A new statistical approach to downscale wind speed distributions at a site in northern Europe

Annemarie Devis,¹ Nicole P.M. van Lipzig,¹ and Matthias Demuzere¹

Received 3 August 2012; revised 28 January 2013; accepted 1 February 2013.

[1] This paper explores a statistical regression approach to downscale large-scale global circulation model output to the wind speed distribution at the hub-height of tall wind turbines. The methodology is developed for Cabauw, using observational, ERA-Interim and ECHAM5 data. The regression analysis is based on the parameters of the probability distribution functions (pdfs) and includes a variable evaluation prior to the development of the statistical models. During winter ECHAM5 performs very well in representing the ERA-Interim wind speed pdf at hub-height. However, during summer, the hub-height wind speed pdf is not well represented by ECHAM5. A regression analysis shows that during summer-day the hub-height wind speed is strongly linked to the wind speed at higher, skillfully represented levels. The summer-day hub-height wind speed can therefore be skillfully predicted using wind speed pdf parameters of higher levels (R^2 of the model using 500 m wind speed scale parameter as a predictor is 0.84). During the summer-night, the stable boundary layer is much shallower and the statistical model shows that solely the higher level wind speed is not able to skillfully predict the hub-height wind speed pdf (R^2 of 0.59). Including temperature information in the downscaling model substantially improves the prediction of the summer-night hub-height wind speed pdf (R^2 adjusted of 0.68).

Citation: Devis, A., N. P. M. van Lipzig, and M. Demuzere (2013), A new statistical approach to downscale wind speed distributions at a site in northern Europe, *J. Geophys. Res. Atmos.*, 118, doi:10.1002/jgrd.50245.

1. Introduction

[2] Because the global climate is warming and fossil fuel resources are diminishing, decreasing or increasing trends in wind speed are highly relevant for the wind energy sector [Pryor *et al.*, 2005a]. Moreover, because the power contained in the wind is proportional to the wind speed cubed [McVicar and Roderick, 2010], even small wind prediction errors result in large power prediction errors, also on a climate scale. Therefore, it is of major importance to have a solid indication of present and future wind climate at the height of the hub of the wind turbine.

[3] There are indications that the wind speed has been changing in recent decades. Vautard *et al.* [2010] used 10 m wind speeds from 822 stations in the Northern Hemisphere and found an average reduction of observed near-surface wind speeds of 0.11 m/s per decade between 1979 and 2008. They argue that this decrease in the 10 m wind speed can partly be explained by an increase in surface roughness. On the other hand, reanalysis wind speed fields at 850 hPa from the National Center for Environmental

Prediction and the National Center for Atmospheric Research indicate that there is a significant increase in annual mean wind speed over the Baltic during the latter half of the 20th century. This increase is attributable to an increase in westerlies, which are in turn related to the prevalence of the positive phase of the North Atlantic Oscillation [Pryor *et al.*, 2005b].

[4] The current state-of-the-art of global circulation models (GCMs) are getting more and more realistic in simulating synoptic-scale weather systems, which allows for the assessment of trends in atmospheric variables. For example, Demuzere *et al.* [2008] and Donat *et al.* [2010] found an increased frequency of westerly flow during winter over central Europe based on present-day and future Intergovernmental Panel on Climate Change SRES A1B scenarios with ECHAM5. However, because typical horizontal and vertical resolutions of the climate models are still very coarse, many practical purposes, e.g., wind energy yield estimations, require to downscale GCM information to finer scales. This downscaling can be undertaken dynamically or statistically. For the dynamical downscaling, a high-resolution (typically 20–50 km) regional climate model (RCM) is integrated for a limited domain, using boundary conditions provided by the GCM simulation [Räisänen *et al.*, 2004]. For the statistical downscaling, a transfer function is developed, which statistically relates large-scale climate parameters generated by the GCM to the observed near-surface parameter of interest [Pryor *et al.*, 2005a]. In contrast to the dynamical method, the

¹Earth and Environmental Sciences, KULeuven, Heverlee, Leuven, Belgium.

Corresponding author: A. Devis, Earth and Environmental Sciences, KULeuven, Celestijnenlaan 200E, 3001, Heverlee (Leuven), Belgium. (annemarie.devis@ees.kuleuven.be)

©2013. American Geophysical Union. All Rights Reserved.
2169-897X/13/10.1002/jgrd.50245

statistical downscaling is relatively easy to use and can be rapidly applied to an ensemble of GCMs [Wilby *et al.*, 2004]. The performance of the statistical technique has shown to be comparable to the dynamical modeling. [Kidson and Thompson, 1998; Solman and Neñez, 1999; Schoof and Pryor, 2001].

[5] Statistical downscaling techniques are used for both short-term wind prediction [de Rooy and Kok, 2004; Salameh *et al.*, 2009; Howard and Clark, 2007; Slougher *et al.*, 2010] and for climate scales [Kaas *et al.*, 1996; Sailor *et al.*, 2000; Pryor *et al.*, 2005a; Klink, 2007; Minvielle *et al.*, 2011; Monahan *et al.*, 2011; Curry *et al.*, 2012, van der Kamp *et al.*, 2012]. Various types of statistical models exist, such as the weather typing approach [Salameh *et al.*, 2009; Minvielle *et al.*, 2011], the use of neural networks [Sailor *et al.*, 2000], and the MOS (model output statistics) technique [Kaas *et al.*, 1996; de Rooy and Kok, 2004; Pryor *et al.*, 2005a; Klink, 2007; Monahan, 2012].

[6] The MOS technique is based on the statistical relationship between small-scale and large-scale circulation variables [Glahn and Lowry, 1972]. This multiple linear regression method is the most simple statistical model to implement and when used with an appropriate cross-validation, it produces statistically robust prediction models [Monahan, 2012]. Most MOS studies focus on downscaling time series (daily values) of the geophysical variable of interest. Alternatively, Pryor *et al.* [2005a] presented the downscaling of the parameters of the probability distribution functions (hereafter abbreviated as pdf) of the wind speed, because (i) GCMs have difficulties in reproducing the time structure of geophysical parameters and (ii) the needs of user communities require a description of the pdf of wind speed.

[7] This paper uses a MOS approach to downscale GCM model output to the small-scale wind speed at the hub-height. As proposed by Pryor *et al.* [2005a] the regression is performed on the first and the second moment of the pdfs. In addition to the arguments of Pryor *et al.* [2005] favoring the downscaling of the pdf parameters, the latter also allows for the use of a linear regression method because the direct use of daily wind speeds are not in agreement with the regression assumptions with regard to the homoscedasticity and the normality of the residuals [Kutner *et al.*, 2005].

[8] The downscaling methodology presented in this paper uses observational, reanalysis, and GCM data. It is in principle applicable to observations, reanalysis, and GCM data of every kind. ECHAM5 data are used because this GCM has shown to perform well for Europe [van Ulden and van Oldenborgh, 2005]. The same is true for the ERA-Interim reanalysis data [Dee *et al.*, 2011]. For the observation data, the wind speeds at 140 m altitude (which is in the remainder of the paper referred to as hub-height) are used. Those are measured by the meteorological mast of Cabauw, which is situated in an open grassland terrain in the Netherlands. The environmental and meteorological conditions of the Cabauw site are representative for other locations in Western-Europe.

[9] A more thorough description of the data sets is given in section 2. Section 3 presents the method of selecting the predictors and developing the multiple regression models while the results of the downscaling of the hub-height wind speed is shown in section 4. Section 5 discusses the link between the predictability of the wind speed and the

boundary layer development throughout the day, including the interpretation of the prevalence of the nocturnal low level jet. The conclusions are summarized in section 6.

2. Data Sources and Treatment

[10] The goal of this paper is to downscale GCM output (from ECHAM5) to a specific location in the Netherlands where wind speed tower measurements are available. To do so, a third data set that acts as a “bridge” between the GCM and the point observations is required. This third data set is reanalysis data (from ERA-Interim). The reanalysis data is useful because it shares characteristics of both the GCM and observations. The reanalysis is a real-weather historical data set that correlates well with the observations at hourly to monthly time scales. Therefore, it is appropriate for developing linear relationships to the observations. The GCM output on the other hand cannot directly be used for developing linear relations between its output data and observations because it is an independent climate simulation that is uncorrelated to observations on hourly to monthly time scales (aside from the seasonal and diurnal cycles). Like the GCM, the reanalysis is a gridded data set (and at matching resolution once it is aggregated to the GCM grid), and is built using a model similar to the GCM. Therefore, it is likely that a transfer function that is developed between the reanalysis and observations, for variables that show similar behavior in both the GCM and reanalysis, can be applied to relate GCM output to the observations.

2.1. Observations

[11] The Cabauw observational data set of the Royal Dutch Meteorological Institute is used in this study. Many studies on boundary layer meteorology have been carried out using the data of this tower [van Ulden and Wieringa, 1996; Baas *et al.*, 2009; Verkaik and Holtslag, 2007; Demuzere and van Lipzig, 2010a, 2010b; Monahan *et al.*, 2011]. A detailed documentation on the Cabauw measurement site and instruments can be found on the Database of CESAR (<http://www.cesar-observatory.nl/>).

[12] The Cabauw mast is situated in the central river delta in the southwestern part of the Netherlands, more than 50 km away from the North Sea (51.97°N, 4.93°E, −0.7 m above sea level). This measurement mast is chosen because of its extensive and reliable profile data set of the atmospheric boundary layer, covering a time span of more than 30 years, in which no significant changes happened to the surroundings. Moreover, the mast is located in flat grassland terrain where surface elevation changes are at most a few meters over 20 km. Near the measurement site, the terrain is open pasture for at least 400 m in all directions, and in the WSW direction for at least 2 km. Further away, the landscape is generally very open in the westerly direction, while the distant east direction is more rough (windbreaks, orchards, low houses). The distant north and south directions are mixed landscapes, much pasture, and some windbreaks [Vermeulen *et al.*, 2011].

[13] The selected observational data set consists of wind speed observations (m/s) from the periods 1989 to 1996 (as 30 min averages) and 2000 to 2009 (as 10 min averages) for levels of 10, 20, 40, 80, 140, and 200 m height. The measuring period was interrupted from 1997 to 2000 to perform a major refurbishment of the tall mast and its installations.

2.2. Reanalysis Data

[14] The ERA-Interim data set, produced by the European Centre for Medium-Range Weather Forecasts, is used in this study. A large number of satellite observations of upper-air wind fields are assimilated in ERA-Interim, based on a 12-hourly four-dimensional variational analysis algorithm [Dee *et al.*, 2011].

[15] Instantaneous 12 and 00 UTC reanalysis atmospheric variables are used for the 18 year period (same period as the observations) for a region of approximately 500 km by 500 km covering the Cabauw study area. The data have a horizontal resolution of approximately $0.75^\circ \times 0.75^\circ$, which corresponds to a resolution of ~ 50 km in meridional direction to ~ 80 km in zonal direction (Figure 1). In this study the 11 lowest model levels are used, with model level heights at an average of 10, 35, 72, 125, 197, 290, 407, 550, 719, 916, and 1142 m.

2.3. Global Circulation Model Data

[16] The GCM used in this project is ECHAM5, which was developed at Max Planck Institute. ECHAM5 has shown to perform well for Europe and unlike many other GCMs, ECHAM5 does not indicate a westerly bias in the large-scale circulation during winter [van Ulden and van Oldenborgh, 2005; Demuzere *et al.*, 2008]. ECHAM5 was used in the Intergovernmental Panel on Climate Change Fourth Assessment Report, alongside many other GCMs from different countries.

[17] Instantaneous 12 and 00 UTC ECHAM5 global climate model data are used for the period from 1989 to 2010 and the model grid cell covering the Cabauw site. This grid cell has a horizontal resolution of ~ 200 km in meridional direction to ~ 150 km in zonal direction (Figure 1). The 1 km lowest atmosphere is covered by 6 levels (at an average of 33, 148, 354, 632, 969, and 1354 m).

Figure 1. Aggregation of the ERA-Interim grid cells (black dotted lines) to the ECHAM5 grid cell covering the Cabauw site (red polygon). The cross-hatched area represents the ERA-Interim grid cells used for aggregation.

2.4. Data Treatment

[18] The 0.75° resolution grid of the reanalysis data set is aggregated to the ECHAM5 grid cell covering the Cabauw study site (4.69°E – 6.56°E , 50.36°N – 52.23°N) (Figure 1). The aggregation uses the weighted average of the ERA-Interim grid cells which are covered by the ECHAM5 grid cell, proportional to their overlapping area with the ECHAM5 grid cell. The model levels of ERA-Interim and ECHAM5 have been interpolated to 10, 140, 500, and 1000 m for temperature and wind speed, and to 140, 500, and 1000 m for specific humidity. These levels are chosen because they describe different atmospheric conditions in the boundary layer and the free atmosphere. Temperature and specific humidity are interpolated using a linear profile. The interpolation of the wind speed at level z is based on the power law, which is commonly used for wind [Archer and Jacobson, 2005; Pryor *et al.*, 2005a] and is typically written as

$$U(z) = U_{\text{ref}} \cdot \left(\frac{z}{z_{\text{ref}}} \right)^\alpha \quad (1)$$

[19] However, in contrast to most previous studies, we do not consider the friction coefficient α as a constant value. For each z -level α was calculated by using the two closest model levels that bracket the z -level:

$$a = \frac{\ln(U(z+1)/U(z-1))}{\ln((z+1)/(z-1))}$$

[20] By doing so, we take into account the variability of α in time and height due to a change in surface roughness and in atmospheric stability [Holt and Wang, 2011]. After the interpolations, temperature and wind speed are used to calculate gradients between 140 m to surface, 500 m to surface, and 1000–500 m. The model variables used in this study are listed in Table 1.

[21] The ERA-Interim and ECHAM5 data sets are split in the following four subsets: summer-day, summer-night, winter-day and winter-night, which will hereafter be abbreviated as respectively SD, SN, WD, WN, whereby

Table 1. List of the Model Variables and Their Short Names

Variable	Short Name
Sea level pressure	SLP
Surface temperature (10 m)	T10
Surface wind speed (10 m)	W10
Hub-height temperature (140 m)	T140
Hub-height specific humidity (140 m)	Q140
Hub-height wind speed (140 m)	W140
Temperature (500 m)	T500
Specific humidity (500 m)	Q500
Wind speed (500 m)	W500
Temperature (1 km)	T1000
Specific humidity (1 km)	Q1000
Wind speed (1 km)	W1000
Temperature gradient between 140 and 10 m	DT (T140-T10)
Temperature gradient between 500 and 10 m	DT (T500-T10)
Temperature gradient between 1 km and 500 m	DT (T1000-T500)
Wind speed gradient between 140 and 10 m	DW (W140-W10)
Wind speed gradient between 500 and 10 m	DW (W500-W10)
Wind speed gradient between 1 km and 500 m	DW (W1000-W500)

summer is composed by the months May to September, winter is covering the months November to March, and day and night being respectively 12 UTC and 00 UTC. Because of the strong diurnal and seasonal dependency of the wind speed characteristics, the statistical downscaling is performed for each subset separately.

3. Method

[22] The statistical downscaling model is based on an empirical relationship between observed local-scale hub-height wind speed pdf parameters (predictands) and the pdf parameters of large-scale atmospheric variables (predictors). The transfer function is derived using ERA-Interim data aggregated to the ECHAM5 grid cell. To select the predictors, the GCM output is compared to the reanalysis data which is aggregated to the GCM grid. Only those variables for which the GCM represents the statistical behavior of the corresponding reanalysis variable are used in the development of the transfer function. The underlying idea is that variables that are skillfully simulated for the present-day climate are more likely to be skillfully predicted for the future; the same is true for a statistical downscaling model based on these variables. Once the statistical downscaling model is derived, it is applied on ECHAM5 data that has been bias-corrected relative to ERA-Interim data aggregated to the ECHAM5 grid cell. Because the predictors are selected after they were evaluated, the bias that needs to be corrected for is small.

3.1. ECHAM5 Variable Evaluation

[23] Global circulation models output is commonly evaluated based on their capability in simulating the mean and standard deviation. However, because the statistical downscaling is based on pdfs, the GCM is required to be skilful in simulating the entire pdf. Evaluating GCM output based on pdfs has the major advantage that if a GCM is able to simulate the entire pdf, it demonstrates a capability to simulate values that are currently rare and that may become more common in the future. Furthermore, establishing the skill of a climate model to simulate the whole pdf is a far harder test for a model than the mean and standard deviation. Thus, by succeeding in such a test, we might have more confidence in projections made with this model [Perkins *et al.*, 2007].

[24] The evaluation of ECHAM5 is performed against the ERA-Interim reanalysis data aggregated to the ECHAM5 grid cell and makes use of the skill score suggested by Perkins *et al.* [2007]. The skill score (equation (2)), measures the common area of two pdfs, by calculating the cumulative minimum value of two distributions of each binned value. In a formula Perkins' skill score is written as

$$S_{\text{score}} = \sum_1^n \min(Z_{\text{modeled}}, Z_{\text{observed}}) \quad (2)$$

where the Z quantities are the histograms of variability and n is the number of bins in the histogram.

[25] This is a very simple measure that provides a robust and comparable score of the relative similarity between the pdfs of two variables. If the S_{score} is close to 0, there is negligible overlap between the two pdfs, while if the two pdfs are exactly the same, the S_{score} will be 1.

Perkins *et al.* [2007] consider the variables with scores smaller than 0.7 as weakly simulated. To verify the validity of the original 0.7 threshold, the results of the Perkins' score are compared to the similar and more widely used Kolmogorov-Smirnov (KS) test-statistic. The KS-test does not depend on the underlying pdf being tested and its null hypothesis states that the samples are drawn from the same distribution [Massey, 1951]. The results of this comparison reveal that the 0.7 Perkins threshold coincides with the 0.05 significance level for the KS-test. Therefore, we adopted this threshold of 0.7 and excluded variables from the analysis with scores beneath this threshold.

[26] It must be noted that because reanalysis data are not necessarily the truth (especially in case of wind speed in the lower atmosphere), this statistical evaluation analysis cannot blindly be regarded as a true evaluation of the ECHAM5 output. A true evaluation would need observational data. Observations are not used in the evaluation because they are not available for all variables and because of the scale difference between the local-scale site observations and the large-scale grid cell of the GCM. The GCM is not supposed to be skilful at point scale; it should be skilful at the scale of the area of the GCM grid cell. Moreover, as the set up of the statistical downscaling model is based on reanalysis data, the GCM predictors used in the model should resemble the reanalysis data and not the observations. Therefore, the statistical downscaling is performed on the variables that are evaluated based on the aggregated ERA-Interim data.

3.2. Probability Density Function Parameter Calculation

[27] Only those variables that are skillfully represented by ECHAM5 are used in the linear regression analysis. As the regression is based on pdf parameters, the data sets of the predictands and possible predictors need to be divided in samples, for which the pdfs are fitted and their parameters are estimated. The number of samples is a trade-off between the strength of the regression analysis and the precision of the pdf parameter estimation. The more samples are used in the regression, the stronger the analysis will be. However, the more samples are derived from the data set, the less records are available to be used to fit the pdf of each sample, and hence the more difficult it is to find a fitting pdf. According to the statistical test of Cohen *et al.* [2003] the minimum required sample size for a multiple regression with two predictors is 33 records (using 0.05 as desired probability level, 0.33 for the moderate anticipated effect size and a statistical power level of 0.8). Consequently, the subsets are divided into 72 samples, each containing 38 records (approximately 1.25 month), so that within each sample the records are at successive points in time.

[28] For each sample, the best distribution (which can be for example Gaussian, Weibull, Gamma, etc.) is fitted and its parameters are estimated, according to the maximum-likelihood method. The best distribution is defined as the distribution for which the KS-goodness-of-fit test is the most significant. When the significance level is below 0.05, no fitting pdf is found. If no fitting pdf is found, a Box-Cox power data transformation [Box and Cox, 1964] is applied as an effort to convert the data to a Gaussian distribution.

For the variable X , the Box-Cox power transformed variable X' is defined as

$$X' = \frac{(X^\lambda - 1)}{\lambda} \text{ for } \lambda \neq 0, X' = \log X \text{ for } \lambda = 0 \quad (3)$$

λ is chosen as the value for which the KS-test of the Gaussian fit to X' is the most significant. If after transforming the data, no fitting pdf is found, the variable is excluded from the further analysis.

[29] The standard pdf used for wind speed is the Weibull pdf [Petersen *et al.*, 1998; Burton *et al.*, 2001; Monahan *et al.*, 2011; Wieringa and Rijkoort, 1983, and others]. The two-parameter Weibull pdf for a random variable U is defined by the scale (A) (which relates to the mean and the range of the pdf) and the shape (k) (which relates to the skewness of the pdf) [Wieringa and Rijkoort, 1983] and is written as

$$p(U, A, k) \equiv \frac{k}{A} \left(\frac{U}{A}\right)^{k-1} \exp \left[-\left(\frac{U}{A}\right)^k \right] \quad (4)$$

[30] The Weibull distributions in Figure 2 fit closely to the observed pdfs for the day-time hub-height (140 m) wind speed, but for the night-time hub-height wind speed, the observed pdf has a broader range and is more skewed than the Weibull pdf (especially during winter). This results in an underestimation of the probability of the relatively strong wind speed during the winter-night. He *et al.* [2010] suggested that this underestimation of the night-time skewness of the pdf is related to a stronger relation between the skewness of the wind speed pdf and the surface buoyancy flux during stable (nightly) stratifications than during unstable (daily) stratifications, especially over rough surfaces. Because Cabauw is situated in open grassland terrain, the Weibull pdf is still adequate in fitting the observed wind speed distribution; however, the fit is better during day than during night.

Figure 2. Frequency distribution plots (full lines), fitted Weibull pdfs (dashed lines) and the estimated scale and shape parameters of observed 140 m wind speeds (in m/s) at Cabauw for SD (green), SN (red), WD (black), and WN (blue).

[31] As seen from Figure 2 the hub-height wind speed characteristics are strongly dependent on the season and the time of the day, with generally stronger wind speed during winter than during summer and stronger during night than during day. To capture the strong diurnal and seasonal variation in wind speed distribution characteristics properly, we perform the statistical downscaling for each time at which the observed Weibull scale and shape parameters are reaching their minimal and maximal values. For WD, WN, and WN, the extremes in scale and shape occur at 00 or 12 UTC. Only for SD, the Weibull parameters reach their minimum around 9 UTC. During SD the scale and shape parameters at 9 UTC are downscaled using an ERA-Interim training data set of 12 UTC.

3.3. Multiple Linear Regression With Monte Carlo Cross-Validation

[32] The observational, the ERA-Interim, and the ECHAM5 data sets are divided in a calibration and a validation period. The calibration period needs to be long enough to cover different kinds of weather situations, and simultaneously the validation period should be long enough to skillfully test the quality of the model (in different types of situations). Therefore, each of the four subsets are randomly divided in two thirds of calibration and one third of validation period. This corresponds to 48 samples for calibration and 24 samples for validation. The random division in calibration and validation periods, unlike to the chronological division, ensures that both periods do not climatologically differ, reduces autocorrelation within the calibration data set during the linear regression analysis and will more likely result in a calibration period which describes a larger range of data, which will therefore be better in predicting extreme situations. The disadvantage of the random division in calibration and validation samples implies that the two data sets are more dependent on each other due to the autocorrelation within the original time series.

[33] The transfer function of the linear regression is derived from the calibration data set. The function describing n samples of the dependent variable Y (being the observed wind speed scale and shape parameters) with predictors X_1 and X_2 (being the pdf parameters of the reanalysis variables) is

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i \quad (5)$$

with i going from 1 to n and ε being the unexplained variation in the dependent variable. To decide upon which variables are used as predictors, a forward stepwise selection is carried out, using a significance level of 0.05. Because complex models are to be avoided, a predictor is only allowed to the model if it causes the R^2 adjusted (R^2 adj) to increase. R^2 adj is a modification of R^2 which, besides measuring the percentage of explained variance, adjusts for the number of explanatory variables in the model). For one predictor in the model, R^2 adj is equal to the determination coefficient R^2 .

$$R^2 \text{ adj} = 1 - (1 - R^2) \cdot \left(\frac{n - 1}{n - p - 1} \right)$$

where p is the total number of predictors in the linear model (not counting the constant term), and n is the sample size [Kutner *et al.*, 2005].

[34] The assumptions of the multiple linear regression technique are verified using the Durbin-Watson test (for independency of the residuals), the Breusch-Pagan test (for the homoscedasticity) and the Shapiro-Wilk test (for the normality of residuals). The adopted significance level for the p -value of the test-statistics is 0.05. For a more detailed description on the forward stepwise selection methodology and the regression assumptions the reader is referred to *Kutner et al.* [2005].

[35] When the test-statistics indicate that one or more of the assumptions are not met, the data are qualitatively checked by visualizing the residuals in terms of time, time lags, estimated values, observed values, etc. If possible, the data are adjusted (using outlier-treatments, transformations, etc.), and the stepwise regression method is performed again. If not, the “bad” predictor combination is excluded before restarting with the stepwise regression.

[36] The performance of the predictive model is estimated by applying the Monte Carlo cross-validation to the linear regression. In other words the random subsampling in a calibration and a validation period is iterated 50 times. For each random split, the regression model is fit to the calibration data and the predictive accuracy is assessed using the validation data. The results are then averaged over the splits. Averaging the quality of the predictions across the validation sets yields an overall measure of prediction accuracy [*Picard and Cook*, 1984].

3.4. Application of the Downscaling Methodology on Bias-Corrected ECHAM5 Data

[37] The downscaling model is applied for each subset on the distribution of the validation period (900 random selected days covering one third of the total period). The predictors of the model are given by the pdf parameters of the ECHAM5 variables. To optimize the result of the downscaling, the difference between the ECHAM5 pdf parameters and the ERA-Interim pdf parameters is removed before performing the downscaling on ECHAM5. For predictor X , being a pdf parameter of the variable of interest, the bias of ECHAM5 is corrected by

$$X_{bc} = X_{echam} - (X_{echam} - X_{ei})$$

where X_{bc} is the ECHAM5 predictor after bias-correction, X_{echam} is the ECHAM5 predictor before bias-correction and X_{ei} is the ERA-Interim predictor, aggregated to the ECHAM5 grid cell.

[38] The downscaling methodology can be summarized as follows:

- (1) Compare 00/12 UTC GCM output to 00/12 UTC reanalysis output (aggregated onto the GCM grid), to see which GCM variables best capture the statistical behavior of the corresponding reanalysis variables. Select the best variables.
- (2) Use the selected variables in the reanalysis (aggregated onto the GCM grid) during a calibration period as predictors, and the tower-observed hub-height wind speed during the calibration period as predictand, in a MOS technique to develop a multilinear transfer function for the probability distribution parameters. Perform a Monte Carlo cross-validation.
- (3) Bias-correct the selected predictors in the GCM relative to the reanalysis (aggregated onto the GCM grid), using the bias between those two data sets during the validation period.
- (4) Apply the multilinear transfer function to the bias-corrected GCM predictors during a validation period, and compare to tower observations during the validation period.

4. Results

4.1. GCM Variable Evaluation

[39] ECHAM5 represents the ERA-Interim pdf of sea level pressure and those of specific humidity and temperature at 10 m, 140 m, 500 m, and 1 km very well (Figure 3). Near the surface, the performance is slightly better during the day than during the night. Although the distributions at the different levels are well represented, the pdfs of the derived vertical gradients are not good. This is attributed to the fact that the pdf of the gradient is not plotted as the difference between the two pdfs, yet it is the pdf of the difference between two levels in the atmosphere. Hence, having similar pdfs for ECHAM5 and ERA-Interim temperature variables does not necessarily mean that the individual records, and consequently the pdfs of the temperature gradients are similar. Additionally, the bad representation of the nightly vertical gradients can also be attributed to the fact that the boundary layer is shallower. Vertical gradients crossing the top of the boundary layer are found to be more difficult to represent.

[40] ECHAM5 is able to represent the ERA-Interim pdf of the wind speed at each level during winter. During summer, the wind speed at higher levels in the atmosphere (500–1000 m) is skillfully represented, but the near-surface (10–140 m) wind is not. The representation of the near-surface wind is worse for SN than for SD.

4.2. Probability Density Function Parameter Calculation

[41] Sea level pressure, temperature, and specific humidity variables are approximately Gaussian distributed, wind speed is shown to be Weibull distributed, and variables describing gradients of temperature and wind speed are best represented by a Gamma pdf. The variables for which no suitable theoretical pdf is fitted ($p < 0.1$), even not after transforming the data, are SD DT (T1000-T500), SN DT (T140-T10) and WN DT (T140-T10), DT (T500-T10), and DT (T1000-T500). The skewness of these distributions is too strong to be skillfully fitted. Therefore, these variables are not taken into account in the linear regression method.

4.3. Multiple Linear Regression

[42] The regression analysis is performed using the pdf parameters of the hub-height (140 m) wind speed as predictands. It results in a transfer function of 1 or 2 selected predictors for each predictand (Table 2). More than 2 predictors are found to be inappropriate. For those predictands that have transfer functions based on 2 predictors, the transfer function using only one (the best) predictor is also analyzed.

[43] During winter, the ERA-Interim wind speed pdf parameters at hub-height averaged over the size of the ECHAM5 grid box ($\sim 25,000 \text{ km}^2$) are the best predictors

Figure 3. Perkins S_{score} for summer (left) and winter (right) day, (grey) and night (black) comparing ERA-Interim and ECHAM5 pdfs for the atmospheric variables. The variables with S_{scores} lower than 0.7 are considered as being unskillfully represented by ECHAM5 and are therefore not taken into account in the multiple linear regression analysis. Unskillful variables are indicated using “asterisk” and “open circle” for respectively day and night variables.

Table 2. Transfer Functions for the SD, SN, WD, and WN Weibull Scale and Shape Parameters

Predictand			Coef 0	Coef 1	Predictor 1	Coef 2	Predictor 2
summer	day	scale	1.05	0.76	W500 scale		
		shape	0.38	0.82	W500 shape		
	night	scale	0.12	0.64	W500 scale	0.48	DT(T500-T10) scale
		shape	-0.7	0.93	W1000 shape	0.24	DW(W1000-W10) scale
Winter	day	scale	1.03	0.93	W140 scale		
		shape	-0.17	1.13	W140 shape		
	night	scale	1.86	0.92	W140 scale		
		shape	0.04	1.08	W140 shape		

to model the observed Cabauw wind speed pdf parameters at hub-height (Table 2).

[44] During summer, ECHAM5 is not able to represent the ERA-Interim wind speed pdf at hub-height adequately. In contrast to the winter, the large-scale hub-height wind speed pdf parameters cannot be used as predictors. The regression analysis reveals that the large-scale ERA-interim wind speed pdf parameters from higher (skillfully represented) levels (500–1000 m) are skilful predictors for the hub-height wind speed pdf parameters during summer.

[45] The boxplots in Figure 4 show the p -values of the diagnostic tests for the selected models performed on 50 random samplings of the calibration period. The testing on different sets of the calibration period, checks the dependency of the test-statistics on the assumptions of the period considered. The boxplots of the diagnostic tests have a reasonable large range, indicating that the skill is not independent of the period selected. Nevertheless, the 0.05 significance level is in most cases still below the 1.5 interquantile-range of the lower quantile, suggesting that the tested assumptions for the selected models are met at significance level of 0.05. Regarding the SN shape parameter, the data are log-transformed in order to meet the assumption of the identically distributed residuals. Besides this parameter, no modifications are conducted on the pdf parameters.

[46] The power to predict the observed wind speed parameters of the validation period is estimated by R^2 adjusted, which is plotted in Figure 5 for each model (and in case of a two predictor model also for the one (best) predictor model). The significance levels of all the models are <0.001 .

[47] The skill of the models to predict the scale parameter is higher than the skill of the models of the shape parameter, for every subset. The predictor(s) in the regression models of the scale parameter are able to explain approximately 60 to 90% of the variation in observed wind speed scale. The remaining 40 to 10% can be explained by unknown variables or inherent variability. For the shape parameters the explained variance varies between 20 and 60% during summer and between 40 and 80% during winter. Overall, the models of the winter are slightly more skilful than the models of the summer. This is probably the effect of the selection of the predictor. As during the winter ECHAM5 is performing better, more variables (including hub-height wind speed pdf parameters) can be selected as predictor. Even though 36 predictors are checked, only during SN, the addition of a second predictor improves the skill of the one predictor model of the scale parameter effectively (the average R^2 adjusted increases from 59% to 68%).

[48] During convective SD conditions, the atmospheric boundary layer is well mixed up to a height of 1 to 1.5 km.

Figure 4. P -values of the (a) Durbin-Watson, (b) Breusch-Pagan, and (c) Shapiro-Wilk tests on the models for SD, SN, WD, and WN Weibull scale and shape parameters. The test-statistics have been performed on 50 random samplings of the calibration period

Consequently the hub-height wind speed can be skillfully predicted using the 500 m pdf wind speed parameters. On the other hand, during SN conditions, the thermal turbulence is absent, resulting in a much shallower boundary layer (100–300 m) over land [Garratt, 1992]. The skill of the 500 m wind speed pdf parameter as a predictor of the hub-height wind speed pdf parameter is lower during SN than during SD. Another phenomena that might play a role is the presence of a nocturnal low level jet (LLJ), which has shown to be present at 20% of the nights (even more during SN) [Baas *et al.*, 2009]. LLJ's originate from an inertial oscillation, which develops after sunset in a layer decoupled from the surface by stable stratification. These structures within the atmospheric boundary layer have typical vertical scales which are in the order of 10 times smaller than the coarse vertical resolution of the employed models. Such a phenomena would make the performance of the statistical model for the hub-height wind speed much worse during SN than during SD. The downscaling methodology however shows that during SN the prediction can be improved by adding a second predictor in the model. For the hub-height

scale parameter, the second predictor is given by the scale parameter of the pdf of the temperature difference between the 500 and 10 m level and for the hub-height shape parameter, it is given by the scale parameter of the pdf of the wind speed difference between the 1000 and 10 m level.

4.4. Application of the Downscaling Methodology on Bias-Corrected ECHAM5 Data

[49] To investigate the performance of the statistical downscaling methodology in representing the hub-height (140 m) wind speed pdf, the downscaling model is applied on bias-corrected ECHAM5 data. The anomaly of the statistically downscaled hub-height wind speed pdf with the observed pdf (ECHAM5downscale) is compared to the anomaly of the bias-corrected ECHAM5 140 m hub-height wind speed pdf without the use of a downscaling model (ECHAM5direct) (Figure 6). The hub-height wind speed pdf of ECHAM5direct is derived according to the power law using the wind speed of the closest model levels.

[50] During winter, the ECHAM5direct hub-height wind speed pdf is a very good approximation of the observed

Figure 5. R^2 adjusted of the one (black boxplots) and two (grey boxplots) predictor models for SD, SN, WD, and WN Weibull scale and shape parameters. The R^2 adjusted is calculated for the models derived from 50 random samplings of the calibration period (left) and applied on 50 random samplings of the validation period (right).

Figure 6. Anomalies from the observed 140m hub-height wind speed pdf for the statistically downscaled ECHAM5 data (full line) and for the bias-corrected not-downscaled ECHAM5 data (dashed line) for SD, SN, WD, and WN conditions. Note the different y-axis range for SN.

hub-height wind speed pdf. A statistical downscaling of these large-scale ECHAM5 wind speed brings very little improvement. During summer ECHAM5downscale shows a clear improvement compared to ECHAM5direct. Using

skilful predictors from ECHAM5 to predict the hub-height wind speed during summer, has comparable performance to the direct prediction of the hub-height wind speed during winter.

Figure 7. R^2 adjusted of the one (black boxplots) and two (grey boxplots) predictor models for SD, SN, WD, and WN Weibull scale and shape parameters for the wind speed at 140 and 40 m. The R^2 adjusted is calculated for the models derived from 50 random samplings of the calibration period.

5. Discussion: The Predictability of the Wind Speed for Nocturnal Stable Boundary Layer Conditions

[51] To obtain more insight in the effect of the boundary layer development throughout the day and the prevalence of the low level jet on the predictability of the wind speed in the atmospheric boundary layer, the downscaling technique is additionally performed on the 40 m observed wind speeds in Cabauw.

[52] The predictors resulting from this analysis are the same as for the 140 m (hub-height) models, except for the following situations: (i) SN scale shares the same first predictor, but no second predictor; (ii) for WN shape DW (W1000-W500) is a second predictor; (iii) for WD scale and shape respectively W10 scale and W10 shape are predictors.

[53] It is generally known that our contemporary understanding and modeling capability of the stable nocturnal boundary layer regime is quite poor [Storm *et al.* 2008]. Winds at 40 m altitude are normally within the stable stratified nocturnal boundary layer, in contrast to the 140 m winds. A comparison of the skill of the models predicting the 40 m wind speeds and the models predicting the 140 m winds (Figure 7) shows that during day conditions, the 140 m models are significantly better than the 40 m models. This is not the case during the night, especially not during SN. During SN the pdf parameters of the wind speed at levels of 500 and 1000 m are better in predicting the wind speed scale parameter at 40 m than the wind speed scale parameter at 140 m. Apparently the higher atmosphere geostrophic wind speed pdf parameters are better in predicting the (often very calm) wind in the nocturnal shallow stable boundary layer than they are in predicting the wind at the top of this layer. This is probably related to the presence of a LLJ at the top of the boundary layer. However, the addition of a second predictor to the SN 500 m wind speed scale parameter is not capable of

improving the prediction for the 40 m, while a second predictor is significantly improving the performance of the prediction for the 140 m. The second predictor for the 140 m scale parameter is given by the scale parameter of the pdf of the temperature difference between the 500 and 10 m level and for the 140 m shape parameter, it is given by the scale parameter of the pdf of the wind speed difference between the 1000 and 10 m level. This points to a relation between the LLJ on the one hand and the temperature difference between inner and outer boundary layer levels and wind speed on the other hand. This result is comparable with the result of Baas *et al.* [2009], who found that the development of a substantial LLJ is most likely to occur for moderate geostrophic forcing and a high radiative cooling.

6. Conclusion

[54] This paper presents a new statistical approach to downscale the large-scale GCM information to the hub-height wind speed Weibull distribution of very tall turbines (140 m). The statistical downscaling technique includes a GCM variable evaluation prior to the development of the regression models. Only those variables that are skillfully represented by the GCM can be selected as predictors in the models. The models are first calibrated using the pdf parameters of ERA-Interim data averaged over the size of the ECHAM5 grid box ($\sim 25,000 \text{ km}^2$) as the predictors and the pdf parameters of the observed wind speed of Cabauw as the explained variables. Afterward the models are applied to the ECHAM5 data, which are first bias-corrected using the aggregated ERA-Interim data as the reference.

[55] The GCM variable evaluation, based on the comparison of the ECHAM5 pdfs with the pdfs of the aggregated ERA-Interim variables, reveals that ECHAM5 represents the ERA-Interim pdfs of sea level pressure and of specific humidity and temperature at 10 m, 140 m, 500 m, and 1 km very well. However, the representation of

the derived variables such as vertical gradients of atmospheric variables relating the inner and outer atmospheric boundary layer regimes is not good. Near the surface, the performance is slightly better during the day than during the night. ECHAM5 is able to represent the aggregated ERA-Interim pdf of the wind speed at each level during winter. During summer, the wind speed at higher levels in the atmosphere (500–1000 m) is skillfully represented, but the lower level (10–140 m) wind is not. During convective summer-day conditions, in which the lower level wind speeds are very much influenced by sensible heat fluxes in the boundary layer and during shallow stable summer-night boundary layer conditions, in which the lower level wind speeds are decoupled from the wind speed above, the ECHAM5 model has the lowest skill in modeling the lower level (10–140 m) wind speeds.

[56] The regression analysis, calibrated on the aggregated ERA-Interim data and the observed 140 m hub-height wind speed for Cabauw, has a better performance for the Weibull scale parameter than for the shape parameter. Suggesting that it is more easy to predict the mean and the range of the wind speed pdfs than its skewness and standard deviation.

[57] During winter the aggregated ERA-Interim 140 m hub-height wind speed is closely related to the observed 140 m wind speed. It is therefore appropriate to directly use ECHAM5 model output to study the wind-climate at the hub-height during winter, in particular during the winter-day.

[58] During summer, ECHAM5 is not able to represent the aggregated ERA-interim wind speed pdf at 140 m adequately. In contrast to the winter, the large-scale 140 m wind speed from ECHAM5 cannot directly be used to study the 140 m wind climate. The regression analysis reveals that the large-scale wind speed from higher (skillfully represented) levels (500–1000 m) can be used as a predictor for the 140 m hub-height wind speed. The predictive capacity of this statistical model, using the 500 m scale parameter as a predictor during summer-day is high (R is 0.84). Because during summer-day conditions the atmospheric boundary layer is well mixed up to levels higher than 500–1000 m, these levels provide skilful predictors for the lower-level wind speed.

[59] On the other hand during the summer-night, the vertical mixing is absent, leading to a shallow stable boundary layer and possibly the occurrence of LLJ's at the height of 140 m. Consequently the 500 m wind speed scale parameter is less skilful in predicting the pdf scale parameter of the wind speed inside and at the top of the nocturnal boundary layer (R^2 is 0.59 for 140 m and 0.69 for 40 m wind speed). However, adding the scale parameter of the pdf of the temperature difference between the 500 and 10 m level as a second predictor improves the skill (R^2 adjusted) of the 140 m wind speed scale parameter model from 0.59 to 0.70. This shows that the prediction of the 140 m hub-height wind speed during nocturnal stable boundary layers, which has been known to be challenging due to the presence of LLJ's at this height, can be substantially improved by including information on temperature and wind speed at other levels. On the other hand, the skill of the models predicting the lower level (40 m) wind speed during summer-night cannot be improved by adding a second predictor. This results points to a relation between on the

one hand the temperature difference between the inner and outer boundary layer levels and the wind speed and on the other hand the presence of a LLJ.

[60] In future research, the possibility of extending the statistical method to sites different than Cabauw will be investigated. In addition, the method will be used on a GCM ensemble to simulate past wind climate and additionally to make an ensemble prediction of future's wind climate at different heights reached by the turbine blades. This will be of interest for wind power estimation studies, for guidelineing political decisions and for tuning the wind power conversion efficiency to the prevailing and future wind characteristics.

[61] **Acknowledgments.** This research was funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology Flanders (IWT-Flanders). ECMWF is acknowledged for providing operational ECMWF data. Furthermore I gratefully thank Fred Bosveld (KNMI) in providing the Cabauw meteorological observations.

References

- Archer, C. L. and M. Z. Jacobson (2005), Evaluation of global wind power, *J. Geophys. Res.*, 108(D9), 1–20, doi: 10.1029/2002JD002076.
- Baas, P., F. C. Bosveld, H. Klein Baltink and A. A. M. Holtslag (2009), A Climatology of Nocturnal Low-Level Jets at Cabauw, *J. Appl. Meteorol. Climatol.*, 48(8), 1627–1642, doi: 10.1175/2009JAMC1965.1.
- Box, G. E. P. and D. R. Cox, (1964), An Analysis of Transformations, *J.R. Stat. Soc. Ser. B Stat. Methodol.*, 26(2), 211–252.
- Burton, T., N. Jenkins, D. Sharpe and E. Bossanyi (2001), Wind Energy Handbook, Second Edition, John Wiley & Sons.
- Cohen, J., P. Cohen, S. G. West and L. S. Aiken (2003), Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, Erlbaum, Mahwah, New Jersey.
- Curry, C. L., D. van der Kamp and A. H. Monahan (2012), Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. I: Predicting wind speed, *Clim. Dyn.*, 38(7–8), 1281–1299, doi: 10.1007/s00382-011-1173-3.
- de Rooy, W. C. and K. Kok (2004), A Combined Physical–Statistical Approach for the Downscaling of Model Wind Speed, *Weather Forecast.*, 19(3), 485–495, doi: 10.1175/1520-0434(2004)019<0485:ACPAFT>2.0.CO;2.
- Dee, D. P., et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, 137(656), 553–597, doi: 10.1002/qj.828.
- Demuzere, M. and N. P. M. van Lipzig (2010a), A new method to estimate air-quality levels using a synoptic-regression approach. Part I: Present-day O₃ and PM₁₀ analysis, *Atmos. Environ.*, 44(10), 1341–1355, doi: 10.1016/j.atmosenv.2009.06.029.
- Demuzere, M. and N. P. M. van Lipzig (2010b), A new method to estimate air-quality levels using a synoptic-regression approach. Part II: Future O₃ concentrations, *Atmos. Environ.*, 44(10), 1356–1366, doi: 10.1016/j.atmosenv.2009.06.019.
- Demuzere, M., M. Werner, N. P. M. van Lipzig and E. Roeckner (2008), An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns, *Int. J. Climatol.*, 29(12), 1796–1810, doi: 10.1002/joc.1821.
- Donat, M. G., G. C. Leckebusch, J. G. Pinto and U. Ulbrich (2010), European storminess and associated circulation weather types: future changes deduced from a multi-model ensemble of GCM simulations, *Clim. Res.*, 42(1), 27–43, doi: 10.3354/cr00853.
- Garratt, J.R. (1992), The atmospheric boundary layer, Cambridge atmospheric and space science series.
- Glahn, H.R. and D.A. Lowry (1972), The use of model output statistics (MOS) in objective weather forecasting, *J. Appl. Meteor.*, 11, 1203–1211, doi: 10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.
- He, Y., A. H. Monahan, C. G. Jones, A. Dai, S. Biner, D. Caya, and K. Winger (2010), Probability distributions of land surface wind speeds over North America, *J. Geophys. Res.*, 115, D04103, doi: 10.1029/2008J010708.
- Holt, E. and J. Wang (2011), Trends of wind speed at wind turbine height of 80 m over the contiguous United States using the North American Regional Reanalysis (NARR), *J. Appl. Meteor. Climatol.*, 51(12), 2188–2202, doi: http://dx.doi.org/10.1175/JAMC-D-11-0205.1.
- Howard, T. and P. Clark (2007), Correction and downscaling of NWP wind speed forecasts, *Meteorol. Appl.*, 14, 105–116, doi: 10.1002/met.12.

- Kaas, E., T.-S. Li and T. Schmith (1996), Statistical hindcast of wind climatology in the North Atlantic and northwestern European region, *Clim. Res.*, 7, 97–110, doi: 10.3354/cr0007097.
- Kidson, J. W. and C. S. Thompson (1998), A comparison of statistical and model-based downscaling techniques for estimating local climate variations, *J. Clim.*, 11(4), 735–753.
- Klink, K. (2007), Atmospheric circulation effects on wind speed variability at turbine height, *J. Appl. Meteorol. Clim.*, 46, 445–456, doi: 10.1175/JAM2466.1.
- Kutner, M. H., C. J. Nachtsheim, J. Neter and W. Li (2005), *Applied Linear Regression Models* (5th Ed.), McGraw-Hill/Irwin, New York.
- Massey, F. J. (1951), The Kolmogorov-Smirnov Test for Goodness of Fit, *J. Amer. Statist. Assoc.*, 46(253), 68–78.
- McVicar, T. R. and M. L. Roderick (2010), Atmospheric science: Winds of change, *Nature*, 3(11), 47–748, doi:10.1038/ngeo1002.
- Minvielle, M., C. Cassou, R. Bourdallé-Badie, L. Terray and N. Najac (2011), A statistical-dynamical scheme for reconstructing ocean forcing in the Atlantic. Part II: Methodology, validation and application to high-resolution ocean models, *Clim. Dyn.*, 36, 401–417, doi: 10.1007/s00382-010-0761-y.
- Monahan, A. H. (2012), Can We See the Wind? Statistical Downscaling of Historical Sea Surface Winds in the Subarctic Northeast Pacific, *J. Clim.*, 25(5), 1511–1528, doi: 10.1175/2011JCLI4089.1.
- Monahan, A. H., Y. He, N. McFarlane and A. Dai (2011), The Probability Distribution of Land Surface Wind Speeds, *J. Clim.*, 24(15), 3892–3909, doi: 10.1175/2011JCLI4106.1.
- Perkins, S. E., A. J. Pitman, N. J. Holbrook and J. McAneney (2007), Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions, *J. Clim.*, 20(7), 4356–4376, doi: 10.1175/JCLI4253.1.
- Petersen, E. L., N. G. Mortensen, L. Landberg, J. Højstrup and H. P. Frank (1998), Wind power meteorology. Part I., Climate and turbulence, *Wind Energy*, 1(S1), 25–45, doi: 10.1002/(SICI)1099-1824(199809)1:1<2::AID-WE15>3.0.CO;2-Y.
- Picard, R. R. and R. D. Cook (1984), Cross-Validation of Regression Models, *J. Amer. Statist. Assoc.*, 79(387), 575–583.
- Pryor, S. C., J. T. Schoof and R. J. Barthelmie (2005a), Empirical downscaling of wind speed probability distributions, *J. Geophys. Res.*, 110 (D19109), doi: 10.1029/2005JD005899.
- Pryor, S. C., R. J. Barthelmie and E. Kjellström (2005b), Potential climate change impact on wind energy resources in northern Europe: analyses using a regional climate model, *Clim. Dyn.*, 25(7–8), 815–835, doi: 10.1007/s00382-005-0072-x.
- Räisänen, J., U. Hansson, A. Ullerstig, R. Döschner, L. P. Graham, C. Jones, H. E. M. Meier, Samuelsson and U. Willén (2004), European climate in the late twenty-first century: regional simulations with two driving global models and two forcing scenarios, *Clim. Dyn.*, 22(1), 13–31, doi:10.1007/s00382-003-0365-x.
- Sailor, D. J., T. Hu, X. Li and J. N. Rosen (2000), A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change, *Renew. Energy*, 19, 359–378, doi: 10.1016/S0960-1481(99)00056-7.
- Salameh, T., P. Drobinski, M. Vrac and P. Naveau (2009), Statistical downscaling of near-surface wind over complex terrain in southern France, *Meteorol. Atmos. Phys.*, 103(1–4), 253–265, doi: 10.1007/s00703-008-0330-7.
- Schoof, J. T. and S. C. Pryor (2001), Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks, *Int. J. Climatol.*, 21(7), 773–790, doi: 10.1002/joc.655.
- Sloughter, J. M., T. Gneiting and A. E. Raftery (2010), Probabilistic wind speed forecasting using ensembles and Bayesian model averaging, *J. Amer. Stat. Assoc.*, 105, 25–35, doi:10.1198/jasa.2009.ap08615.
- Solman, S. A. and M. N. Neñez (1999), Local estimates of global climate change: A statistical downscaling approach, *Int. J. Climatol.*, 19(8), 835–861, doi: 10.1002/(SICI)1097-0088(19990630)19:8<835::AID-JOC401>3.0.CO;2-E.
- Storm B., J. Dubhia, S. Basu, A. Swift and I. Giammanco (2008), Evaluation of the Weather Research and Forecasting model on forecasting low-level jets: implications for wind energy, *Wind Energy*, 12(1), 81–90, doi: 10.1002/we.288.
- Ulden, A. P. and J. Wieringa (1996), Atmospheric boundary layer research at Cabauw, *Bound.-Lay. Meteorol.*, 78(1–2), 39–69, doi: 10.1007/BF00122486.
- van der Kamp, D., C. Curry and A. H. Monahan (2012), Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. II: Predicting wind components, *Clim. Dyn.*, 38(7–8), 1301–1211, doi:10.1007/s00382-011-1175-1.
- van Ulden, A. P. and G. J. van Oldenborgh (2005), Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for regional climate scenarios: a case study for West-Central Europe, *Atmos. Chem. Phys. Discuss.*, 5(4), 7415–7455, doi:10.5194/acpd-5-7415-2005.
- Vautard, R., J. Cattiaux, P. Yiou, J. Thépaut and P. Ciais (2010), Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness, *Nature*, 3(11), 756–761, doi: 10.1038/ngeo979.
- Verkaik, J. W. and A. A. M. Holtslag (2007), Wind profiles, momentum fluxes and roughness lengths at Cabauw revisited, *Bound.-Lay. Meteorol.*, 122(3), 701–719, doi: 10.1007/s10546-006-9121-1.
- Vermeulen, A. T., A. Hensen, M. E. Poppa, W. C. M. van den Bulk, and P. A. C. Jongejan (2011), Greenhouse gas observations from Cabauw Tall Tower (1992–2010), *Atmos. Meas. Tech.*, 3(5), 617–644, doi: 10.5194/amtd-3-4169-2010.
- Wieringa, J. and P. J. Rijkoort (1983), *Windklimaat van Nederland*, Staatsuitgeverij's Gravenhage, Den Haag.
- Wilby, R. L., S. P. Charles, E. Zorita, B. Timbal, P. Whetton and L. O. Meams (2004), IPCC report, *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods*, published on-line, 27.