<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>An end-user led approach for assessing the feasibility of small-scale wind energy projects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Miu, Luciana M.; Hayes, Barry P.; Djokic, Sasa Z.</td>
</tr>
<tr>
<td><strong>Publication Date</strong></td>
<td>2017-06-18</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>NUI Galway</td>
</tr>
<tr>
<td><strong>Link to publisher's version</strong></td>
<td><a href="https://doi.org/10.13025/S8192N">https://doi.org/10.13025/S8192N</a></td>
</tr>
<tr>
<td><strong>Item record</strong></td>
<td><a href="http://hdl.handle.net/10379/6847">http://hdl.handle.net/10379/6847</a></td>
</tr>
<tr>
<td><strong>DOI</strong></td>
<td><a href="http://dx.doi.org/10.13025/S8192N">http://dx.doi.org/10.13025/S8192N</a></td>
</tr>
</tbody>
</table>
An End-User Led Approach for Assessing the Feasibility of Small-Scale Wind Energy Projects

Luciana M. Miu  
Dept. of Chemical Engineering  
Imperial College  
London, UK  
l.miu16@imperial.ac.uk

Barry P. Hayes  
Engineering and Informatics  
NUI Galway  
Galway, Ireland  
barry.hayes@nuigalway.ie

Sasa Z. Djokic  
School of Engineering  
The University of Edinburgh  
Edinburgh, UK  
sasa.djokic@ed.ac.uk

Abstract—The successful realization of a wind energy project, i.e. a wind farm, wind plant or, in its simplest form, a single wind turbine generator, WTG, strongly depends on accurate assessment of wind energy resources at the target location. While larger wind energy projects include extensive wind resource measurements, this is typically not feasible for small-scale wind applications. This paper presents a simple linear regression model for evaluating wind energy resources at unmonitored sites, based on the correlational analysis of a limited number of sites where wind energy measurements are available. The methodology is illustrated on two regions in Scotland, where wind speed and wind direction measurements are publicly available from 26 UK Met Office stations. The presented approach requires only the coordinates of the end-user’s WTG location to estimate the available wind energy and incorporates a method to identify the best-fit WTGs for the calculated wind energy resource. For the selection of WTG, a database of 200 small-scale wind turbines available to UK end-users was constructed, reflecting the current state of the market.

Index Terms—Correlational analysis; spatio-temporal analysis; wind energy; wind resource; wind generation turbine.

I. INTRODUCTION

The ambitious renewable energy targets set by the Scottish Government have been one of the main catalysts for various microgeneration projects [1], within which small-scale wind energy projects (typically a single wind turbine installations) played an instrumental role. However, the development of micro and small-scale wind energy projects is strongly dependent on the accurate estimation of the wind energy resource. The typical end-users undertaking small-scale wind turbine generator (WTG) projects are neither able to perform a detailed measurements of the wind energy resource at the target WTG location, nor to afford the installation of a WTG without certainty of payback.

Typically, freely available wind energy resource estimation tools, such as wind atlases or input-based tools (e.g. [2]) only offer a first-pass overview of a site’s suitability for the development of a wind energy installation. Most tools available to Scottish users are either based on the Numerical Objective Analysis of Boundary Layer (NOABL, [3]), or Met Office Integrated Data Archive System (MIDAS, [4]), i.e. on two databases which have been reported to generate errors in the estimation of average wind speed between +10% and +20% [5], [6]. The use of these two tools could, therefore, lead to a significant overestimation of wind energy resource, which is very important, if not critical for assessing the return on investment of planned small-scale projects.

This paper presents a simple linear regression model for evaluating wind energy resources at unmonitored sites, based on a spatio-temporal correlational analysis of a limited number of sites where wind energy measurements are available. The model produces estimated wind speed time series (WSTS) and is illustrated on two regions in Scotland, for which publicly available hourly wind speed and wind direction measurements are available from a number of the UK Met Office stations [4]. The presented approach requires only the geographic coordinates of the planned location for installation of a WTG (i.e. end-users’ location).

The methodology is based on an assessment of inter-site distance and directional alignment of wind regimes as determining factors of the correlation between pairs of WSTS at different sites. The analysis was performed for both raw wind data and for “detrended” data (where mean values are removed), in order to identify the best data-processing method for the correlation and regression of the WSTS’. The presented results demonstrate that the developed model works well for sites experiencing uniform regional-scale wind regimes, as it was assumed previously in [7]. In such cases, the WSTS and annual average wind speeds are simulated with an accuracy of 90%, or higher. However, the model performance deteriorates for sites with highly individual and non-uniform wind climates.

Following the evaluation of the WSTS at a user-defined unmonitored location, further analysis is conducted to identify the best-fit WTG for installation. This part of the analysis is informed by an extensive study of the small-scale WTG market, where a 200-turbine database is built for further use.

The paper is structured as follows: Section II provides a brief literature review; Section III describes the methodology; Section IV gives the results; Section V lists main conclusions.
II. A BRIEF OVERVIEW OF EXISTING METHODS

Regression, autoregressive (AR) and autoregressive moving average (ARMA) models, as well as artificial neural networks (ANN) have been previously used for evaluating wind energy resources (typically wind speeds) at non-measured locations, all to varying levels of success. Examples range from the early use of inter-site correlation in [8] to the development of ANN models, capable of simulating hourly WSTS with reported errors below 5% [9]. Regression models form an important class of estimation methods, with perhaps the most well-known being the Measure-Correlate-Predict approach [10]. The AR and ARMA models have also been successfully used for the simulation of a wide range of meteorological parameters, including wind speed (e.g. [11]), with some variations, such as polynomial AR models [12] or ARMA-GARCH models [13], being developed to further improve model performance.

As mentioned, this paper uses a simple linear regression model, which can be easily constructed and tested with widely available non-specialized software and permits full manipulation of explanatory variables and input data. Accordingly, all presented results are obtained using [14] for the correlational analysis, and [15] for model development.

III. METHODOLOGY

A. Regions Selected for Analysis

Two regions from Scotland, UK, are selected for the analysis: the Central Belt lowlands (marked as “CB”) and the NE of Scotland (marked as “NES”), roughly located between the Firth of Forth and the Moray Firth, Fig. 1. In each region, 13 suitable Met Office weather stations were identified, containing records of hourly average wind speed and wind direction from 2006 to 2016, [16].

Fig. 1. Two regions from Scotland selected for the analysis [16] (numbers indicate station IDs in the Met Office MIDAS database).

Two selected regions exhibit different regional-scale wind regimes. The NES region is characterized by the interaction of multiple weather fronts and features a range of different topographical characteristics, while the CB region shows a strong presence of mass airflow zones, with a well-defined prevailing wind direction. These characteristics offer a basis for comparison of different regional wind climate conditions.

B. Analytical and Computational Framework

The raw WSTS data from weather stations in two selected regions were averaged (“smoothed”) over 1-day, 1-week and 1-month periods, in order to investigate whether the removal of periodic trends improves the correlational analysis. To determine whether correlational behaviour changes according to wind speed, the raw and smoothed WSTS’ were separated (i.e. ‘banded’) into six speed ranges, mimicking a typical WTG power curve, Fig. 2.

![Fig. 2. Typical WTG power curve and designated wind speed bands [17].](image)

A series of metrics were used to analyse the wind climates of selected stations/sites, strength of inter-site correlational relationships, prediction accuracy of the developed model and performance of WTGs from the database. The strength of the correlation between all pairs of sites in considered region is assessed using Pearson’s coefficient, \( R_{xy} \), for synchronized measurements and using cross-correlation coefficient for time-lagged measurements, \( R_{xy}(m) \), representing strength of the maximum cross-correlation coefficients, calculated using raw, smoothed and banded WSTS as:

\[
R_{xy} = \frac{(N) \sum_{n=1}^{N} (x_n - \mu_x)(y_n - \mu_y)}{\sigma_x \sigma_y} \quad (1)
\]

\[
R_{xy}(m) = E\{x_{n+m}y_n^*\} = E\{x_n^*y_{n-m}\} \quad (2)
\]

where: \( x \) and \( y \) are two WSTS, \( N \) is the number of data points at each site, \( \mu \) and \( \sigma \) are the mean wind speed and standard deviation of each time series, \( m \) is the lag and \( E \) is the expectation.

Both synchronized and lagged WSTS are used to check whether the wind climates of the sites within the same region influence each other in:

a) synchronous manner, i.e. if the wind climate at a selected location at a particular hour can be correlated with the contributions of wind climates at surrounding locations registered during that same hour, or
b) asynchronous manner, i.e. the wind energy will be transported from one site to another, resulting in a delay in two sites experiencing the same wind climate, depending on the inter-site distance and actual wind speed/direction. In that case, the wind climate at a selected location during a particular hour is better correlated with the contributions of wind climates at surrounding locations registered during earlier or later hours, expressed as a lag (m) between related WSTS.

The above analysis was performed using two major explanatory variables, referred to as ‘temporal separation’ and ‘directional alignment’. Temporal separation defines the expected lag/delay between one site experiencing a particular wind climate and the neighbouring site experiencing the same wind climate as the shortest distance between two sites divided by the wind speed during a particular hour:

\[ T_{sep} = \frac{d \times 1000}{U \times 3600} \]  

where: \( T_{sep} \) is temporal separation between two sites (in hr), \( d \) is distance between two sites (in km), and \( U \) is wind speed at the reference site during a particular hour (in m/s).

The surface distance between any two sites was computed from station coordinates, using the haversine formula [18] and assuming it is equal to distance at their anemometer heights.

Directional alignment defines how much the wind climate at a contributing station is expected to impact the affected station. This impact depends on how well-aligned the direction of wind at the contributing station is with the shortest distance between the contributing and affected stations. It is defined as the difference between inter-site angular distance and the prevailing wind direction of the contributing station:

\[ D_{align} = \alpha - WD \]  

where: \( D_{align} \) is degree of directional alignment (in degrees), \( \alpha \) is angular distance between two sites (in degrees) and \( WD \) is the wind direction at the first site, during a particular hour (degrees). The inter-site angular distance (\( \alpha \)) was defined as the angle formed by the line of shortest distance and the east-west distance between them. The value of \( \alpha \) was calculated for each pair of sites as:

\[ \alpha = \arccos \left( \frac{x_{EW}}{x} \right) \times \frac{180}{\pi} \]  

where: \( x \) is the shortest distance between two sites and \( x_{EW} \) is projection of the shortest distance on the east-west parallel (longitudinal distance).

An extensive regression analysis was performed to determine whether the wind speed, temporal separation and directional alignment time series of the contributing sites were good predictors for the WSTS of the affected site. The analysis was run for both synchronized and lagged raw, detrended and banded WSTS, to determine the effect of lagging on the accuracy of prediction by contributing sites. The regression equations are shown for synchronous WSTS’ (6) and for lagged WSTS’ (7).

\[ WSTS_{affected} = \sum_{k=1}^{n} a_k \times WSTS_k + \sum_{k=1}^{n} c_k \times D_{align(k)} \]  

\[ WSTS_{affected} = \sum_{k=1}^{n} d_k \times WSTS_k + \sum_{k=1}^{n} e_k \times T_{sep(k)} + \sum_{k=1}^{n} f_k \times D_{align(k)} \]  

where: \( WSTS_{affected} \) is WSTS of the affected site (response variable), \( WSTS_k \) is WSTS of the contributing site of index \( k \) (predictor), \( D_{align(k)} \) is directional misalignment of the contributing site of index \( k \) (predictor), \( T_{sep(k)} \) is temporal separation of the contributing site of index \( k \) (predictor), \( n \) is total number of contributing sites used as predicting variables, \( a_k – f_k \) are model regression coefficients.

C. Development of WTG Database

To provide a comprehensive list of existing small-scale WTGs, an existing database, produced in 2012 based on [19], was reviewed and updated. This database originally contained 194 different WTGs, with power ratings of up to 100 kW, available from 68 different manufacturers. The revised database found 217 WTGs, available from 47 UK-based and international manufacturers in 2016. An initial comparison shows that the WTG market is volatile, with 69% of manufacturers active in 2012 going out of business, or being acquired by larger companies by 2016.

The 2016 database provides the rated power (W), cut-in wind speed (m/s), rated wind speed (m/s), swept area (m²), turbine hub height (m) and purchase price (where available, £), for each WTG. A visual inspection of the manufacturer power curves also allowed the calculation of power output (kW) and normalized power densities (power output divided by the WTG swept area, in kW/m²). Because not all WTG manufacturers provided reliable power curves, power densities were computed for 200 of the machines.

Fig. 3 compares the original (2012) and updated (2016) WTG databases, illustrating a reduced share of WTGs with rated power in the 0.5-5 kW and 5-10 kW ranges and an increased share of WTGs with rated power in the 10-50 kW and 50-100 kW range. It was found that the average normalized power (per swept area) decreased by around 6% between 2012 and 2016, indicating that there have been no improvements in the design efficiency of WTGs, i.e. that the market still lacks compact and highly efficient machines.
The final part of the analysis selects the optimal WTGs for a particular site with specific wind energy resources (i.e. simulated or actual WSTS). The WTGs in the database were classified into four generic power curve categories developed in [19], with category WTG-1 further divided in two subcategories A and B, representing the best-performing, and category WTG-4 the worst-performing machines, as used e.g. in [7] and illustrated in Fig. 4.

For the selected (unmonitored) site, the annual electricity production (AEP) of all candidate WTGs from the database was estimated by, first, allocating corresponding generic WTG curve (Fig. 4) and, then, using simulated WSTS data, best-fit Rayleigh and Weibull PDFs approximations and average annual wind speed (AAWS) to calculate AEP values. The results of this analysis are shown in Section IV.

IV. RESULTS

The presented model was tested by applying the developed linear regression equation to the contributing sites in each sample area, assuming that affected site is unmonitored. Afterwards, the model is validated by comparing calculated AAWS, WSTS and key statistical metrics: average wind speed, standard deviation, variance, Weibull scale and shape parameters at affected site with the corresponding actually recorded WSTS, AAWS and all related metrics, as well as by calculating the strength of correlation between the calculated and recorded data series.

A. CB Region

1) Comparison of modelled and measured AAWS': The model was successful in simulating the WSTS’ of all sites in the CB region, with the average annual wind speed (AAWS) of 4 out of 5 sites being simulated with errors of less than 9%, and less than 2% at 3 out of 5 sites, Table I. There was also a high cross-correlation, $R_{x,y}$ (2), between the simulated and actual WSTS (over 0.94 in all cases).

2) Comparison of modelled and measured WSTS': Figs. 5 and 6 compare simulated and measured time series for a period of 144 hours at two sites in CB region, showing an overall very good match. Regarding all sites in CB area, the model performed better when using synchronous, rather than lagged WSTS.

3) Comparison of statistical metrics: In CB region, the model underestimates the variance of simulated WSTS, also reflected in an overestimation of Weibull shape parameters, a known shortcoming of linear regression models. The percentage errors between the statistical metrics for sites in the CB region are shown in Table II.
### TABLE II. ERRORS IN THE ESTIMATION OF KEY METRICS (CB AREA)

<table>
<thead>
<tr>
<th>Site ID</th>
<th>% difference between simulated and actual metric values</th>
<th>Std. Dev.</th>
<th>Variance</th>
<th>Weibull C</th>
<th>Weibull k</th>
</tr>
</thead>
<tbody>
<tr>
<td>246</td>
<td>-20.33 -36.53</td>
<td>-1.78</td>
<td>24.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>971</td>
<td>-10.74 -20.32</td>
<td>-8.50</td>
<td>22.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>987</td>
<td>-19.18 -34.67</td>
<td>-32.59</td>
<td>36.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1006</td>
<td>-22.75 -40.32</td>
<td>-1.61</td>
<td>26.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1007</td>
<td>-34.31 -56.84</td>
<td>-16.07</td>
<td>28.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B. NES Region**

1) **Comparison of modelled and measured AAWS:** The developed model and regression equations generated much higher errors in the NES region for AAWS, Table III, although there was relatively strong cross-correlation, $R_{x,y}$ (2), between the simulated and actual WSTS (over 0.85).

| Site ID | Comparison of AAWS’ $R_{x,y}$ (2) Measured AAWS (m/s) Simulated AAWS (m/s) Difference (%) |
|---------|---------------------------------------------------------------|---------------------------------|-----------------|-----------------|
| 117     | 0.85                                                          | 14.35                          | 2.85            | -80.13          |
| 145     | 0.86                                                          | 10.97                          | 3.17            | -71.10          |
| 161     | 0.90                                                          | 4.67                           | 2.26            | -51.61          |
| 177     | 0.91                                                          | 6.64                           | 2.18            | -67.16          |

2) **Comparison of modelled and measured WSTS’**:

Fig. 7 compares simulated and measured time series for a period of 144 hours at one site in NES region, showing an overall poor match. For all sites in NES area, the model performed better when using lagged, rather than synchronous WSTS data.

![Simulated and measured WSTS’ for site 145 (144 hours).](image1)

3) **Comparison of statistical metrics:** In the NES region, as expected from the WSTS performance, the model heavily underestimates all key statistical parameters, except Weibull shape parameter, which is overestimated, Table IV.

### TABLE IV. ERRORS IN THE ESTIMATION OF KEY METRICS (NES AREA)

<table>
<thead>
<tr>
<th>Site ID</th>
<th>% difference between simulated and actual metric values</th>
<th>Std. Dev.</th>
<th>Variance</th>
<th>Weibull C</th>
<th>Weibull k</th>
</tr>
</thead>
<tbody>
<tr>
<td>117</td>
<td>-88.93 -98.78</td>
<td>-79.84</td>
<td>73.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>-87.46 -98.42</td>
<td>-84.60</td>
<td>96.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>161</td>
<td>-81.19 -96.46</td>
<td>-54.04</td>
<td>122.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>177</td>
<td>-64.14 -87.12</td>
<td>-49.79</td>
<td>38.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. **WTG Selection: Annual Energy Production**

This section gives the results for selection of optimal WTGs described in Section III.C. As the WSTS data had the highest temporal resolution, they were used as the reference estimation methods, against which the estimates produced by the AAWS and PDFs were compared. There are significant differences in the estimated AEP values by the four different metrics at all sites in both considered regions (only the Weibull distribution at sites 145 and 987 estimated the AEP within a 10% error of the WSTS estimate). Fig. 8a shows mean absolute percentage errors (MAPE) for sites in CB region, with overall average error for AAWS of -52.5%, Rayleigh of 46.2% and Weibull of 52%. Fig. 8b shows the same MAPE values for sites in NES region, where overall average error for all sites are -30.7% for AAWS, 54% for Rayleigh and 59.5% for Weibull.

![Estimated AEP by four different metrics in two analysed regions.](image2)

The analysis of WTG payback periods showed that 57% of WTGs in the database are viable (i.e. they will pay back for their installation in under 20 years) in the CB region, while this was the case for only 29% of WTGs in the NES region. The results for individual sites are shown in Fig. 9. At all sites, the turbines generating the highest estimated AEP were from the highest WTG generic categories (WTG-1 or WTG-2).
It should be noted that the model presented in this paper assumes single-WTG installation in an open area and does not take into account micro-siting issues (e.g. WTG height, surrounding terrain, presence of turbulences and wake/array losses), which typically have strong effect on WTG outputs, especially for small-scale WTGs installed in urban built environment.

V. CONCLUSIONS

The accurate estimation of wind energy resources at unmonitored sites from measurements at a limited number of monitored sites proved to be a difficult task. This paper presents a simple linear regression model, based on the spatio-temporal correlational analysis, which under certain conditions can produce reasonably accurate estimates. The methodology is illustrated on two relatively large regions in Scotland, where publicly available wind speed and wind direction measurements are available from the UK Met Office stations. The presented approach requires only the coordinates of the end-user’s WTG location, is simple to implement and requires low computational resources. It further incorporates a method to identify the best-fit WTGs for the considered location with specific wind energy resources. For that purpose, a database of 200 wind turbines was constructed, reflecting the current market for small-scale wind turbines in the UK.

The analysis showed that inter-site distance and, to some extent, directional alignment impact the strength of correlation between measured sites. The classification of station WSTS into wind speed bands showed evidence of inter-site wind energy transport, which would cause a temporal delay in the correlation of station wind climates. To test whether stations in both study areas are subject to temporal delay, or, rather, behave synchronously, the regression models were fitted using both synchronous and lagged time series as predictors. The results showed that the correlation coefficients were higher when using lagged WSTS in the NES region and, conversely, higher when using synchronous WSTS in the CB region. Consequently, better predictions are obtained in CB region, which exhibits a more uniform regional-scale wind regime.

The developed model is effectively an important tool for empowering end-users and potential WTG developers to make an independent and more confident decision on small-scale WTG projects, helping to increase further implementation of renewable-based energy technologies and, in that way, provide wider environmental, social and economic benefits. However, further research is needed to determine whether the presented “end-user led approach” can be applied to other areas and regions in Scotland, UK and elsewhere.

REFERENCES