Markov Random Field for Wind Farm Planning

Hale Cetinay∗, Taygun Kekeç†, Fernando A. Kuipers∗ and D. M. J. Tax†

∗Network Architectures and Services, Delft University of Technology, Delft, The Netherlands
Email: {H.Cetinay-iycil, F.A.Kuipers}@tudelft.nl
†Pattern Recognition and Bioinformatics Laboratory, Delft University of Technology, Delft, The Netherlands
Email: taygunkekec@gmail.com, D.M.J.Tax@tudelft.nl

Abstract—Many countries aim to integrate a substantial amount of wind energy in the near future. This requires meticulous planning, which is challenging due to the uncertainty in wind profiles. In this paper, we propose a novel framework to discover and investigate those geographic areas that are well suited for building wind farms. We combine the key indicators of wind farm investment using fuzzy sets, and employ multiple-criteria decision analysis to obtain a coarse wind farm suitability value. We further demonstrate how this suitability value can be refined by a Markov Random Field (MRF) that takes the dependencies between adjacent areas into account. As a proof of concept, we take wind farm planning in Turkey, and demonstrate that our MRF modeling can accurately find promising areas.

Keywords—quality of wind; criteria for wind farms; spatial relations in wind farm investments; planning for wind power integration

I. INTRODUCTION

Over the last decades, our society has developed a more comprehensive understanding of environmentally-friendly approaches to energy generation, urging us to focus more on sustainable energy sources, such as wind energy. As a result, the integration of renewable energy goals into their long-term policies has been the priority of many countries. One example is the policy by the Ministry of Energy and Natural Resources of the Turkish Republic [1], which aims to attain a wind farm capacity of 20 GW by 2023.

A large-scale integration of wind farms will challenge the main power grid, which once was built without renewables. Thus, power grid operators must carefully analyze the expansion scenarios for wind farms and plan the necessary improvements to ensure that the electric power grid will not succumb to a large reflex to wind energy.

In recent years, many studies have been conducted to evaluate potential geographic areas for wind farms [2], [3]. A subset of these studies considers only wind speed measurements as a basis for the assessment [3], [4], ignoring any economic or environmental restrictions for wind farms. Some studies propose including a list of environmental criteria for a more realistic integration [5], [6], [7], but the accuracy of these criteria-based methods depends directly on the input data, such as the wind power characteristics, which are hard to determine exactly. Inaccuracies in input parameters can propagate easily leading to imprecise modeling. Moreover, the assessments are usually carried out for each area independently, ignoring any neighboring relations, while the surrounding geographic factors and investments play a role in deciding on the investment in a wind farm [8]. Motivated by such reasons, we have developed a novel spatially-aware model for the wind farm suitability of areas.

The remainder of this paper is organized as follows. In Section II, we construct a grid-based model on the Cartesian plane for representing a geographic area. Subsequently, we model the indicators of wind farms via fuzzy sets and obtain an initial suitability value using multiple-criteria decision analysis. In Section III, we explain our Markov Random Field (MRF) approach for providing a refined spatially-aware suitability value for wind farms. To the best of our knowledge, we are the first to combine the fuzzy logic and multiple-criteria decision analysis with MRF to find promising areas for wind farms. Finally, a comprehensive case study is provided in Section IV. The results of the case study suggest that our wind farm suitability methodology provides fine-grained information for wind farm investment.

II. MODELING WIND FARM SUITABILITY

In this section, we adopt a grid-based reconstruction of geographic areas and quantify the key criteria involved in wind farm investment.

A. A grid-based model on the two-dimensional Cartesian plane

We use a two-dimensional grid-based model of equally-sized rectangles to represent the spherical geographic area under consideration (See Figure 1). We assume that we are given a set \( N \) of points \( k \) in spherical-world coordinates, composed of latitude \( \phi(k) \) and longitude \( \lambda(k) \) values. Each \( k \in N \) of these spherical points is projected onto a two-dimensional plane using a linear mapping, where the horizontal \( X(k) \) coordinate is obtained using the degree of longitude \( \lambda(k) \) of \( k \) and the vertical \( Y(k) \) coordinate of point \( k \) can be computed based on the degree of the latitude \( \phi(k) \):

\[
X(k) - 1 = \frac{\lambda(k) - \beta_X}{\alpha_X}
\]

\[
Y(k) - 1 = \frac{\phi(k) - \beta_Y}{\alpha_Y}
\]

where \( \alpha_X (\alpha_Y) \) is the scaling between the degree of longitude (latitude) and the horizontal (vertical) coordinate, and \( \beta_X (\beta_Y) \).
defines the value of longitude (latitude) of the first coordinate on the two-dimensional plane.\footnote{Although the linear projection is not an accurate representation of the Earth’s surface, the projection has the advantage of being geometrically simple and therefore is widely used [9].}

### B. Quantifying the elementary criteria for wind farms

The decision to invest in a wind farm at a certain location depends on two main criteria: wind power potential and investment disincentives. The wind power generative potential of an area can be captured by indicators such as average wind speed, wind power density, and the capacity factor of a prospective wind turbine. On the other hand, disincentive indicators can include high values of land cost or altitude levels, and the proximity to urban areas.

Ideally, an investor should review all $M$ indicators $\{r_1(k), \ldots, r_M(k)\}$ before investing in a wind farm at an area $k$. However, in practice, this review process is often not performed due to the difficulty in dealing with the uncertainty, vagueness, or the lack of information in the practical decision process. In this paper, we model those indicators of a wind farm investment using fuzzy sets [10], which enables us to explicitly deal with uncertainty. Different than Boolean logic, in which the truth can only be the integer values 0 or 1, fuzzy logic can handle the concept of partial truth during a decision process.

We use increasing fuzzy function $F(r_i(k))$ in (3) and decreasing fuzzy function $F(r_i(k))$ in (4) to evaluate the satisfaction degree of each indicator $r_i(k)$ for a wind farm in area $k$. The increasing fuzzy function represents the incentive indicators, whereas the decreasing fuzzy function represents the disincentive indicators. The resulting fuzzy membership degrees take values between 0 and 1 corresponding to the unsatisfactory and full-satisfactory evaluations of an area $k$, respectively.

\[
F(r_i(k)) = \begin{cases} 
0 & \text{if } r_i(k) < q_i, \\
\frac{r_i(k) - q_i}{p_i - q_i} & \text{if } q_i \leq r_i(k) \leq p_i, \\
1 & \text{if } r_i(k) > p_i, 
\end{cases} \quad (3)
\]

\[
F(r_i(k)) = \begin{cases} 
1 & \text{if } r_i(k) < p_i, \\
\frac{r_i(k) - p_i}{q_i - p_i} & \text{if } p_i \leq r_i(k) \leq q_i, \\
0 & \text{if } r_i(k) > q_i, 
\end{cases} \quad (4)
\]

where for each indicator $r_i$, $q_i$ and $p_i$ correspond to the thresholds of unsatisfactory and full-satisfactory evaluations, respectively.

### C. Multiple-criteria decision analysis of wind farms

Since we have to deal with and optimize for multiple fuzzy parameters, we focus on multiple-criteria decision analysis in this section.

The perspective of an investor is important when assessing the criteria for a wind farm. For instance, an investor could consider a worst-case scenario of the related indicators or could, as the other extreme, consider a best-case scenario. Following [5], [6], we employ fuzzy logic aggregation operators to allow for variability in perspective. We use the and $\land$ and the or $\lor$ aggregation operators to map two extreme cases of an investor’s stance on multiple-criteria decisions. The and operator $\land$ of the fuzzy membership degrees requires the satisfaction of all desired criteria, in other words, a conservative perspective when evaluating the satisfaction degrees of related indicators:

\[
\land(k) = \min_{1 \leq i \leq M} F(r_i(k)) \quad (5)
\]

The or operator $\lor$ is appropriate to model a more optimistic or lenient perspective. The implementation of the or operator in (6) passes over the less satisfactory indicators:

\[
\lor(k) = \max_{1 \leq i \leq M} F(r_i(k)) \quad (6)
\]

Lastly, to model the perspective of an investor in between those two extreme cases, we can use a weighted mean operator $\mu$ in (7):

\[
\mu(k) = \sum_{i=1}^{M} w_i F(r_i(k)) \quad (7)
\]

where the ultimate decision is the convex combination of the satisfaction degrees of the decision indicators, such that $\sum_i w_i = 1$.

By applying this aggregation operator to each rectangle $k \in \mathcal{N}$, we obtain an elementary suitability value $\hat{z}_k \in [0, 1]$ of an individual area bounded by that rectangle $k$. However, such elementary suitability values are not fully representative yet, due to reasons mentioned in the next section, where we describe a random field approach to model a more fine-grained spatially-aware suitability.

### III. Spatial Suitability Modeling with Markov Random Field

The suitability values $\hat{z}_k$ calculated in the previous section provide an initial suitability estimate for a rectangle $k$ in the grid-based model. However, we have to take the following into account: Firstly, input parameters to calculate an elementary suitability can exhibit significant measurement noise and the parameters related to wind energy potential can deviate due to inaccurate measuring instruments [11]. Secondly, the proposed elementary suitability value may not be unique, since different degrees of freedom exist in the specification of the decision making process. Lastly, the construction of a grid-based model requires the projection of a geolocation onto the Cartesian plane, introducing quantification errors that must be dealt with.
where $p$ denotes the discrete suitability states a node can take, and $I$ refers to an element of the Markovian graph over our grid-based model (See Figure 3).

Motivated by these reasons, in this section, we refine the computed elementary suitability values $\hat{z}_k$ by modeling the true suitability values in a MRF, which can characterize the local spatial interactions in elementary suitability values.

We first assume that there exist some underlying unobserved suitability values $x = \{x_1, x_2, ..., x_k, ..., x_N\}$, but we observe a noisy version of them: $\hat{z} = \{\hat{z}_1, \hat{z}_2, ..., \hat{z}_k, ..., \hat{z}_N\}$. Thus, the correlation among $x_k$ and $\hat{z}_k$ is expected to be high. Let $B(k)$ be the Markov blanket of a node $k$. Different Markov blanket functions are shown in Figure 2. Then, if we assume the conditional independence of $x_k$, given the suitability values of the neighboring nodes $B(k)$, $x_k$ can take a value independently of the other nodes: $N \setminus B(k)$.

For our scenario, we adopt a 4-node neighborhood, and denote the Markov blanket of $k$ as $B_4(k)$. This defines a Markov Random Field over our grid-based model. Figure 3 illustrates a Markov Random Field modeling for Turkey. Horizontal links are pairwise interactions between nodes and vertical links represent the terms that force similarity between observed elementary and unknown spatially-aware suitability values.

Due to the conditional independence properties of the Markov blanket, we can write the likelihood $p(\hat{z}|x)$ as

$$p(\hat{z}|x) = \prod_{k=1}^{N} \prod_{j=1}^{S} p(\hat{z}_k|x_k = j) \mathbb{1}(x_k)$$  

(8)

where $p(.)$ is the probability function, $S$ is the number of discrete suitability states a node can take, and $\mathbb{1}(x_k)$ is the indicator vector, where all components are zero except for component $x_k$, which is one.

To measure the suitability-state compatibility between neighboring nodes in the Markovian graph, we specify a prior, data-independent, rule. For this purpose, we use the Ising model [12] [13], which penalizes the state incompatibility between different nodes. Deciding this measure of compatibility for all links of the graph defines the unary potential $p(\hat{z}_k|x_k = j) = \exp(-|x_k - z_k|)$ in (8) and the marginal probability $p(x)$ in (9):

$$p(x) = \prod_{k=1}^{N} \sum_{s \in B_4(x_k)} \psi(x_k, x_s)$$  

(9)

where potential $\psi$ encourages smoother solutions by forcing $x_s$ and $x_k$ to be in the same suitability-state configuration:

$$\psi(x_k, x_s) = \frac{\exp(-\gamma |x_k - x_s|)}{G}$$  

(10)

where $G$ is the normalization term that sums over all possible state configurations of $(k, s)$ ensuring that $\psi(x_k, x_s)$ probabilities sum to one, and $\gamma$ is the smoothness factor. The smoothness factor controls the strength of the imposed prior. For instance, $\gamma = 0$ corresponds to using no prior at all.

Adopting the pairwise interaction model along with Ising priors, our goal is to maximize the posterior $p(x|\hat{z})$, which can be computed with the help of Bayes’ theorem:

$$p(x|\hat{z}) = \frac{p(\hat{z}|x)p(x)}{p(\hat{z})}$$  

(11)

Since we maximize over $x$, we ignore the term $p(\hat{z})$ and write the Maximum A Posteriori (MAP) estimate $x_{MAP}$ of $x$ as:

$$x_{MAP} = \arg\max_x p(\hat{z}|x)p(x)$$  

(12)

Such a spatial modeling through MRF refines the local elementary suitability estimates $\hat{z}$ to achieve globally-consistent suitability values $x$.

To derive MAP suitability estimates, (12) has to be maximized: a brute-force search is out of the question even for medium-sized problems, where $N$ is in the order of hundreds, since $S$ discrete suitability states lead to $S^N$ different configurations. Thus, for the solution of the multi-state case ($S > 2$), an exact MAP solution is often not applicable. In that case, we resort to an Iterated Conditional Modes (ICM) algorithm for finding the MAP solution. ICM uses a greedy strategy to find the local maximum of (12). The idea can be stated as follows: the algorithm starts with an initial estimate of the suitability values, and then for each node $k \in N$, the state configuration that gives the highest increase in the posterior probability is chosen to be the current suitability state. This state-update procedure is continued until there are no changes in the state configuration of the nodes. This convergence is guaranteed by the ICM algorithm [16]. Even for problems with many states and nodes, the convergence of the ICM algorithm is fast, since the convergence rate is linear in $S$ and $N$.

3Rectangle $k$ refers to an element of the grid-based model on the two-dimensional Cartesian plane, whereas node $k$ refers to the corresponding element of the Markovian graph over our grid-based model (See Figure 3).

4Although performing maximization over the general random fields is shown to be NP-Hard [14], in the binary problem case, i.e., when $S = 2$, it was shown that maximization in (12) can be treated as a combinatorial maximum-flow minimum-cut problem on a graph [15].
IV. Case Study

In this section, we present a case study to demonstrate the merits of our framework. We obtained country-wide wind data of Turkey. Additionally, we collected the geographic locations in Turkey for which licenses for wind farm construction are held by an investor. More details on our data collection procedure can be found in [17].

A. A grid-based model of Turkey

Based on our wind measurement data set, each rectangle in the grid-based model corresponds to a 6 km × 6 km area. As the length of a degree of latitude does not change (approximately 111.2 km); the scaling factor $\alpha_Y$ in (2) is taken as $\frac{6}{111.2} \simeq 0.053$. On the other hand, the length $l_{\lambda}(\phi')$ of a degree of a longitude depends on its degree of latitude $\phi'$ and can be approximated as

$$l_{\lambda}(\phi') = \cos(\phi') \times 111.3 \text{ km} \quad (13)$$

Using (13), around the southern points of Turkey at $36^\circ$, the length of a degree of a longitude is approximately 90 km. For ease of demonstration, in this study, we set the length of a degree of a longitude at 90 km, thus, the scaling factor $\alpha_{X}$ in (1) is taken as $\frac{29}{90} \simeq 0.066$. For other purposes, it is possible to decrease the projection error by choosing variable lengths of a degree of a longitude.

The degrees of the latitude and longitude at the first coordinate on the two-dimensional plane are defined according to the position of the geographic area. To fully enclose Turkey in our projection, we choose $43^\circ$ N latitude and $25^\circ$ E longitude as the first coordinate $(1, 1)$. The final equations to construct the grid-based model onto a two-dimensional Cartesian plane are given in (14) and (15).

$$X(k) - 1 = \frac{\lambda(k) - 25^\circ}{0.066}$$

$$Y(k) - 1 = \frac{43^\circ - \phi(k)}{0.053} \quad (15)$$

B. Quantifying the wind farm potential in Turkey

Key indicators [18] to capture the wind energy potential at an area $k$ are the average wind speed $\nu(k)$, the wind power density $\rho(k)$, and the capacity factor $\eta(k)$ of the probable wind turbine at that area. Due to the positive correlation between the promising wind energy potential and the investment criteria for wind farms, the increasing fuzzy function in (3) is used to calculate corresponding satisfaction degrees of those indicators.

The landscape of Turkey contains heterogeneously distributed mountainous regions with varying altitudes. High altitude regions and high slope lands are undesirable for establishing wind farms. Thus, we use the altitude $\alpha(k)$ of a geographic area $k$ as a disincentive indicator for wind farms. Due to the negative correlation between the altitude and the investment criteria for wind farms, the decreasing fuzzy function in (4) is used.

![Figure 4: The membership functions of the selected indicators of wind farms.](image)

![Figure 5: The CDFs of the membership degrees of selected indicators in Turkey.](image)

The resulting membership functions of selected indicators are shown in Figure 4. The full-satisfactory and unsatisfactory thresholds of indicators are determined based on related work [7], [17].

Next, for each area $k$, we calculate the membership degrees of the selected indicators for wind farms. The Cumulative Distribution Functions (CDF) of all areas in Turkey are shown in Figure 5. For each indicator, approximately 15% of all areas have 0 membership values, corresponding to the unsatisfactory evaluations for wind farms.

Finally, to represent the preference of an investor for the multiple-criteria decision, we use the three aggregation operators presented in Section II-C. In the weighted mean operator (7), each of the 4 indicators is given an equal weight of 0.25. Figure 6 depicts the Cumulative Density Functions of the elementary suitability values with different aggregation operators. The and operator represents the conservative evaluation: 35% of the areas in Turkey have the minimum (0) elementary suitability value, whereas the or operator represents the optimistic evaluation: 25% the areas in Turkey have the maximum (1) elementary suitability value. On the other hand, the weighted mean operator represents a smoother
evaluation: Under this operator, almost none of the areas has an extreme \( \{0, 1\} \) elementary suitability value.

C. Spatially-aware suitability for wind farms in Turkey

We apply the Markov Random Field described in Section III to obtain the spatially-aware suitability values of wind farms in Turkey. Elementary suitability values for wind farms in Turkey are determined using the weighted mean operator. Our ICM algorithm visits the nodes sequentially. The number of nodes in the Markovian graph \( N = 21,983 \) and the number of suitability states a node can take is set to \( S = 256 \).

**Qualitative Results:** Figure 7 depicts local patches extracted from distinctive regions in Turkey. The rectangles in the grid-based model are colored according to their suitability value for wind farms. Lighter colors represent higher suitability values. We observe that the suitability values for wind farms are particularly high in the Izmir region. This Aegean region benefits from the strong south-westerly wind, Lodos. On the other hand, the Muş region has lower suitability values as a consequence of its relatively high altitude and continental climate.

The spatially-aware suitability values for wind farms in the second row of Figure 7 include not only the individual characteristics of a specific area, but also its neighboring areas. Thus, the suitability map by MRF seems more smooth and more globally consistent. As an example, even though an area surrounded by disincentives (such as high mountains) can have higher values of elementary suitability, its spatially-aware suitability value could be lowered due to the neighboring disincentives (See Figure 7 (c)).

**Quantitative Results:** To assess the practicability of the proposed spatially-aware suitability values for wind farms, we investigate the suitability values of licensed wind farm locations in Turkey. Figures 8 and 9 depict the histograms of the suitability values in all geographic areas and in the licensed areas for wind farms in Turkey, respectively.

The distribution of spatially-aware suitability values for licensed wind farm locations in Figure 9 follows an increasing behavior. In particular, most of the licensed wind farm locations have high suitability values. However, there are few regions where the suitability values are extraordinarily small. The calculated suitability values are insufficient for a full explanation of the license acquisition behavior in those regions. These might be overcome by including more socio-economic indicators in the suitability analysis.
Next, we compare the elementary and spatially-aware suitability values of the licensed geographic locations of wind farms to analyze whether a spatially-aware suitability modeling with MRF captures additional clues about the investing behavior. Using all areas, we first compute the expected $E[x]$ suitability of all areas. The expected $E[x]$ suitability for wind farms in Turkey, the vertical dashed-red line in Figure 8, corresponds to the suitability for a wind farm given that an investor made a random choice for a geographic area.

Subsequently, we calculate the tail probability $F_T$ in (16) which refers to the license acquisition event of an area whose suitability is smaller than the expected $E[x]$ suitability. We hypothesize that investors have access to a diverse set of common and privileged sources of information, such that their license acquisition behavior is a measure of true suitability and they make a better decision than a random choice for an area for wind farms. In that sense, we expect a lower value of tail probability $F_T$ in the licensed geographic locations of wind farms if we capture the true suitability values using MRF:

$$F_T = p(x_1 < E[x])$$  \hspace{1cm} (16)

where $x_1$ is the suitability value of a licensed area.

The tail probabilities $F_T$ in (16) are calculated as 0.11 and 0.08 in elementary and spatially-aware suitability models for the licensed geographic locations of wind farms. Thus, the suitability model with MRF decreases the tail probability more than 12% compared to an elementary suitability model for wind farms in Turkey. In addition to the reasons in Section III (i.e., the decrease in measurement noise and quantification errors), this decrease in the tail probability can be further explained by the spatial dependence of wind characteristics and wind farm investments. Investors could have a tendency to acquire a license of a region where the neighboring areas are already licensed and have a high suitability value, whereas they are less likely to invest in an area surrounded by disincentives. As a result, we can conclude that our spatially-aware modeling can help to estimate the effects of the unknown socioeconomic factors on the wind farm suitability and investment decision process.

V. Conclusion

In this paper, we have proposed a framework to calculate the suitability of a geographic area for wind farms. Given the wind energy potential and the incentive indicators of wind farms, initial suitability values for wind farms have been formed via fuzzy logic and multiple-criteria decision analysis. Subsequently, we refined those suitability values by a Markov Random Field model that can include the effects of the spatial relations on the wind farm suitability. Our results from a case study in Turkey show that such a spatially-aware modeling can estimate the suitability of a geographic area for wind farms.

For future work, a similar MFR approach could be developed to investigate the potential areas for solar parks. Another future direction could include incorporating the results of suitability analysis into a medium- or a long-term expansion planning of power grids. In particular, the calculated suitability values for wind farms, which demonstrate the likelihood of a wind farm investment in an area, could be used to construct possible wind power integration scenarios for power grids.

ACKNOWLEDGEMENT

This work was supported in part by Alliander N.V. and the Dutch Organization for Scientific Research (NWO; grant 612.001.301).

REFERENCES