UNIVERSITY OF LEEDS

This is a repository copy of *Long-term wind resource assessment for small and medium-scale turbines using operational forecast data and measure-correlate-predict*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/84631/

Version: Accepted Version

Article:
Weekes, SM, Tomlin, AS, Vosper, SB et al. (3 more authors) (2015) Long-term wind resource assessment for small and medium-scale turbines using operational forecast data and measure-correlate-predict. Renewable Energy, 81. 760 - 769. ISSN 0960-1481

https://doi.org/10.1016/j.renene.2015.03.066

© 2015, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International
http://creativecommons.org/licenses/by-nc-nd/4.0/

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Long-term wind resource assessment for small and medium-scale turbines using operational forecast data and measure-correlate-predict

S. M. Weekes*[1,2], A. S. Tomlin[2], S. B. Vosper[3], A. K. Skea[3], M. L. Gallani[3] and J. J. Standen[3]

[1]Doctoral Training Centre in Low Carbon Technologies, [2]Energy Research Institute, School of Process, Environmental and Materials Engineering, University of Leeds, Leeds, LS2 9JT, UK

[3]Met Office, Fitzroy Road, Exeter, EX1 3PB, UK

*E-mail: a.s.tomlin@leeds.ac.uk

Highlights

- Wind resource predicted at 37 sites using operational forecast data and MCP
- Forecast reference data highly competitive with nearby meteorological observations
- Systematic improvement when using forecast data at coastal sites

Abstract

Output from a state-of-the-art, 4 km resolution, operational forecast model (UK4) was investigated as a source of long-term historical reference data for wind resource assessment. The data were used to implement measure-correlate-predict (MCP) approaches at 37 sites throughout the United Kingdom (UK). The monthly and hourly linear correlation between the UK4-predicted and observed wind speeds indicates that UK4 is capable of representing the wind climate better than the nearby meteorological stations considered. Linear MCP algorithms were implemented at the same sites using reference data from UK4 and nearby meteorological stations to predict the long-term (10-year) wind resource. To obtain robust error statistics, MCP algorithms were applied using onsite measurement periods of 1-12 months initiated at 120 different starting months throughout an 11 year data record. Using linear regression MCP over 12 months, the average percentage errors in the long-term predicted mean wind speed and power density were 3.0% and 7.6% respectively, using UK4, and 2.8% and 7.9% respectively, using nearby meteorological stations. The results indicate that UK4 is highly competitive with nearby meteorological observations as an MCP reference data source. UK4 was also shown to systematically improve MCP predictions at coastal sites due to better representation of local diurnal effects.

Keywords: measure-correlate-predict, wind resource assessment, operational forecast data, numerical weather prediction

1 Introduction

Small and medium-scale wind turbines, typically defined as < 500 kW rated power [1, 2], have exciting prospects as we move towards a low carbon energy future. As a source of renewable, low carbon energy, such turbines have the potential to contribute to both carbon savings and improved diversity of supply. The global market has seen rapid growth in the last decade with an increase in capacity of 18% reported for 2012, compared to the previous year [1]. Within the UK, the small and medium-scale wind energy industry is predicted to contribute £241 million Gross Value Added (a measure of the contribution of an individual market sector) to the UK economy in 2014. While changes to the UK Feed-in Tariff have slowed growth since 2013, the industry has the potential to contribute up to £864 million to the UK economy in 2023, given appropriate political support [2]. However, in order for small and medium-scale wind energy to continue to flourish, methods for rapid, accurate and low-cost wind resource assessment are required [1].

In the large-scale wind energy industry, estimates of the long-term wind resource are generally achieved using the measure-correlate-predict (MCP) method [3]. In a typical MCP approach, short-term wind data are obtained at the location and height of the potential wind turbine site (target site) over a training period and these data are correlated to a nearby reference site where a long-term historical data record is available. The correlated data are then used to make a long-term prediction at the target site, under the assumption that the historical wind resource is an adequate predictor of the future resource.

The length of the short-term measurement period and long-term prediction period will vary depending on the size of the project and the rigour of the site assessment procedure. For large-scale wind projects, onsite measurement periods of 1-3 years are generally required along with long-term prediction over several decades [3, 4]. For small and medium-scale installations, the lower investment costs may justify shorter measurement and prediction periods, reflecting the reduced financial risk [5]. In many cases, long-term reference data are sourced from established monitoring stations operated, for example, by national meteorological institutions or airports. However, in recent years there has been increasing interest in the use of output from numerical weather prediction (NWP) models, as well as derived atmospheric reanalysis data sets, as a source of long-term reference data for MCP [6]. Both NWP and reanalysis involve the assimilation of large amounts of data including observations from satellites, weather balloons, aircraft, ships, buoys and surface meteorological stations. These data are used to initialise numerical models which produce a time-evolving, three-dimensional grid of modelled atmospheric variables [6]. Reanalysis data are so termed because they represent a second analysis, using a consistent assimilation and analysis model, as well as incorporation of observations not available to real-time operational forecasts [7].

Such data are attractive in that (i) they are available globally and (ii) compared to nearby surface wind observations, they may be less affected by changes in land-use and local obstructions, and they may also cover a longer historical period.

While these data are increasingly being used in long-term wind resource assessment [8], there are relatively few rigorous studies considering the suitability of NWP and derived datasets in MCP applications. Brower [6] carried out a detailed study using NCEP/NCAR (National Centers for Environmental Prediction / National Center for Atmospheric Research) [7] reanalysis as an MCP reference data source for sites located in the United States. It was found that in some cases, reanalysis data can be subject to spurious trends and internal inconsistencies, particularly when considering data lengths of greater than 10 years. Liléo and Petrick [9] compared NCEP/NCAR with MERRA [10] (Modern Era Retrospective Analysis for Research) as well as a more recent NCEP release NCEP/CFSR (Climate Forecast System Reanalysis) [11], with improved spatial and temporal resolution, using observations at 24 meteorological stations in Sweden. Their results indicated that MERRA performed better in MCP analysis, due to the data's higher spatial and temporal resolution, and was less prone to the spurious trends observed in other reanalysis data. Similar studies have also been reported by Pinto *et al.* [12] as well as Jimenez *et al.* [13]. The emerging picture from these preliminary studies is that data sets with high spatial and temporal resolution are required for the successful implementation of MCP and that long-term inconsistencies may affect predictions on time-scales greater than 10 years. Since reanalysis data generally have low spatial resolution (tens to hundreds of kilometres), and variable temporal resolution, this presents a challenge to their use in MCP.

One possible solution is the use of mesoscale models with improved spatial and temporal resolution driven either by reanalysis or some alternative data containing consistent observations of atmospheric variables [14]. Operational forecast data from high resolution NWP models may be considered as a natural choice in this regard. Since forecast data can be obtained at the location of the target sites (subject to the model resolution), the data may offer improved representation of localised climates compared to nearby surface measurements located tens of kilometres away, or low resolution reanalysis data. Currently, there is a lack of rigorous studies investigating the use of high resolution, operational forecast data in MCP. In this study, the Met Office Unified Model (UM) is investigated as an MCP reference source in the context of small and medium-scale wind installations.

The UM [15] is a state-of-the-art operational weather and climate forecast system, used for both global and regional prediction. It is currently operated with horizontal grid spacings of approximately 25 km globally and 1.5 km (previously 4 km) within the UK. As a terrain-following, mesoscale model, the UM is capable of producing, local, site-specific forecasts through progressively higher resolution models whose boundary conditions are provided by the global model. Wilson and Standen [16] recently demonstrated that due to the higher spatial and temporal resolution of the 4 km model (UK4), the data are capable of outperforming reanalysis in wind resource assessments using downscaling methods.

In the current study, UK4 is investigated as a source of long-term reference data for MCP and its performance is compared with alternative reference data obtained from nearby meteorological stations. Linear MCP approaches are used to predict the long-term (10 year) wind resource at 22 target sites (later extended to 37) located in four different terrain types. A range of error metrics are used to compare the accuracy of the predictions using the two sources of reference data. The study is particularly relevant to small and medium-scale wind resource assessment due to the range of heights considered (10 - 22.5 m above ground level) and the length of the long-term predictions (10 years). The main objectives of the study are (i) to investigate the utility of using UK4 data as a long-term reference data source for MCP, (ii) to determine the most appropriate forecast height to use in this context, and (iii) to investigate factors that may impact the performance of UK4 within the MCP approach, including local terrain and the use of hindcast data.

2 Methodology

2.1 MCP algorithms

A large range of MCP approaches have been investigated in a research context. These include two-dimensional, vector and non-linear regression techniques [17-20], matrix approaches [21, 22] and more recently, artificial neural networks [23-25] and joint probability distributions [26-28]. A recent review by Carta *et al.* [3] considered over 150 studies demonstrating the wide range of available techniques. Despite the large number of alternatives, linear approaches [29, 30] are currently the most widely used in the wind industry, presumably due to their simplicity and effectiveness [3]. The current study is concerned with investigating the utility of operational forecast data as an MCP reference data source, rather than investigating specific MCP algorithms. Hence, two established linear MCP approaches are used in this work, as described below.

2.1.1 Linear regression

In a linear regression approach, the target and reference site wind speeds may be related by the linear expression:

$$\hat{u}_{tar} = \alpha + \beta u_{ref} + \varepsilon$$

Equation 1

where $\hat{u}_{tar}$ is the predicted wind speed at the target site, u_{ref} is the observed wind speed at the reference site, α and β are the regression coefficients obtained using a least squares fit to the training data and ε is an error term which represents the residual scatter.

Previous studies [5, 28] have indicated that ε can be modelled using a zero mean Gaussian distribution of the form:

$$\varepsilon = N(0, \sigma_{res}^2)$$

Equation 2

where σ_{res} is the sample standard deviation of the residuals about the predicted target site wind speeds $\hat{u}_{tar}$, as calculated from the N training observations using [31]:

$$\sigma_{res} = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}\left(u_{tar,i} - \hat{u}_{tar,i}\right)^2}$$

Equation 3

The approach described above is referred to as LR in the remainder of this study.

2.1.2 Variance ratio method

An alternative linear approach is the variance ratio method (VR). The approach was proposed by Rogers *et al.* [30] based on the observation that for simple linear regression, without consideration of the residual scatter, the target site variance is underestimated by a factor $1/r$ where r is the linear correlation coefficient. By forcing the predicted and observed target site wind speeds to have the same variance, Rogers *et al.* derived the expression:

$$\hat{u}_{tar} = \left[\bar{u}_{tar} - \left[\frac{\sigma_{tar}}{\sigma_{ref}}\right]\bar{u}_{ref}\right] + \left[\frac{\sigma_{tar}}{\sigma_{ref}}\right]u_{ref}$$

Equation 4

where σ_{tar} and σ_{ref} represent the standard deviation about the mean wind speeds at the target ($\bar{u}_{tar}$) and reference ($\bar{u}_{ref}$) sites respectively, as calculated from the short-term training data. Several studies have confirmed the utility of the VR approach, and along with linear regression, it is widely used for assessing the success of new techniques [25, 30, 32].

2.2 Meteorological observations

To implement the MCP approaches and enable comparison between the predicted and observed wind resource, hourly averages of wind speed and direction (resolution 10° and 0.51 ms^{-1} = 1 knot), were obtained from the Met Office anemometer network [33] for an 11 year period covering August 2001 to July 2012. A total of 22 UK target sites were chosen to represent the four terrain types of urban, suburban, rural and coastal, as determined from satellite images. These terrains were chosen to be representative of the range that may be encountered by small and medium-scale wind energy developers. An additional 15 nearby sites were chosen as long-term reference sites for implementation of the MCP algorithms. Reference sites were chosen on the basis of data availability and proximity to the target sites. Wherever possible, reference sites were located in open terrain, and in the case of pairings with coastal target sites, coastal reference sites were given preference. In some cases the same reference site was used for more than one target site, hence, the total number of unique, individual sites, reference plus target, was 37. Anemometers were assumed to be located at 10 m above ground level, as is the standard Met Office practice, unless otherwise indicated. Details of all reference and target sites are shown in Table 1 and their locations are shown in Figure 1. The same combination of reference and target sites were also the subject of previous MCP studies [5, 28].

Reference sites (Rf)			Target sites			
Site	OS grid	Elev (m)	Site	OS grid	Elev (m)	*d* (km)
Rf1	SD6614	440	U1*	SJ8396	33	25
Rf2	SU5501	9	U2**	SU4210	26	16
Rf3	NO4620	10	SU1	NJ8712	65	101
Rf4	TF0049	63	SU2	SK5045	117	49
Rf5	SU3039	90	SU3	SU8554	65	58
Rf5	SU3039	90	SU4	SU1344	132	17
Rf5	SU3039	90	SU5	SU1740	126	13
Rf1	SD6614	440	SU6	SD8812	110	22
Rf6	SP2186	96	SU7	SP3180	119	12
Rf7	NJ2169	7	C1	NK1345	15	96
Rf3	NO4620	10	C2	NU2514	23	133
Rf8	TA0243	7	C3	TA1967	15	30
Rf9	NR6622	10	C4	NM8834	3	113
Rf10	SM8905	44	C5	SN2452	133	59
Rf11	SX4952	50	C6	SX9456	58	46
Rf12	SD3131	10	C7	SD3000	9	31
Rf13	NJ0662	5	R1	NH8914	228	51
Rf14	SE4961	14	R2	SE5238	8	24
Rf4	TF0049	63	R3	SK5026	43	55
Rf6	SP2186	96	R4	SO9749	35	44
Rf5	SU3039	90	R5	SU7349	118	45
Rf15	NT2302	236	R6	NS8264	277	74

Table 1: Meteorological observation sites used in this study. Terrain types are identified by the site names Urban, SubUrban, Coastal and Rural. Reference sites are labelled Rf. The reference-target site separation (d) and elevation above sea level (Elev) are also shown.

Starred sites denote building mounted anemometers at heights *H = 20.6 m, **H = 22.5 m above ground level.

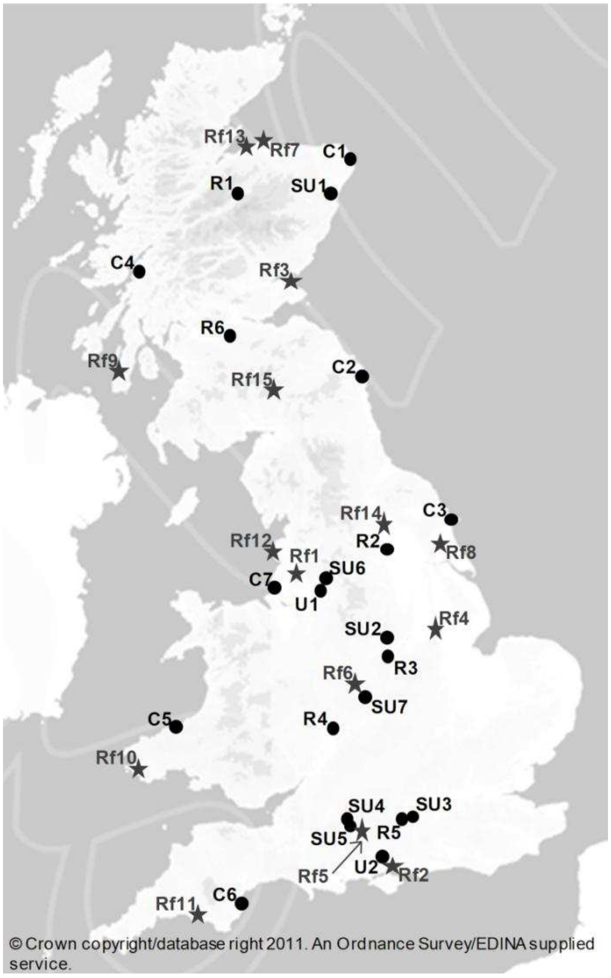

Figure 1: Approximate geographical locations of the meteorological observation sites used in this study. Terrain types are identified by the site names Urban, SubUrban, Coastal and Rural. Reference sites are labelled Rf, [28].

2.3 UK4 forecast model output

In addition to the meteorological observations detailed above, long-term wind data from the Met Office UK4 model were obtained as a source of MCP reference data. In this study, the observed wind data at the reference sites (as detailed in Section 2.2) are referred to as Rf, while the UK4 data used as an alternative reference source are referred to as UK4. The UK4 data covered the same 11 year period outlined above and were sourced directly from the Met Office as part of a collaborative project. The data consisted of a time series of hourly wind speed and direction predictions at eight heights, (10, 20, 35, 50, 100, 150, 200 and 500 m above ground level).

The data were obtained at the location of each target site using bilinear interpolation in the horizontal plane and interpolated in the vertical assuming a logarithmic profile between model levels. The UK4 model has 70 vertical levels with a model top of 40 km; there are 11 model levels below 500m (2.50, 13.3, 33.3, 60.0, 93.3, 133, 180, 233, 293, 360 and 433 m). Since operational UK4 forecasts only started in 2007, additional hindcasts were used to cover the period 2001-2006. These involved running the UK4 forecast model over a historical time period with the following distinctions:

(i) For the hindcasts, a series of nested configurations were run. Firstly, a global N216 resolution (approximately 60 km horizontal resolution at mid-latitudes) UM configuration with 50 vertical levels was run where the initial conditions were obtained from ECMWF (European centre for medium-range weather forecasts) ERA-i reanalysis [34]. The global hindcast provided the initial and boundary conditions for a 38-level 12 km resolution western European UM configuration, which in turn provided the initial and boundary conditions for the 70-level UK4 UM configuration. For the operational forecast period (2007 onwards), the UK4 forecasts were initialized with the UK4 analysis (generated by variational data assimilation) and boundary condition data were obtained from the operational 12 km resolution North Atlantic and European (NAE) UM forecast up until 17 January 2012 and from the N512 resolution (approximately 25 km horizontal resolution at mid-latitudes) global UM forecast after this.

(ii) For the hindcast period, the time between reinitialisation was 48 hours compared to 24 hours for the post 2007 period.

The effect of these differences is considered in more detail in the following sections.

2.4 Independence of UK4 and observed wind data

The success of the Rf and UK4 reference data were quantified by comparing the long-term MCP predictions with the observed wind resource at target sites that form part of the Met Office anemometer network. Along with data from a large number of other atmospheric observations, data from this network were assimilated for use in initialising the background field of the operational UK4. Hence, an important question is whether the UK4 data can be considered largely independent of the observations at the target sites. Here ‘independent’ is used to mean that the UK4 data are not significantly influenced by the target site observations since any such influence will tend to cause the success of the MCP approaches to be overstated. This issue is only relevant to UK4 data post-2007 since the ERA-i used for initialisation of the pre-2007 hindcast data did not include observations from this network [35]. There are several factors that must be clarified in addressing this issue.

Firstly, UK4 assimilates observations of a range of atmospheric variables from a large number of sources. These observations are combined in the background model field that is optimised to be meteorologically consistent. Hence, while surface wind speed observations may influence the background field, the UK4 data are not forced to fit them. Secondly, observations are only used to initialise UK4 at the start of a run. As the forecast evolves it knows nothing of the actual time evolving wind climate at the locations used for initialisation. Thirdly, the UK4 data used in the current study are extracted starting at T+2 hours, where T+0 is the time at which the forecast is initialised. Hence, there is a gap of approximately two hours between any assimilated observations and the first forecast point. Several studies [35-37] have shown that assimilation of the 10 m surface wind speeds has little influence on the forecast output beyond the analysis time and that other variables related to temperature and humidity are of greater importance. However, due to the importance of this issue in objectively assessing the performance of UK4 as a reference data source, further tests were performed.

To investigate the effect of assimilated wind data, the linear correlation coefficients between hourly wind speeds (r_u) for the forecast period (Jan 2007 – July 2012), were compared for each site using the full data set, and after excluding the target site and reference (UK4) data between T+2 and T+5 hours inclusive (T+0 and T+1 are not included in the UK4 dataset). This represents the first 6 hours of the forecast, after which any influence from the assimilated surface wind speeds is expected to be lost. The average value of r_u across the 22 sites differed by less than 1% using the two data sets. In addition, the errors in the MCP wind resource predictions using the two data sets were found to be almost identical indicating that for the current purposes, the UK4 and observed wind data can be considered independent.

2.5 Error metrics

The predicted wind resource parameters of mean wind speed ($\bar{u}$), mean Betz power density ($\bar{p}_d$), standard deviation of hourly wind speeds (σ) and Weibull shape factor (k) across all 22 target sites were compared with onsite measurements using a sliding window technique in order to obtain robust error statistics. The technique allows the use of multiple training and test periods shifted throughout the entire 11 year data record in order to account for intra- and inter-annual variability. Training lengths of between 1 and 12 months were used, along with test periods of 10 years, such that the training and test periods never overlapped. For each training length, the sliding window was used to obtain 120 test predictions corresponding to different starting months throughout the 11 year data record. Full details of the technique can be found in reference [5].

Quantitative comparisons with onsite measurements were made using the metrics of mean absolute percentage error (%Error), mean absolute error (MAE), mean bias error (MBE) and standard deviation in the MBE (σ_{MBE}). These are defined below for the predicted mean wind speed, equivalent error metrics may be defined for the remaining parameters of interest.

$$\%\text{Error} = 100 \sum_j \frac{\left|\bar{u}_{obs,j} - \bar{u}_{pred,j}\right|}{\bar{u}_{obs,j}} / n$$

Equation 5

$$\text{MAE} = \sum_j \left|\bar{u}_{obs,j} - \bar{u}_{pred,j}\right| / n$$

Equation 6

$$\text{MBE} = \sum_j (\bar{u}_{pred,j} - \bar{u}_{obs,j}) / n$$

Equation 7

where j represents the j^{th} site, $\bar{u}_{obs}$ and $\bar{u}_{pred}$ are the long-term observed and predicted mean wind speeds respectively and n is the total number of target sites.

3 Results and Discussion

3.1 Representativeness of the reference data

To be suitable for making long-term target site predictions, the meteorological trends at the reference site must be consistent with those at the target site. The linear correlation between the target site and reference wind speeds based on monthly wind indices and hourly wind speeds were used to investigate the degree of similarity between the target site observations and the Rf and UK4 reference data. The wind index (WI) describes the short-term deviations from a long-term mean. Here, the monthly WI was used to compare the monthly mean wind speed to the11 year mean defined by:

$$WI_m = \frac{\bar{u}_m}{\bar{u}_{11yr}}$$

Equation 8

where WI_m and $\bar{u}_m$ are the wind index and mean wind speed for month m respectively and $\bar{u}_{11yr}$ represents the mean wind speed over 11 years.

As an example, Figure 2 compares the monthly WI for a typical target site (SU3) and the UK4 reference data (height 50 m) between August 2001 and July 2012. For this site, the UK4 data successfully predicts the monthly variability in the target site wind speed. The degree of agreement can be quantified using the linear correlation coefficient between the observed and predicted WI values (r_{WI}) [38]. This process can be repeated using Rf reference data in place of UK4 to compare the relative success of the two data sources.

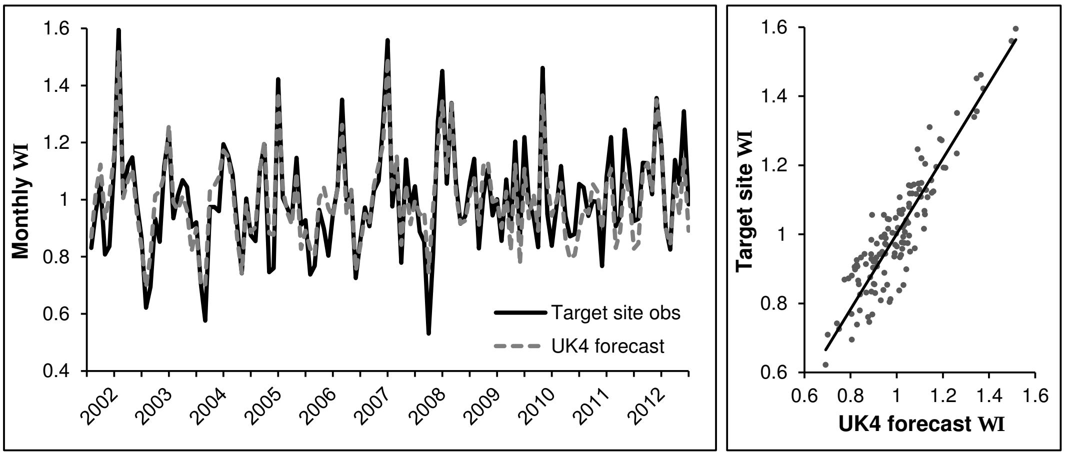

Figure 2: Left - Monthly wind indices for target site observations (solid line) and UK4 data (dotted line) at site SU3. Right – The equivalent data as a scatter plot, the line indicates a linear fit to the data.

The r_{WI} metric is useful in that it reveals differences in the long-term trends between the reference and target data and because it imposes the restriction of linearity only on these long-term trends, rather than the instantaneous wind speeds. However, the linear correlation between hourly wind speeds (r_u) is also of interest in investigating correlations on hourly timescales. Table 2 compares both the r_{WI} and r_u metrics across all target sites using the UK4 and Rf reference data.

	Wind Index $\mathbf{r}_{\mathrm{WI}}$		Hourly wind speed $\mathbf{r}_{\mathrm{u}}$	
Site	**UK4**	**Rf**	**UK4**	**Rf**
U1	0.90	0.81	0.81	0.79
U2	0.88	0.89	0.80	0.87
SU1	0.87	0.78	0.79	0.55
SU2	0.92	0.89	0.81	0.82
SU3	0.91	0.90	0.80	0.85
SU4	0.86	0.94	0.77	0.88
SU5	0.96	0.95	0.82	0.92
SU6	0.75	0.73	0.75	0.73
SU7	0.84	0.89	0.76	0.81
C1	0.92	0.64	0.76	0.51
C2	0.95	0.79	0.78	0.66
C3	0.87	0.71	0.74	0.68
C4	0.89	0.92	0.68	0.70
C5	0.98	0.90	0.85	0.79
C6	0.94	0.82	0.76	0.67
C7	0.87	0.94	0.82	0.88
R1	0.76	0.67	0.68	0.53
R2	0.94	0.94	0.81	0.88
R3	0.92	0.89	0.80	0.79
R4	0.94	0.94	0.79	0.85
R5	0.97	0.94	0.85	0.86
R6	0.96	0.91	0.85	0.73
Average	**0.90**	**0.85**	**0.79**	**0.76**

Table 2: Linear correlation coefficients r_{WI} and r_u for UK4 (50 m height) and Rf observations. The shaded cells represent the reference data source with the highest correlation to the target sites.

For r_{WI} there is a clear preference for the UK4 data with 15 out of 22 sites achieving higher r_{WI} values compared to the Rf observations and 2 sites showing no preference. This indicates that on average the UK4 data better represents the long-term climate for these sites. For r_u, the results are mixed with half the sites achieving higher r_u values when using the UK4 data, although the UK4 data achieves a higher overall average. The two metrics r_{WI} and r_u are most consistent at coastal sites with both metrics showing a preference for UK4 reference data at 5 out of 7 sites. This indicates that for these sites the UK4 data exhibits a stronger linear correlation to the target sites on both monthly and hourly timescales. This is likely because the coastal Rf sites, all of which are located at least 30 km from the target sites, are less able to represent the complex seasonal and diurnal variability in the coastal zone [39].

The results presented in Table 2 regarding the relative preference for Rf or UK4 data, as judged by the r_{WI} and r_u metrics, should not be considered definitive since they will depend on the choice of Rf site, which in turn will depend on data availability. However, they demonstrate the advantages of using co-located UK4 data in cases where nearby reference observations are either not available or exhibit different climatology to the target site.

3.2 Effect of forecast height

Since NWP models produce predictions of wind speed and direction at multiple levels, it is necessary to determine the most appropriate forecast height for use in MCP. Although 20 out of the 22 target sites are located at 10 m above ground level, the 10 m forecast data do not necessarily best describe the local wind climate. To account for orography with scales smaller than can be resolved on the model grid, NWP models generally include a parameterisation which represents the effect of the turbulent form drag due to sub-grid hills and valleys. While this improves the accuracy of weather and climate predictions, such schemes are known to result in unrealistically low wind speeds close to the surface [40]. Hence, the most appropriate forecast level is one of sufficient height that effects of the form drag parameterisation do not dominate, while remaining within the boundary layer such that the forecast winds remain representative of the near-surface winds.

To investigate the effect of forecast height on the success of the MCP predictions, the LR MCP algorithm was applied to predict the 10 year wind resource at the 22 target sites using UK4 data from eight forecast heights as a reference data source. Figure 3 shows the %Error metrics averaged across all 22 target sites as a function of the forecast height and training period. The %Error using Rf reference data (termed 'baseline') is also included.

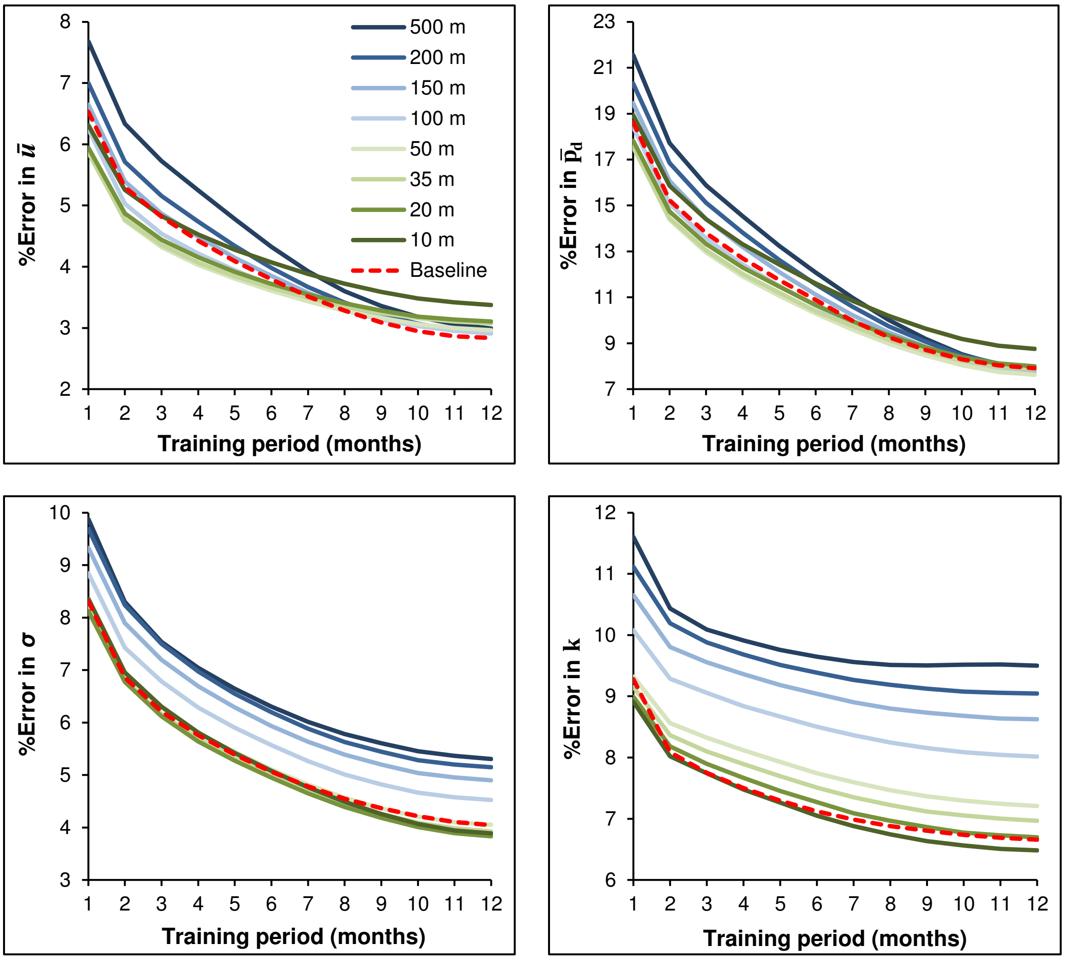

Figure 3: %Error metrics for $\bar{u}$, $\bar{p}_d$, σ and k using the LR MCP algorithm and UK4 data at different heights as a reference source. Lines show the mean values for each training period averaged across 22 site pairs and 120 starting months. The baseline using Rf data as a reference source is also shown.

The %Error in $\bar{u}$ and $\bar{p}_d$ generally decreases with decreasing UK4 height before increasing again below 35 m. For σ and k the trend is more straightforward with a decrease in error with decreasing height, indicating that the width of the wind speed distribution is better predicted when using UK4 data at 10 m. Similar trends were also observed for the VR approach.

The results are summarized in Figure 4 which shows the %Error metrics as a function of UK4 forecast height for training periods of 3 and 12 months. The values have been normalised by the baseline, hence, a value of less than one indicates improved performance of UK4 compared to Rf. For $\bar{u}$ and $\bar{p}_d$, the lowest %Error are observed for UK4 heights close to 50 m. In addition, UK4 performs slightly better than Rf in terms of the %Error in $\bar{u}$ and $\bar{p}_d$. These trends are stronger when using the shorter training period of three months, possibly due to seasonal dependent stability effects. Such effects could result in decoupling across relatively short lateral and vertical distances, thus enhancing the effect of forecast height as well as the errors resulting from using Rf sites located some distance from the target sites. The %Error in σ and k are generally higher when using UK4, except at the lowest forecast heights. Further investigation of the diurnal variability in the hourly wind speeds revealed that for UK4 data obtained at 50 m and below, the UK4 predicted and observed wind speeds exhibit very similar trends, with a maximum in the early afternoon. At 100 m, the diurnal trend in the UK4 data becomes significantly weaker, with complete reversal observed at 500 m. It seems likely that this is due to the height dependence of the net downward transfer of horizontal momentum during turbulent mixing within the boundary layer [41]. The reversal leads to a reduced correlation between data from higher levels in UK4 and observed surface winds, and is likely related to the increase in MCP error when using UK4 data above 50 m. For the MCP analysis detailed in the following sections, a fixed forecast height of 50 m is used.

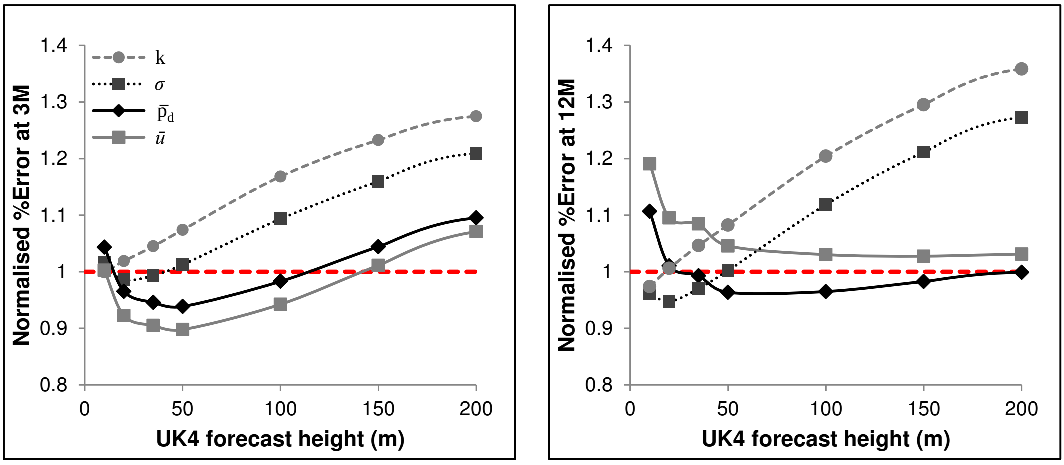

Figure 4: %Error metrics normalised by the baseline as a function of the UK4 forecast height for 3 month (left) and 12 month (right) training periods. Results are averaged across 22 site pairs and 120 starting months. Lines are a guide to the eye.

Table 3 shows the error metrics of %Error, MAE, MBE and σ_{MBE} for the MCP approaches of LR and VR using UK4 reference data as well as the Rf baseline. The errors across all metrics using LR are very similar to the baseline, demonstrating that UK4 reference data performs similarly to nearby meteorological observations. There is also an indication that the UK4 data may lead to slightly reduced errors when using the shorter training period of 3 months. The errors using VR exhibit similar trends to LR, albeit with slightly higher errors in $\bar{u}$ and $\bar{p}_d$ and slightly lower errors in σ and k, as has been observed previously [5].

3 M	Method	$\bar{u}$	$\bar{p}_d$	σ	k
%Error	Baseline	4.8	14	6.2	7.8
	LR	4.3	13	6.3	8.3
	VR	4.8	15	4.9	4.6
		$\bar{u}$ (ms^{-1})	$\bar{p}_d$ (Wm^{-2})	σ (ms^{-1})	k
MAE	Baseline	0.21	11	0.16	0.14
	LR	0.18	9.3	0.15	0.15
	VR	0.20	10	0.12	<0.1
MBE	Baseline	<0.1	-2.8	-0.10	0.13
	LR	<0.1	-3.1	-0.11	0.14
	VR	0.11	5.7	<0.1	<0.1
σ_{MBE}	Baseline	0.29	18	0.19	0.12
	LR	0.23	14	0.16	0.11
	VR	0.25	17	0.16	<0.1

12 M	Method	$\bar{u}$	$\bar{p}_d$	σ	k
%Error	Baseline	2.8	7.9	4.0	6.7
	LR	3.0	7.6	4.1	7.2
	VR	3.3	9.2	3.3	4.1
		$\bar{u}$ (ms^{-1})	$\bar{p}_d$ (Wm^{-2})	σ (ms^{-1})	k
MAE	Baseline	0.12	5.8	0.10	0.12
	LR	0.12	5.2	<0.1	0.13
	VR	0.14	6.1	<0.1	<0.1
MBE	Baseline	<0.1	-1.6	<0.1	0.11
	LR	<0.1	-2.2	<0.1	0.13
	VR	<0.1	1.5	<0.1	<0.1
σ_{MBE}	Baseline	0.15	8.8	0.11	<0.1
	LR	0.15	7.4	<0.1	<0.1
	VR	0.16	8.7	0.11	<0.1

Table 3: Error metrics for the wind resource parameters of $\bar{u}$, $\bar{p}_d$, σ and k using UK4 reference data at 50 m and training periods of 3 months (left) and 12 months (right). The baseline using Rf reference data is also shown. Values are averaged across 22 site pairs and 120 starting months. σ_{MBE} represents the standard deviation across all sites and starting months.

3.3 Effects of terrain

To explore the performance of the UK4 data in different terrains, the average %Error in $\bar{p}_d$ averaged across all seasons using 3 and 12 month training periods is shown in Figure 5 for each site using UK4 and Rf reference data.

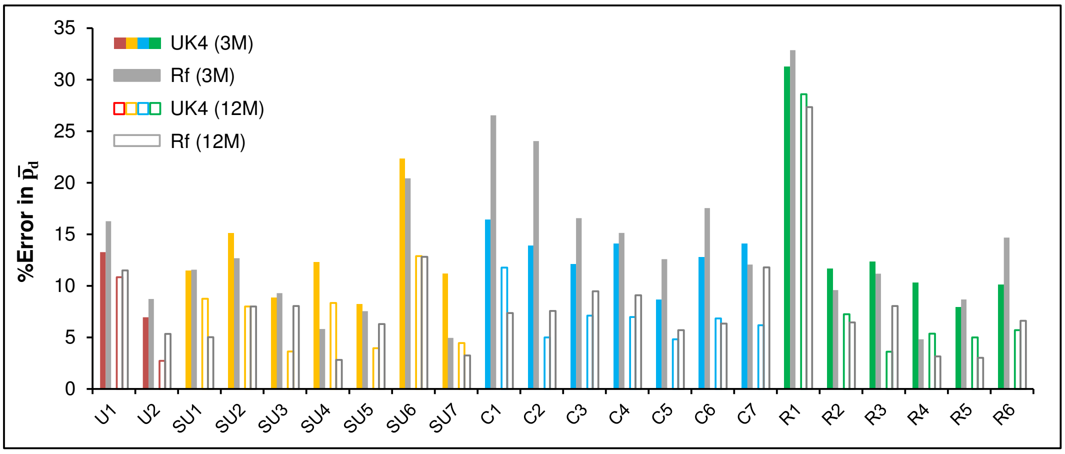

Figure 5: %Error in $\bar{p}_d$ for each of the 22 target sites using training periods of 3 or 12 months and UK4 or Rf reference data. Values are averaged across 120 starting months.

Comparison with Table 2 shows that the r_{WI} and r_u metrics are reasonable predictors for the best reference data source when using a 3 month MCP training period, although they are less successful for 12 month training periods where the absolute error differences between UK4 and Rf are smaller. For the urban, suburban and rural sites, the results are variable with the UK4 data resulting in reduced errors at 7 sites compared to 8 sites for the Rf reference data, (for both 3 and 12 month training periods). This indicates that while UK4 performs adequately overall, it may not consistently outperform Rf. This is likely because UK4 data are not of sufficiently high resolution to represent local effects caused by complex terrain or the built environment. However, for coastal sites there appears to be a clear preference for UK4 with reduced errors at 6 out of 7 sites (5 out of 7 when using a 12 month training period).

Due to the complexity of coastal wind flows, these locations may be particularly subject to highly localised climates [42]. For example, the wind speed and direction at a nearby Rf site are likely to depend on the coastal orientation, proximity of the sea and local stability, attributes that may vary significantly over short distances. Since the UK4 data are co-located with the target sites, this reduces the impact of such localised effects. The larger error reduction when using UK4 compared to Rf for a 3 month training period indicates that UK4 may also reduce errors related to seasonal variability. Since stability effects, as well as the thermal contrast between land and sea that gives rise to sea breezes, are likely to have strong seasonal components [39], the improved performance of UK4 may be an indication that it is better able to represent seasonally varying local conditions and fetch effects, compared to Rf data located tens of kilometres away. To investigate this further, the diurnal variation in r_u at the coastal sites was compared using UK4 and Rf reference data. If, as suggested, UK4 better represents variable stability conditions, this should result in reduced diurnal variability in r_u when using UK4. To allow multiple sites to be easily compared, the diurnal r_u values were first normalised by the average value for each site shown in Table 2. The normalised values, averaged across the 11 year data record, are presented in Figure 6.

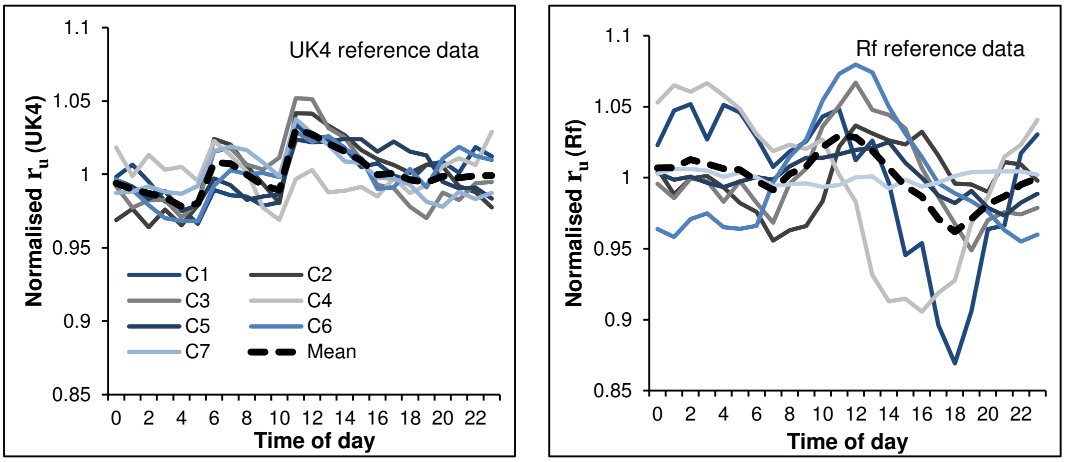

Figure 6: Diurnal variability in the normalised linear correlation coefficient r_u between the reference and target site hourly wind speeds using UK4 (left) and Rf (right) at coastal sites. Values are averaged over the 11 year data record.

Figure 6 demonstrates that the use of UK4 markedly reduces the diurnal variation in r_u for the majority of coastal sites. This implies that UK4 is better able to represent the diurnal changes at such sites, and by extension, the seasonal variability related to these effects. Equivalent analysis of non-coastal sites revealed relatively small differences in the diurnal variability in r_u between UK4 and Rf. This is likely because Rf data are more capable of representing the stability conditions at sites not subject to complex coastal climates, or equally, because the impact on r_u is less significant in non-coastal areas.

3.4 Effect of hindcast

As mentioned previously, the UK4 reference data pre-2007 were obtained from a hindcast initialised every 48 hours using ERA-i. This may affect the pre-2007 data in two ways; (i) through the longer time between re-initialisation (48 hours compared to 24) since forecast skill generally decreases with forecast duration and (ii) through differences in ERA-i compared to the UK4 analysis. To investigate this further, the linear correlation coefficient r_u, between observed and UK4 predicted hourly wind speeds, was calculated for all 37 observational sites (22 target plus 15 Rf) using the following procedure:

(i) The 11 year data sets (UK4 and observed) were split into two periods covering the hindcast (Aug 2001 – Dec 2006) and the forecast (Jan 2007 – July 2012).

(ii) For the hindcast period, r_u was calculated before and after removing data every second day corresponding to hours 25 – 48, where hour 1 corresponds to the time of re-initialisation.

To serve as a control, data from every second day was also removed from the forecast period. Since these data are re-initialised every 24 hours, this was expected to have no effect on r_u.

For the forecast period, the average r_u was found to be 0.83, and as expected, no change was observed after removing data from every second day. For the hindcast period, the average r_u values were 0.77 and 0.81 before and after removal of hours 25 – 48 respectively. This implies that the effect of the 48 hour re-initialisation period is slightly larger (reduction of 0.04 in r_u) than the effect of using ERA-i for re-initialisation in place of the UK4 analysis (reduction of 0.02 in r_u). It should be noted that this conclusion is based on the assumption that no additional factors are responsible for the reduction in r_u pre- and post-2007.

To investigate the impact of the 48 hour re-initialisation period on the accuracy of the long-term MCP predictions, the MCP algorithms were applied at the 22 target sites using the full 11 year data record under two conditions: (i) after removing hours 1 - 24 (the higher skill period) and (ii) after removing hours 25 - 48 (the lower skill period) from the hindcast data. These conditions were chosen to provide the highest contrast possible between the two data sets. MCP training periods were restricted to the post-2007 data to avoid removing entries from the short-term training periods, while the full 11 year data record was used for prediction. Using this approach, the error metrics were averaged over 48 sliding window positions, compared to 120 used previously. The results showed an average change in the %Error for $\bar{u}$ and $\bar{p}_d$ of < 0.3 percentage points under the two conditions, for both 3 and 12 month training periods. This indicates that the 48 hour re-initialisation period used for the hindcast does not significantly impact on the accuracy of the long-term predicted wind resource using MCP. Note that the MCP approach involves extrapolation of a fixed reference/target site correlation, obtained from a short-term training period, across a long-term prediction period. Since this correlation will exhibit intra- and inter-annual variability, as reflected in σ_{MBE}, it is not surprising that the additional variability due to the use of hindcast data does not significantly affect the prediction errors. While it is not possible to use the same approach to isolate the effect of using ERA-i for initialising the hindcast UK4, given that this appears to have a lesser impact on r_u compared to the time between re-initialisation, it is likely that this also has a lesser impact on the long-term predicted wind resource parameters.

3.5 Extension to 37 test sites

When using UK4 data exclusively as a reference data source, there is an opportunity to expand the test sites from 22 to include all 37 target and reference (Rf) sites listed in Table

1. Using this approach, both the target and reference sites are treated as locations where we wish to predict the wind resource and the reference data are obtained solely from UK4. The additional Rf sites are located mostly in rural or coastal areas. The error metrics for both data sets (22 target sites and 37 reference plus target sites) are compared in Table 4 for training periods of 3 and 12 months using the LR algorithm.

		No. Sites	$\bar{u}$	$\bar{p}_d$	σ	k
%Error	**3M**	22	4.3	13	6.3	8.3
		37	4.1	12	6.1	8.1
	12M	22	3.0	7.6	4.1	7.2
		37	2.8	7.3	4.0	7.0
			$\bar{u}$ (ms^{-1})	$\bar{p}_d$ (Wm^{-2})	σ (ms^{-1})	k
MAE	**3M**	22	0.18	9.3	0.15	0.15
		37	0.18	10	0.15	0.15
	12M	22	0.12	5.2	<0.1	0.13
		37	0.12	5.7	0.10	0.13
MBE	**3M**	22	<0.1	-3.1	-0.11	0.14
		37	<0.1	-3.5	-0.11	0.14
	12M	22	<0.1	-2.2	<0.1	0.13
		37	<0.1	-2.6	<0.1	0.13
σ_{MBE}	**3M**	22	0.23	14	0.16	0.11
		37	0.23	14	0.16	0.11
	12M	22	0.15	7.4	<0.1	<0.1
		37	0.15	7.7	<0.1	<0.1

Table 4: Error metrics for the wind resource parameters of $\bar{u}$, $\bar{p}_d$, σ and k for training periods of 3 months (3M) and 12 months (12M) using UK4 reference data and LR. The error metrics are shown for the 22 target sites and the combined 37 target plus reference sites, averaged across 120 starting months.

It is clear that the average error metrics remain broadly unchanged even after incorporating the additional 15 sites for both 3 and 12 month training periods, thus increasing the confidence that these results are broadly representative of UK sites.

4 Conclusions

Output from a state-of-the-art UK forecast model has been investigated in terms of its suitability for use as long-term reference data in MCP. The ability of the UK4 data to represent long-term trends in the target site wind speeds was investigated through the calculation of a monthly wind index. Based on the linear correlation between UK4 predicted and observed monthly wind indices (r_{WI}), the UK4 data were shown to be capable of representing the long-term trends in the target site wind speeds, on average, slightly better than observations at nearby meteorological stations. In addition, the use of hindcast data to extend the UK4 data set was found to have little impact on the accuracy of the predicted long-term wind resource using MCP despite a reduction in the linear correlation r_u between UK4 predicted and observed hourly wind speeds.

A forecast height of 50 m was shown to result in the lowest errors in the predicted $\bar{u}$ and $\bar{p}_d$ using linear MCP algorithms. Based on this, MCP approaches were used to calculate the 10 year wind resource at 22 target sites (later extended to 37) using both the UK4 data and nearby meteorological observations as reference sources. The UK4 data were shown to be highly competitive with nearby meteorological observations when used as a long-term reference source for a range of training periods. At coastal sites, a systematic improvement in the predicted $\bar{p}_d$ was observed when using UK4 reference data. Further analysis revealed that this is likely due to improved representation of the local stability conditions and associated wind flows when using UK4 compared to nearby meteorological observations.

These results indicate that UK4 could provide a valuable source of reference data for the implementation of MCP approaches, with particularly utility at coastal sites.

5 Acknowledgements

This work was financially supported by the Engineering and Physical Sciences Research Council through the Doctoral Training Centre in Low Carbon Technologies. Additional funding was received from the University of Leeds, Climate and Geohazard Services

6 References

1. World Wind Energy Association, *Small Wind World Report*, March 2014.
2. RenewableUK, *Small and Medium Wind Strategy – The current and future potential of the sub-500 kW wind industy in the UK*, November 2014.
3. Carta, J.A., S. Velázquez, and P. Cabrera, *A review of measure-correlate-predict (MCP) methods used to estimate long-term wind characteristics at a target site.* Renewable and Sustainable Energy Reviews, 2013. **27**: p. 362-400.
4. AWS Scientific Inc. & National Renewable Energy Laboratory (U.S.), *Wind Resource Assessment Handbook - fundamentals for conducting a successful monitoring program*, 1997.
5. Weekes, S.M. and A.S. Tomlin, *Data efficient measure-correlate-predict approaches to wind resource assessment for small-scale wind energy.* Renewable Energy 2014. **63**: p. 162-171.
6. Brower, M.C., *The use of NCEP/NCAR reanalysis data in MCP.* European Wind Energy Association annual conference, Athens, 2006.
7. Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, *The NCEP/NCAR 40-year reanalysis project.* Bulletin of the American Meteorological Society, 1996. **77**(3): p. 437-471.
8. J. M. Rodrigo, *State-of-the-art wind resource assessment, WAUDIT Deliverable D-7*, 2010, CENER National renewable energy centre.
9. Liléo, S. and O. Petrick, *Investigation on the use of NCEP/NCAR, MERRA and NCEP/CFSR reanalysis data in wind resource analysis.* European Wind Energy Association annual conference, Brussels, 2011.

10. Rienecker, M.M., M.J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M.G. Bosilovich, S.D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W. Gu, J. Joiner, R.D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P. Pegion, C.R. Redder, R. Reichle, F.R. Robertson, A.G. Ruddick, M. Sienkiewicz, and J. Woollen, *MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications.* Journal of Climate, 2011. **24**(14): p. 3624-3648.

11. Saha, S., S. Moorthi, H.L. Pan, X.R. Wu, J.D. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H.X. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y.T. Hou, H.Y. Chuang, H.M.H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H.L. Wei, R.Q. Yang, S. Lord, H. Van den Dool, A. Kumar, W.Q. Wang, C. Long, M. Chelliah, Y. Xue, B.Y. Huang, J.K. Schemm, W. Ebisuzaki, R. Lin, P.P. Xie, M.Y. Chen, S.T. Zhou, W. Higgins, C.Z. Zou, Q.H. Liu, Y. Chen, Y. Han, L. Cucurull, R.W. Reynolds, G. Rutledge, and M. Goldberg, *The NCEP climate forecast system reanalysis.* Bulletin of the American Meteorological Society, 2010. **91**(8): p. 1015-1057.

12. Pinto, C., R. Guedes, P. Pinto, and M. Ferreira, *NCEP/NCAR reanalysis data for the Portuguese mainland.* European Wind Energy Association annual conference, Athens, 2006.

13. Jimenez, B., K. Moennich, and F. Durante, *Comparison between NCEP/NCAR and MERRA reanalysis data for long-term correction in wind energy assessment.* European Wind Energy Association annual conference, Copenhagen, 2012.

14. Taylor, M., J. Freedman, K. Waight, and M. Brower, *Using simulated wind data from a mesoscale model in MCP*, 2009, AWS Truewind.

15. Davies, T., M.J.P. Cullen, A.J. Malcolm, M.H. Mawson, A. Staniforth, A.A. White, and N. Wood, *A new dynamical core for the Met Office's global and regional modelling of the atmosphere.* Quarterly Journal of the Royal Meteorological Society, 2005. **131**(608): p. 1759-1782.

16. Wilson, C. and J. Standen, *The added value and validation of mesoscale models compared to atmospheric reanalyses for resource assessment.* European Wind Energy Association annual conference, Vienna, 2013.

17. Nielsen, M., L. Landberg, N.G. Mortensen, R.J. Barthelmie, and A. Joensen, *Application of measure-correlate-predict approach for wind resource measurement.* European Wind Energy Association Annual Conference, Copenhagen, 2001.

18. Achberger, C., M. Ekstrom, and L. Barring, *Estimation of local near-surface wind conditions - a comparison of WASP and regression based techniques.* Meteorological Applications, 2002. **9**(2): p. 211-221.

19. Riedel, V., M. Strack, and H. Waldl, *Robust approximation of functional relationships between meteorological data: Alternative measure-correlate-predict algorithms.* European Wind Energy Association annual conference, Copenhagen, 2001.

20. Clive, P., *Non-linearity in MCP with Weibull distributed wind speeds.* Wind Engineering, 2008. **32**(3): p. 319-323.

21. Mortimer A., *A new correlation/prediction method for potential wind farm sites.* Proceedings of the BWEA, 1994, 1994.

22. Woods, J.C. and S.J. Watson, *A new matrix method of predicting long-term wind roses with MCP.* Journal of Wind Engineering and Industrial Aerodynamics, 1997. **66**(2): p. 85-94.

23. Bechrakis, D.A., J.P. Deane, and E.J. McKeogh, *Wind resource assessment of an area using short term data correlated to a long term data set.* Solar Energy, 2004. **76**(6): p. 725-732.

24. Monfared, M., H. Rastegar, and H.M. Kojabadi, *A new strategy for wind speed forecasting using artificial intelligent methods.* Renewable Energy, 2009. **34**(3): p. 845-848.

25. Velazquez, S., J.A. Carta, and J.M. Matias, *Comparison between ANNs and linear MCP algorithms in the long-term estimation of the cost per kW h produced by a wind turbine at a candidate site: A case study in the Canary Islands.* Applied Energy, 2011. **88**(11): p. 3869-3881.

26. Carta, J.A., P. Ramirez, and C. Bueno, *A joint probability density function of wind speed, and direction for wind energy analysis.* Energy Conversion and Management, 2008. **49**(6): p. 1309-1320.

27. Romo Perea, A., J. Amezcua, and O. Probst, *Validation of three new measure-correlate-predict models for the long-term prospection of the wind resource.* Journal of Renewable and Sustainable Energy, 2011. **3**(2).

28. Weekes, S.M. and A.S. Tomlin, *Comparison between the bivariate Weibull probability approach and linear regression for assessment of the long-term wind energy resource using MCP.* Renewable Energy, 2014. **68**: p. 529-539.

29. Derrick, A., *Development of the measure-correlate-predict strategy for site assessment.* British Wind Energy Association annual conference, 1992: p. 259-265.

30. Rogers, A.L., J.W. Rogers, and J.F. Manwell, *Comparison of the performance of four measure-correlate-predict algorithms.* Journal of Wind Engineering and Industrial Aerodynamics, 2005. **93**(3): p. 243-264.

31. Ellison, S.L.R., V.J. Barwick, and T.J.D. Farrant, *Practical statistics for the analytical scientist, a bench guide*. 2009, RSC Publishing.

32. Bass, J.H., M. Rebbeck, L. Landberg, M. Cabre, and A. Hunter, *An improved measure-correlate-predict algorithm for the prediction of the long term wind climate in regions of complex environment*, 2000, Joule Project JOR3-CT98-0295.

33. Met Office, *Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853 - current). NCAS British Atmospheric Data Centre 2012.*

34. Dee, D.P., S.M. Uppala, A.J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M.A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S.B. Healy, H. Hersbach, E.V. Holm, L. Isaksen, P. Kallberg, M. Koehler, M. Matricardi, A.P. McNally, B.M. Monge-Sanz, J.J. Morcrette, B.K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.N. Thepaut, and F. Vitart, *The ERA-Interim reanalysis: configuration and performance of the data assimilation system.* Quarterly Journal of the Royal Meteorological Society, 2011. **137**(656): p. 553-597.

35. Met Office, *Global assimilation of air temperature, humidity, wind and pressure from surface stations: practice and performance.* Forecasting Research Technical Report No: 582, 2013.

36. Macpherson, B., B.J. Wright, W.H. Hand, and A.J. Maycock, *The impact of MOPS moisture data in the U.K. Meteorological Office mesoscale data assimilation scheme.* Monthly Weather Review, 1996. **124**(8): p. 1746-1766.

37. Auría, A.C. and B.N. Fernández-Victorio, *Usage of SYNOP 10 meters wind observations in HIRLAM.* HIRLAM Newsletter, 2009. **50**(03): p. 8-17.

38. Watson, S.J., P. Kritharas, and G.J. Hodgson, *Wind speed variability across the UK between 1957 and 2011.* Wind Energy, DOI: 10.1002/we.1679, 2013.

39. Barthelmie, R.J., *The effects of atmospheric stability on coastal wind climates.* Meteorological Applications, 1999. **6**(1): p. 39-47.

40. Howard, T. and P. Clark, *Correction and downscaling of NWP wind speed forecasts.* Meteorological Applications, 2007. **14**(2): p. 105-116.

41. Wieringa, J., *Shapes of annual frequency distributions of wind speed observed on high meteorological masts.* Boundary-Layer Meteorology, 1989. **47**(1-4): p. 85-110.

42. Barthelmie, R.J. and J.P. Palutikof, *Coastal wind speed modelling for wind energy applications.* Journal of Wind Engineering and Industrial Aerodynamics, 1996. **62**(2–3): p. 213-236.