells over Indiana that exist in hotspots:
21 (15.11%)

A4: Example console output from running WiFSS-LRCA's Cellular Automata component:

Console output

```
---------- DATASET SELECTION AND SETUP ----------

NOTE: The Logistic_Regression_Model.py script MUST be executed
prior to running the Cellular_Automata_Model.py script.
If one wishes to execute the model over states that
contain zero commercial wind farms (Louisiana, Mississippi
Alabama, Georgia, South Carolina, Kentucky), states that possess
wind farms in only one grid cell at all but the highest spatial
resolutions (Arkansas, Florida, Virginia, Delaware, Connecticut,
New Jersey, Tennessee), or states at low spatial resolutions at which
too many predictors were removed due to collinearity (Rhode Island
at the 100th or 80th percentile), the Logistic_Regression_Model.py
script must be executed for the CONUS.

Specified study region: Indiana
Specified wind farm density: 85 acres/MW
Specified wind power capacity: 100th percentile (525 MW)

The constraints and neighborhood effects from the previous model run
have not been changed. A hexagonal neighborhood range of 3
grid cell(s) is thus retained, along with the following constraints:
['No wind farms within 2000 meters of an airport.', 'No wind farms within 15000 meters of a
power plant.', 'No wind farms within 500 meters of a major road.', 'No wind farms more than
20000 meters from a major road.', 'No wind farms within 500 meters of a major transmission
line.', 'No wind farms more than 20000 meters from a major transmission line.', 'No wind
farms in grid cells with an average wind speed 80 meters above ground less than 3 meters
per second.', 'Wind farm development is prohibited in grid cells shared by military bases.',
'Wind farm development is prohibited in grid cells shared by national parks.', 'Wind farm
development is prohibited in grid cells shared by USFWS critical habitats.', 'Wind farm
development is prohibited in grid cells shared by USFWS wildlife refuges.', 'Wind farm
development is prohibited in grid cells shared by tribal land.']

The following predictor configurations were selected by the user:
['Full', 'No_Wind']

The user specified a custom gained wind farm capacity of 2500 MW every 5 years,
which based on model resolution (100th percentile, 525MW) translates to 5 new wind farms
per model iteration.

---------- SCENARIO CONSTRUCTION ----------

The following are the scenarios selected by the user
(see Model Instructions for scenario details):
['CLIMATE_CHANGE', 'DEMOGRAPHIC_CHANGES', 'SOCIOPOLITICAL_LANDSCAPE',
'NATURAL_AND_CULTURAL_PROTECTION']

---------- MODEL PROJECTION: Null ----------

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\Constraints_and_Neighborhood_Effects\Hexago
n_Grid_85_acres_per_MW_100th_percentile_Indiana.gdb\Hexagon_Grid_85_acres_per_M
W_100th_percentile_Indiana_Constraints_Neighborhoods
```

---------- MODEL PROJECTION: Full ----------

Coefficient changes under the selected scenario when applying the Full predictor configuration:

	Predictors	Coeff_Change_(%)	...	Coeff_2045	Coeff_2050
0	Avg_Elevat	0	...	-0.523482	-0.523482
1	Avg_Temp	10	...	-0.294820	-0.265338
2	Avg_Wind	10	...	2.118074	2.329882
3	Bat_Count	-10	...	-0.130263	-0.143289
4	Bird_Count	-10	...	-1.610006	-1.771007
5	Critical	-10	...	0.241107	0.216996
6	Type_15_19	0	...	0.793625	0.793625
7	Historical	-10	...	-0.157192	-0.172912
8	Military	0	...	0.589572	0.589572
9	Mining	0	...	-1.431692	-1.431692
10	Nat_Parks	-10	...	-0.618871	-0.680758
11	Near_Air	0	...	0.285588	0.285588
12	Near_Plant	0	...	0.080833	0.080833
13	Near_Roads	0	...	0.156328	0.156328
14	Near_Sch	0	...	-0.258780	-0.258780
15	Near_Trans	0	...	0.523574	0.523574
16	Near_Hosp	0	...	0.284254	0.284254
17	Fem_15_19	10	...	0.006586	0.007245
18	Hisp_15_19	10	...	0.529305	0.582236
19	Avg_25	10	...	-0.069519	-0.062567
20	Whit_15_19	-10	...	0.506229	0.455606
21	Dens_15_19	0	...	-0.034292	-0.034292
22	Plant_Year	0	...	-0.324657	-0.324657
23	Dem_Wins	10	...	-0.285427	-0.256884
24	supp_2018	10	...	1.457957	1.603753
25	Prop_Rugg	0	...	-1.061243	-1.061243
26	Trib_Land	-10	...	-0.175184	-0.192702
27	Unem_15_19	0	...	0.498987	0.498987
28	Wild_Refug	-10	...	0.413504	0.372153
29	Farm_Year	0	...	-0.196794	-0.196794

[30 rows x 9 columns]

Predictors removed from the model based on having a constant value in all grid cells: None

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\Constraints_and_Neighborhood_Effects\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana.gdb\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Constraints_Neighborhoods

QADI table produced by comparing projections forced by the
Null versus Full predictor configurations:

Grid Cell Projections (Null Configuration)

Grid Cell Projections (Full Configuration)

	NoFarm	Y(2025)	Y(2030)	Y(2035)	Y(2040)	Y(2045)	Y(2050)	Sum
NoFarm	357	0	1	1	2	2	1	364
Y(2025)	0	3	2	0	0	0	0	5
Y(2030)	0	1	2	2	0	0	0	5
Y(2035)	0	0	0	1	3	1	0	5
Y(2040)	0	1	0	1	0	1	2	5
Y(2045)	3	0	0	0	0	0	2	5
Y(2050)	4	0	0	0	0	1	0	5
Sum	364	5	5	5	5	5	5	394

Quantity Disagreement: 0 *Allocation Disagreement: 31*

QADI Index: 0.079

---------- MODEL PROJECTION: No_Wind ----------

Coefficient changes under the selected scenario when applying the No_Wind predictor configuration:

	Predictors	Coeff_Change_(%)	...	Coeff_2045	Coeff_2050
0	Avg_Elevat	0	...	0.074983	0.074983
1	Avg_Temp	10	...	-0.320944	-0.288849
2	Bat_Count	-10	...	-0.255154	-0.280669
3	Bird_Count	-10	...	-1.802695	-1.982965
4	Critical	-10	...	0.222165	0.199948
5	Type_15_19	0	...	0.674383	0.674383
6	Historical	-10	...	-0.235184	-0.258702
7	Military	0	...	0.601035	0.601035
8	Mining	0	...	-1.223741	-1.223741
9	Nat_Parks	-10	...	-0.407217	-0.447939
10	Near_Air	0	...	0.256439	0.256439
11	Near_Plant	0	...	0.151401	0.151401
12	Near_Roads	0	...	0.112543	0.112543
13	Near_Sch	0	...	-0.207303	-0.207303
14	Near_Trans	0	...	0.524056	0.524056
15	Near_Hosp	0	...	0.442366	0.442366
16	Fem_15_19	10	...	-0.026683	-0.024015

17	Hisp_15_19	10	...	0.704906	0.775397
18	Avg_25	10	...	-0.101550	-0.091395
19	Whit_15_19	-10	...	0.516311	0.464680
20	Dens_15_19	0	...	-0.179760	-0.179760
21	Plant_Year	0	...	-0.320183	-0.320183
22	Dem_Wins	10	...	-0.264187	-0.237768
23	supp_2018	10	...	1.826726	2.009399
24	Prop_Rugg	0	...	-1.605261	-1.605261
25	Trib_Land	-10	...	-0.113987	-0.125385
26	Unem_15_19	0	...	0.566459	0.566459
27	Wild_Refug	-10	...	0.138113	0.124301
28	Farm_Year	0	...	-0.175766	-0.175766

[29 rows x 9 columns]

Predictors removed from the model based on having a constant value in all grid cells: None

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\Constraints_and_Neighborhood_Effects\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana.gdb\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Constraints_Neighborhoods

QADI table produced by comparing projections forced by the Null versus No_Wind predictor configurations:

Grid Cell Projections (Null Configuration)

Grid Cell Projections (No_Wind Configuration)

	NoFarm	Y(2025)	Y(2030)	Y(2035)	Y(2040)	Y(2045)	Y(2050)	Sum
NoFarm	357	0	1	2	2	2	0	364
Y(2025)	0	3	2	0	0	0	0	5
Y(2030)	0	1	2	1	1	0	0	5
Y(2035)	0	1	0	1	1	1	1	5
Y(2040)	2	0	0	0	1	0	2	5
Y(2045)	2	0	0	1	0	1	1	5
Y(2050)	3	0	0	0	0	1	1	5
Sum	364	5	5	5	5	5	5	394

Quantity Disagreement: 0 *Allocation Disagreement: 28*

QADI Index: 0.071