WIND ENERGY
Wind Energ. 2007; **10**:471–482
Published online 19 July 2007 in Wiley Interscience
(www.interscience.wiley.com) DOI: 10.1002/we.237

Optimal Combination of Wind Power Forecasts

Henrik Aa. Nielsen*, Torben S. Nielsen and Henrik Madsen, Technical University of Denmark, Informatics and Mathematical Modelling, 2800 Kongens Lyngby, Denmark
Maria J. San Isidro Pindado and Ignacio Marti, CENER, National Renewable Energies Centre, Wind Energy Department, Ciudad de la Innovación, 31621 Sarriguren (Navarra), Spain

Key words:
wind power forecasting; combination of forecasts; combined forecasting; correlation of forecast errors

We consider wind power forecasts based on a number of different meteorological forecasts originating from three different global meteorological models. Wind power forecasts based on these meteorological forecasts have fairly similar performance. However, in the paper, we show that the wind power forecast errors are relatively uncorrelated. For this reason, we can combine the forecasts and obtain a final forecast which performs better than any of the individual forecasts. Optimal weights are found based on the bias of the individual forecasts and the variance–covariance matrix of the individual forecast errors. In the paper, we show that quite significant improvements can be obtained using only a few different meteorological forecasts. Copyright © 2007 John Wiley & Sons, Ltd.

Received 6 November 2006; Revised 6 June 2007; Accepted 11 June 2007

Introduction

When forecasting wind energy production, one is often faced with the possibility of using several different meteorological forecasts and several different methods of transforming the meteorological forecasts into wind power forecasts. Considering all possible combinations of these two methods results in a large number of possible ways in which a wind power forecast can be produced. Combined forecasting supplies weights by which the individual forecasts are multiplied and summed in order to arrive at a final forecast. The concept was originally proposed by Bates and Granger[1] for two individual forecasts. Since then, a large number of procedures have been suggested; see e.g. the article by de Menezes and Taylor[2] for a review and the article by Zou and Yang[3] and the references therein for more recent work.

Linear methods for combined forecasting include the simple average, Bayesian methods and regression-based methods with or without restricted weights. The most appropriate method depends on the criteria considered and can be based on the properties of the forecast errors such as variance, asymmetry and serial correlation.[2] The properties of the individual forecast errors can strongly influence the characteristics of the error of the combination.[2] In this paper, only linear combinations are considered. The criterion used to select the optimal combination is to minimize the mean squared error (MSE) between the combined forecast and the measurements. Following this aim, and given known properties of the individual forecasts, the weights obtained will also minimize the variance and the bias of the errors.

The method of Bates and Granger can easily be extended to more than two forecasts and to account for an eventual bias. For clarity, this is included in the section describing the method below. Tambke *et al*.[4] observed

*Correspondence to: H. Aa. Nielsen, Technical University of Denmark, Informatics and Mathematical Modelling, 2800 Kongens Lyngby, Denmark.
E-mail: han@imm.dtu.dk
Contract/grant sponsor: Danish Public Service Obligation (PSO); contract/grant number: 104922 (FU-4101), 101295 (FU-2101).

minor improvements when combining forecasts of wind speed from the European Centre for Medium-Range Weather Forecasts (ECMWF) and Deutscher Wetterdienst (DWD) for a site in the North Sea. It is not clear what led Tambke *et al*. to use the weights specified in their paper. Sánchez[5] applied combined forecasting within the field of wind energy forecasting and showed that significant improvements can be obtained by combining several forecasts. In this paper, we use the basic method for combining forecasts by Bates and Granger, extended with a bias correction. We then study the circumstances under which improvements can be expected within the field of wind power forecasting.

Data

Data from two wind farms are used in the investigations presented in this paper. One farm, Klim, is located in flat terrain near the shore in north-western Denmark, while the other farm, Alaiz, is located in complex terrain in the north of Spain. In both cases, several wind power forecasts are available; both in terms of meteorological models and in terms of wind power forecast models. More details of the farms and the data are given below.

Data from Klim

The wind farm at Klim in north-western Denmark consists of 35 Vestas V44 turbines each with a capacity of 600 kW, resulting in a total installed capacity of 21 MW. The latitude is 57·06°N and the longitude is 9·15°E. The data available for this location are:

- Forecasts from the power-curve model of WPPT[6,7] when using meteorological forecasts from Deutscher Wetterdienst. These forecasts are called DWD below. The meteorological forecasts are produced using a local model nested within a global model, both run by Deutscher Wetterdienst.
- Forecasts from the power-curve model of WPPT when using meteorological forecasts from the HIRLAM model of the Danish Meteorological Institute. These forecasts are called HIR below. The HIRLAM model is a nested model, where the boundary conditions are obtained from the global model of ECMWF.
- Forecasts from the power-curve model of WPPT when using meteorological forecasts based on MM5 modelling[8] and the global meteorological model GFS (Global Forecast System) from the National Centers for Environmental Prediction (NCEP) in the USA. These forecasts are called MM5(grd) below.
- Forecasts from LocalPred[9,10] when using meteorological forecasts based on MM5 modelling and the global meteorological model GFS from NCEP. These forecasts are called MM5(lp) below.

The common data period is from February to November 2003, and therefore this period is used here. Only the meteorological forecast for which the initialization is 00Z (i.e. 00:00 UTC) is used together with horizons up to 24 h.

Data from Alaiz

The wind farm in Alaiz is situated 15 km south of Pamplona in the Navarra region of Spain in very complex terrain between 910 and 1120 m above sea level. It is a large wind farm with a rated capacity of 33·1 MW provided by 49 Gamesa G47-660 wind turbines and one Lagerwey LW750 turbine. The geographical coordinates are latitude 42·67°N and longitude 1·57°W. The RIX[11] (Ruggedness IndeX) value for this wind farm is 15; see Figure 1 for an illustration of the complexity of the terrain.

For Alaiz, nine different wind power forecasts are available. These are based on three meteorological models: Two of these are from the Spanish Meteorological Institute, these are both HIRLAM models with a spatial resolution of 0·2 and 0·5 degrees, respectively.* The HIRLAM models use the global meteorological model

*Actually the National Meteorological Institute of Spain now runs two new versions of HIRLAM at higher resolution.

DOI: 10.1002/we

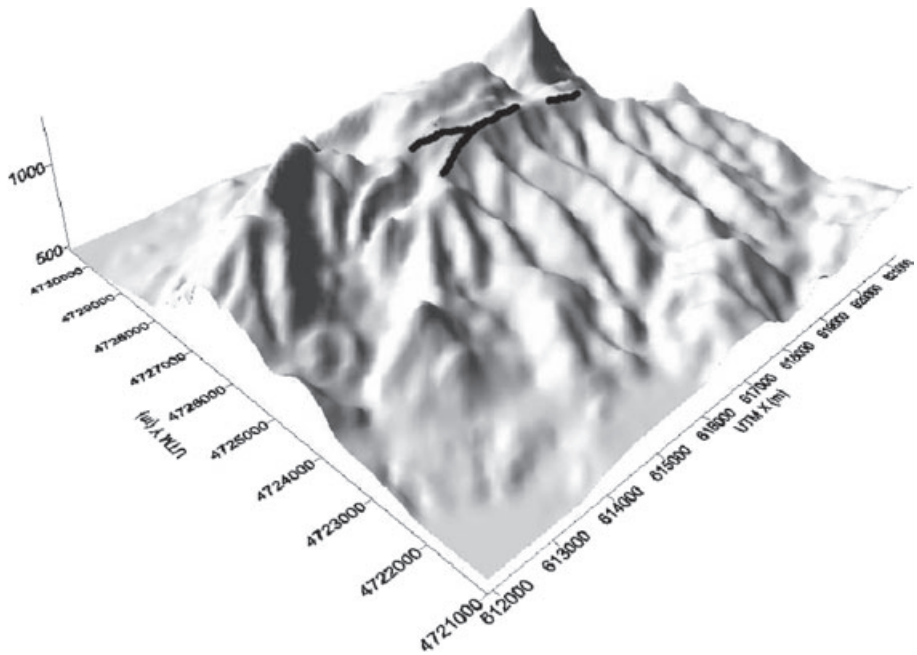

Figure 1. 3D view of the area surrounding the wind farm in Alaiz. The turbine positions are marked in black

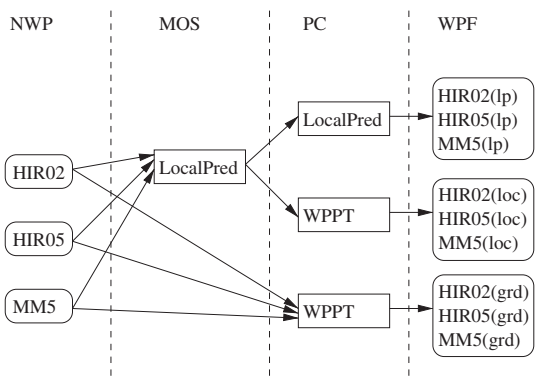

Figure 2. Illustration of the numerical weather prediction (NWP), model output statistics (MOS) and power curve (PC) resulting in nine wind power forecasts (WPF) for Alaiz

from ECMWF as boundary conditions. These two HIRLAM models are called HIR02 and HIR05 below. The remaining meteorological forecast for Alaiz is based on MM5 modelling with nesting of domains to obtain an area around the wind farm with 3 km of resolution. The MM5 model uses the global meteorological model GFS from NCEP as boundary conditions. The meteorological forecasts are filtered through a MOS-step (Model Output Statistics) and a PC-step (Power Curve) to obtain the final wind power forecasts. For the MOS-step, LocalPred is used. Also, forecasts skipping the MOS-step are generated. For the PC-step, either LocalPred or the power-curve model of WPPT is used. The process generating the nine wind power forecasts is illustrated in Figure 2, where the naming convention of the nine wind power forecasts can also be seen. Details about the MOS- and PC-step can be found in the references.[7,9,10] For all wind power forecast methods, data are available from the middle of May 2004 until the middle of October 2005. The forecasts are initiated once a day at 00Z and are available for horizons up to 24 h.

DOI: 10.1002/we

Methods

Combined Forecasting Assuming Known Bias and Variance

In this section, the basic theory of combined forecasting is considered and it is shown how the optimal weights can be found. The theory assumes that the mean vector and the variance–covariance matrix of the individual forecast errors are known. In practice, these quantities need to be estimated. This is considered in the next subsection and the practical results of combined forecasting included in this paper demonstrate the consequence of using estimates instead of the true (unknown) values.

Assume that the individual forecasts $\hat{y}_j$; $j = 1, \ldots, J$ are available and that these are to be combined into one single forecast. If this is done by adjusting for bias μ_j of the individual forecasts and assigning a weight w_j to each, then it is natural to define the combined forecast as

$$\hat{y}_c = \sum_{j=1}^{J} w_j(\hat{y}_j - \mu_j), \tag{1}$$

where the time index and forecast horizon have been excluded in order to simplify the notation. The bias terms can be collected into one single term and without loss of generality the combined forecast can be expressed as

$$\hat{y}_c = \mu_0 + \sum_{j=1}^{J} w_j \hat{y}_j. \tag{2}$$

Given the observation y, the error of the combined forecast $e_c = y - \hat{y}_c$ can then be expressed as

$$e_c = -\mu_0 + \sum_{j=1}^{J} w_j e_j + y\left(1 - \sum_{j=1}^{J} w_j\right), \tag{3}$$

where $e_j = y - \hat{y}_j$ is the error of forecast j. Note that if the weights are chosen so that $\sum_{j=1}^{J} w_j = 1$, then the combined forecast error can be expressed as a bias-corrected weighted sum of the individual forecast errors. This is attractive since it allows us to understand the properties of the combined forecast on the basis of the mean and variance–covariance of the individual forecast errors. Hence, in the remaining part of the paper, the above restriction will be imposed on the weights. The restriction is implemented by expressing w_J as $1 - \sum_{j=1}^{J-1} w_j$ and determining optimal values for w_j; $j = 1, \ldots, J-1$ and for μ_0.

Let $\mathbf{e}$ be a multivariate random variable, arranged as a column vector and containing the errors of the individual forecasts $1, \ldots, J$ and let $\boldsymbol{\mu}$ and $\mathbf{V}$ be the mean vector of $\mathbf{e}$ and variance–covariance matrix of $\mathbf{e}$, respectively. Using the definitions

$$\underset{J\times 1}{\mathbf{u}_J} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad \underset{(J-1)\times J}{\mathbf{A}} = \begin{bmatrix} 1 & 0 & \ldots & 0 & -1 \\ 0 & 1 & \ldots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -1 \end{bmatrix} \quad \underset{(J-1)\times 1}{\mathbf{w}} = \begin{bmatrix} w_1 \\ \vdots \\ w_{J-1} \end{bmatrix} \tag{4}$$

weight J can be expressed as one minus the sum of the first $J-1$ weights and e_c can be written as

$$e_c = -\mu_0 + \mathbf{u}_J^T \mathbf{e} + \mathbf{w}^T \mathbf{A}\mathbf{e}, \tag{5}$$

where $\mathbf{w}$ can be determined without considering the restriction.

Here we will follow common practice and choose $\mathbf{w}$ so that the MSE of the combined forecast error is minimized. This corresponds to minimizing the second moment of e_c:

$$E[e_c^2] = \mu_c^2 + \sigma_c^2, \tag{6}$$

DOI: 10.1002/we

where μ_c and σ_c^2 are the mean and variance of e_c. From equation (5), it follows that

$$\mu_c = -\mu_0 + \left(\mathbf{u}_J^T + \mathbf{w}^T\mathbf{A}\right)\mu, \tag{7}$$

and

$$\sigma_c^2 = \mathbf{u}_J^T\mathbf{V}\mathbf{u}_J + 2\mathbf{u}_J^T\mathbf{V}\mathbf{A}^T\mathbf{w} + \mathbf{w}^T\mathbf{A}\mathbf{V}\mathbf{A}^T\mathbf{w} \tag{8}$$

Naturally, μ_0 only affects μ_c. To find μ_0 and $\mathbf{w}$ minimizing equation (6), we must find the solution to the following J equations:

$$\frac{\partial}{\partial\mu_0}E\left[e_c^2\right] = 0 \quad \text{and} \quad \frac{\partial}{\partial\mathbf{w}}E\left[e_c^2\right] = 0 \tag{9}$$

The first equation results in

$$\mu_0 = \left(\mathbf{u}_J + \mathbf{w}^T\mathbf{A}\right)\mathbf{m} = \sum_{j=1}^{J} w_j\mu_j, \tag{10}$$

whereby, as mentioned in the Introduction, $\mu_c = 0$ and we must therefore select $\mathbf{w}$ so that σ_c^2 is minimized:

$$\frac{\partial\sigma_c^2}{\partial\mathbf{w}} = 2\mathbf{A}\mathbf{V}\mathbf{u}_J + 2\mathbf{A}\mathbf{V}\mathbf{A}^T\mathbf{w} = 0, \tag{11}$$

due to the symmetry of $\mathbf{V}$. Assuming, $\mathbf{A}\mathbf{V}\mathbf{A}^T$ is invertible, the solution to equation (11) can be expressed as

$$\mathbf{w} = -\left(\mathbf{A}\mathbf{V}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{V}\mathbf{u}_J, \tag{12}$$

$w_J = 1 - 1^T\mathbf{w}$, and μ_0 can be found using equation (10). Equation (12) can be used together with equation (8) to find the minimal $E[e_c^2]$, since $\mu_c = 0$ at the solution. If $\mathbf{A}\mathbf{V}\mathbf{A}^T$ is singular, the optimal weights are not uniquely defined, and in this case a generalized inverse can be used. Using a Lagrange-multiplier approach[12], it is possible to realize that all J weights can also be expressed as $\mathbf{V}^{-1}1/(1^T\mathbf{V}^{-1}1)$.

To illustrate the improvement which can be achieved, we now consider the simplest possible case of two unbiased forecasts. The variances of the individual forecast errors are denoted σ_1^2 and σ_2^2 and the correlation between the individual forecast errors is denoted ρ. As noted by Bates and Granger[1], this results in

$$\sigma_c^2 = \frac{\sigma_1^2\sigma_2^2(1-\rho)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, \tag{13}$$

which can also be deduced using equations (8) and (12). To measure the relative performance of the forecasts, we assume that $\sigma_1 < \sigma_2$ and define the relative performance of method 2 compared to method 1 as $r_1 = (\sigma_2 - \sigma_1) / \sigma_2$. With this it is easy to show that the relative improvement of the combined forecast over the best individual forecast, i.e. $r_c = (\sigma_1 - \sigma_c) / \sigma_1$, can be expressed as

$$r_c = 1 - \sqrt{\frac{1-\rho^2}{(1-r_1)^2 - 2\rho(1-r_1) + 1}}, \tag{14}$$

which is defined for $|\rho| < 1$.

Figure 3 (left panel) shows r_c as a function of ρ and r_1 for some realistic ranges of these. It is seen that given two forecasts, with identical performance ($r_1 = 0$), for which the individual forecast errors are correlated with a correlation coefficient of $\rho = 0{\cdot}7$, the combined forecast will perform 8% better than the best individual forecast. If the forecasts do not perform equally well ($r_1 > 1$), the correlation must be lower in order to achieve the same improvement. Note also that improvements can also be obtained for highly correlated forecast errors

DOI: 10.1002/we

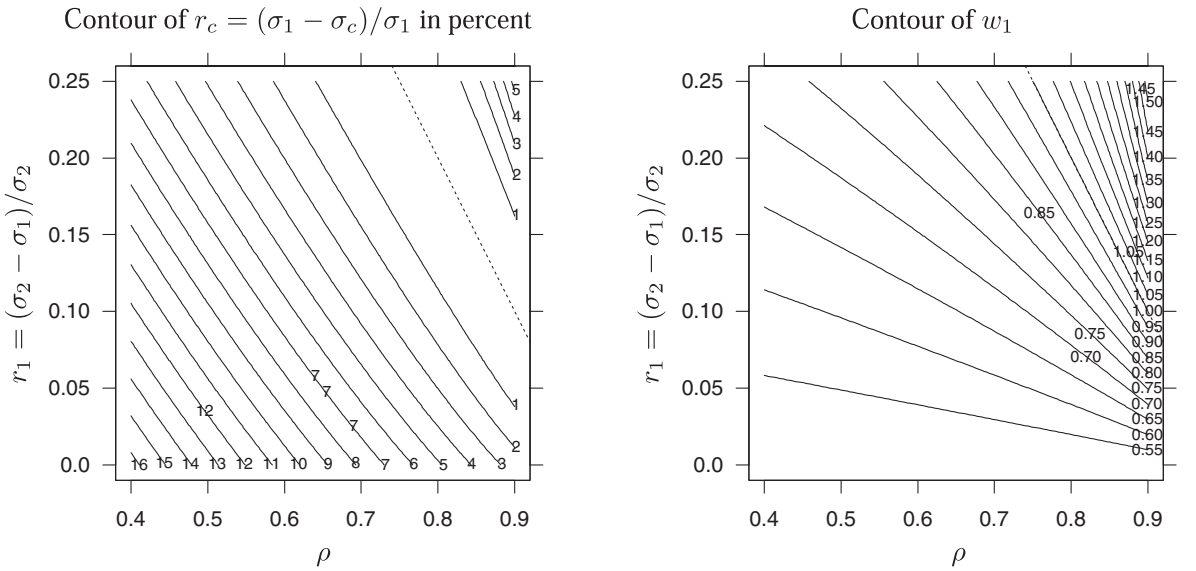

Figure 3. Combination of two unbiased individual forecasts with performance $\sigma_1 < \sigma_2$. The plot in the left panel shows the relative improvement r_c as a function of the correlation ρ and the relative performance of method 1 as compared to method 2. To the left/below the dotted line, both weights are positive. Correspondingly, the plot in the right panel shows the weight on the best forecast

and large differences in performance. However, we doubt that this will be of relevance when combining well-proven individual forecasts.

The weight on the best of the two individual forecasts, i.e. w_1, can be expressed as

$$w_1 = \frac{1-\rho(1-r_1)}{(1-r_1)^2 - 2\rho(1-r_1)+1}\,. \tag{15}$$

Figure 3 (right panel) shows w_1 as a function of ρ and r_1 for the same ranges as above. It is seen that for highly correlated forecast errors and large differences in performance, the improvements mentioned above are obtained when $w_1 > 1$ and hence $w_2 < 0$.

Estimation of Mean and Variance–Covariance

In the previous section, the mean vector $\boldsymbol{\mu}$ and the variance–covariance matrix $\mathbf{V}$ of the errors of the individual forecasts were assumed to be known. In practice, these have to be estimated from observed forecast errors. The method suggested here is to use separate estimates for each forecast horizon h and to allow these estimates to be adaptive over time. Let $\hat{y}_{j,t+h|t}$ be the h-step forecast using the individual method j, generated with the information available at time t.

For the h-step forecast, at time t, the errors $e_{j,t-s|t-s-h} = y_{t-s} - \hat{y}_{j,t-s|t-s-h}$; $s = 0, 1, 2, \ldots$ are known. We denote the estimates of the mean and the variance–covariance available at time t for horizon h by $\hat{\boldsymbol{\mu}}_t^{(h)}$ and $\hat{\mathbf{V}}_t^{(h)}$, respectively. The corresponding μ_0 and weights will be denoted $\hat{\mu}_{0,t}^{(h)}$ and $\hat{w}_{j,t}^{(h)}$; $j = 1, \ldots, J$. These weights are then used to produce the h-step combined forecast using

$$\hat{y}_{c,t+h|t} = \hat{\mu}_{0,t}^{(h)} + \sum_{j=1}^{J} \hat{w}_{j,t}^{(h)} \hat{y}_{j,t+h|t}\,, \tag{16}$$

i.e. at time t, we assume that the most recent estimates of the mean vector and variance–covariance matrix for the individual h-step forecasts are appropriate at time $t + h$ also.

DOI: 10.1002/we

When $\mathbf{e}_{t+1|t+1-h} = [e_{1,t+1|t+1-h} \ldots e_{J,t+1|t+1-h}]^T$ becomes available at time $t+1$, the estimates of mean and variance are updated using the exponential weighting:[13]

$$\hat{\boldsymbol{\mu}}_{t+1}^{(h)} = \lambda \hat{\boldsymbol{\mu}}_{t}^{(h)} + (1-\lambda) e_{t+1|t+1-h} \tag{17}$$

$$\hat{\mathbf{V}}_{t+1}^{(h)} = \lambda \hat{\mathbf{V}}_{t}^{(h)} + (1-\lambda)\left(\mathbf{e}_{t+1|t+1-h} - \hat{\boldsymbol{\mu}}_{t+1}^{(h)}\right)\left(\mathbf{e}_{t+1|t+1-h} - \hat{\boldsymbol{\mu}}_{t+1}^{(h)}\right)^T \tag{18}$$

where $\lambda \in (0, 1]$ is the forgetting factor. Initialization of $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{V}}$ is described below. Note that the estimates defined in this way are adaptive and therefore the method is able to adjust to changes in bias, performance and correlation.

Results

This section describes the results obtained for two wind farms, Klim in north-western Denmark, and Alaiz in the north of Spain. The latter is located in complex terrain, whereas the first is located in very flat terrain close to the North Sea. The meteorological forecasts as well as the power forecasts available differ between the two sites.

Results for Klim

Data from before 1 April 2003 are used to initialize the estimates of the bias of the individual forecasts and the variance–covariance matrix of the individual forecast errors. When evaluating the performance of the forecasts, only data from 1 April 2003 and onwards are used. The individual forecasts contain a few missing values. When comparing the performance, only time points for which none of the forecasts are missing are used in the calculations of the root mean squared (RMS) forecast error.

In order to select the forgetting factor, we consider the two forecasts which are commercially available for Klim, i.e. DWD and HIR. Figure 4 shows the RMS together with the effective number of observations, i.e. $1/(1-\lambda)$. One line is displayed for each horizon. While the minimum is not well defined, it is evident that the effective number of observations should not be below 25. In fact from 50 and onwards, the curves are very

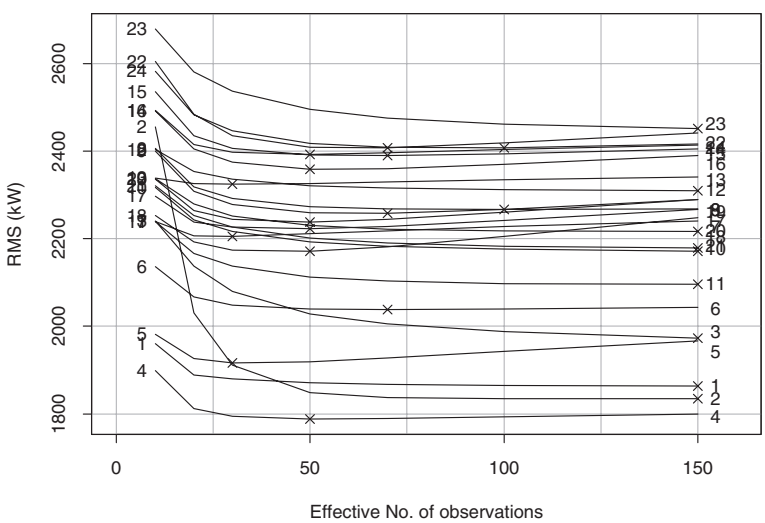

Figure 4. Performance of the combined forecast based on DWD and HIR against the effective number of observations $N_{eff} = 1/(1-\lambda)$, where λ is the forgetting factor. The minimum on each curve is marked by '×' and the numbers to the left and right of each curve are the number of hours since 00Z

DOI: 10.1002/we

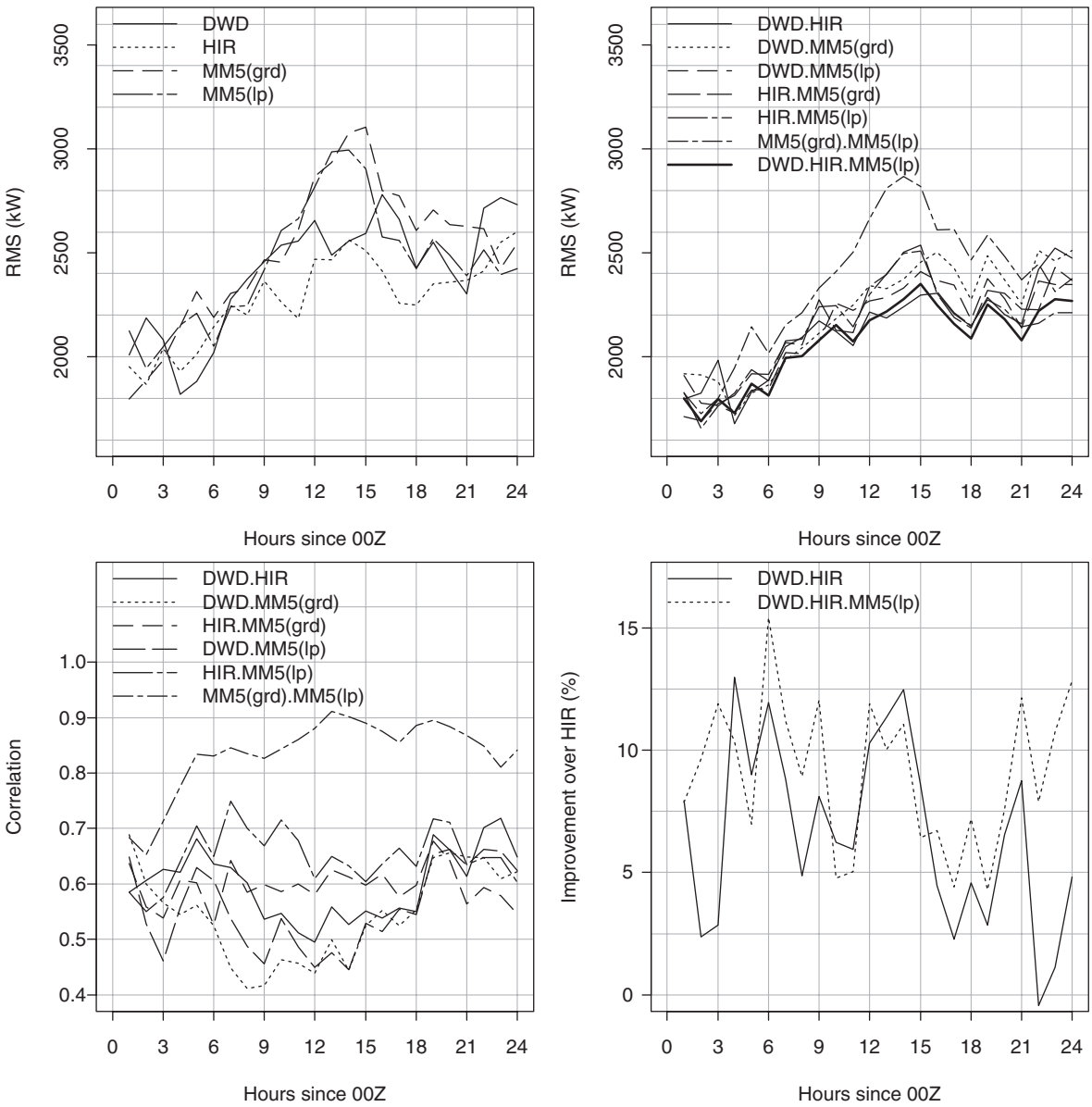

Figure 5. Klim wind farm: Performance of individual (top, left) and combined (top, right) forecasts, together with the pairwise correlations of the individual forecast errors (bottom, left) and relative improvement of two selected combined forecasts over a forecast based on HIRLAM (bottom, right)

flat and for this reason, a forgetting factor corresponding to 50 effective observations is selected. This allows for a relatively quick adaptation to changes in e.g. the correlation structure of the forecast errors.

The two plots on the left in Figure 5 show the performance of the four individual forecasts and all pairwise correlations. It is seen that in general, wind power forecasts based on HIR are superior to the remaining forecasts, especially for horizons from 9 to 21 h. The pairwise correlations are in general between 0·5 and 0·7, but the correlation between forecast errors based on MM5(grd) and MM5(lp) is as high as 0·9. This is presumably because both forecasts are based on the same global meteorological model.

DOI: 10.1002/we

To illustrate the influence of correlation and performance of the individual forecast methods, all possible combinations of two forecasts are considered. Also, a combination of three forecasts with low correlation is considered [DWD, HIR, MM5(lp)]. The result is shown in Figure 5 (top, right). Compared with the performance of the individual forecasts, the improvement is evident. One combined forecast does not perform very well; this is the forecast based on MM5(grd) and MM5(lp) for which the correlation is high. Note that this is also true for horizons between 6 and 9 h where the four individual methods show similar performance and where the correlation is high (approximately 0·85). The combination of DWD and HIR performs well in general. The combination based on the three forecasts with low correlation of the individual forecast errors also performs well and in a more consistent manner across horizons. However, for some horizons, the combination of three forecasts is slightly inferior to the DWD/HIR combination. If the true bias and variance–covariance matrix were available, this could not happen. Hence, the observation is an illustration of the effect of estimation error.

Finally, Figure 5 also shows the relative improvement in RMS of the combined forecasts DWD/HIR and DWD/HIR/MM5(lp) over HIR which overall is the best individual forecast. Improvements as high as 15% are observed, with an overall level of approximately 9% for the combination of three forecasts and slightly less for the combination of two forecasts. Again, it is observed that the combination of three forecasts yields an improvement which is rather consistent across horizons. Furthermore, the improvement obtained for DWD/HIR and the observed performances and correlation of the individual forecasts are consistent with the theoretical limits as depicted in Figure 3.

The combined forecast may be beyond the range of possible power productions. In the results presented, these forecasts are not truncated. For the combination of DWD and HIR from 6 to 60 negative values, out of a total of approximately 300 values, are observed depending on the horizon. The average of these ranges from –0·07 to –0·32 MW. Values larger than the installed capacity are only observed in one case (the value is 21·13 MW).

Results for Alaiz

Data from before 8 July 2004 are used to initialize the estimates of the bias of the individual forecasts and the variance–covariance matrix of the individual forecast errors. When evaluating the performance of the forecasts, only data from 8 July 2004 and onwards are used. The individual forecasts contain a number of missing values. When comparing the performance, only time points for which none of the forecasts are missing are used in the calculations of the RMS forecast error. The forgetting factor used is not optimized for Alaiz; instead, the value selected using data from Klim is used.

To illustrate the properties of combined forecasting, we first focus on combining the three forecasts based on HIR02. The performance is depicted in the upper-left corner of Figure 6. It is seen that the combined forecast performs similarly to the best of the individual forecasts. To understand why, look at the pairwise correlations of the forecast errors of the individual forecasts, which are depicted in the lower-left corner of Figure 6 (dotted lines). It is seen that the correlation between HIR02(grd) and HIR02(lp) is rather low, but since the performance of HIR02(grd) is poor, this does not improve the performance of the combined forecast as compared to the best individual forecast. Similarly, for the two best performing individual forecasts, HIR02(loc) and HIR02(lp), the correlation is high. In conclusion, it is not possible for the combined forecast to improve over the best individual forecast. In fact, it is seen that the combined forecast is often slightly inferior to HIR02(loc); as in the previous section this illustrates the effect of estimation error.

For the three forecasts based on different meteorological models, but the same MOS and power-curve models in each case, the results are depicted in the upper-right corner of Figure 6. It is seen that for many horizons, the combined forecast performs significantly better than any of the individual forecasts. The correlation between the two forecast errors of the two best performing forecasts HIR02(loc) and MM5(loc) is rather low (Figure 6, lower-left corner).

Finally, the lower-right corner of Figure 6 depicts the performance of the best individual forecast [HIR02(loc)], the combined forecast based on the three forecasts using different meteorological models

DOI: 10.1002/we

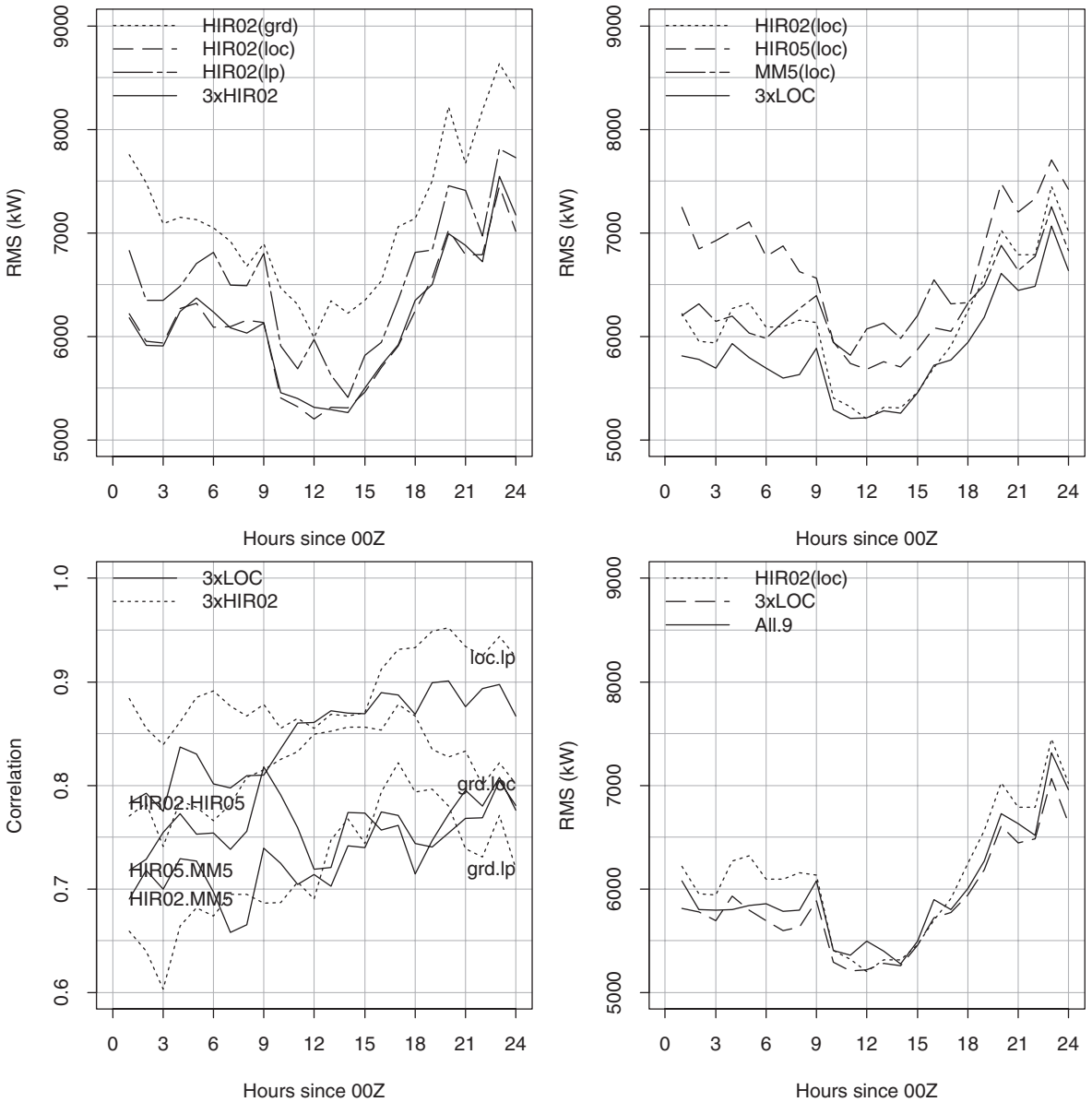

Figure 6. Alaiz wind farm: Performance of the three forecasts based on HIR02 and the performance of the combination of these (top, left). Performance of the three forecasts based on LOC and different MET models and the performance of the combination of these (top, right). Pairwise correlations within the three HIR02 and the three LOC forecasts (bottom, left). Performance of the individual forecast based on HIR02(loc), the combination of the three forecasts based on LOC and the combination of all nine forecasts (bottom, right)

[HIR02(loc), HIR05(loc) and MM5(loc)], and the combined forecast based on all nine available forecasts (cf. the section about data). It is seen that it is beneficial to use combined forecasting, but also that the combination of all nine forecasts is inferior, for some horizons even compared to HIR02(loc). Again, this is an effect of the estimation error. For the combination of the three forecasts mentioned, the improvement is up to approximately 9% over HIR02(loc) with an overall level of approximately 4%. The improvement obtained is comparable with the theoretical level expected from Figure 3, indicating that the forgetting factor is also appropriate for Alaiz.

DOI: 10.1002/we

The overall best combination [HIR02(loc), HIR05(loc) and MM5(loc)] only occasionally produces negative forecasts; for most horizons, the number is below 3 out of approximately 375 values. The maximum of 6 occurs 5 h after 00Z. For four horizons, values larger than the installed capacity are seen in one case for each horizon.

Conclusion and Discussion

In this paper, it is demonstrated that wind power forecasts obtained using different meteorological models may have similar performance even though the correlation of the forecast errors is relatively low. For this reason, a weighted and bias-corrected sum of the individual forecasts can perform better than any of the individual methods. Here, a mean square criterion is considered and the optimal weights depend on the underlying mean and variance–covariance of the individual forecast errors. If the true values of these are known, the combination will always perform at least as well as the best individual forecast and often better if different meteorological models are used.

However, the mean and variance–covariance have to be estimated and therefore the combination can be suboptimal. In general, a few good wind power forecast methods should be used and these should use different meteorological forecasts. As an example of this aspect, combination of three meteorological forecasts filtered trough WPPT, resulted in improvements as high as 15%, with an overall level of 9%, for the wind farm near Klim in Denmark. For the wind farm near Alaiz, the corresponding numbers are 9 and 4%, respectively. However, for Alaiz if one meteorological forecast and three different combinations of MOS and power-curve are used, then no improvement is obtained.

The method used occasionally produces combined forecasts outside the range of possible power productions. In the results presented here, we did not attempt to handle such forecasts separately, but in practice such forecasts should be truncated. This procedure will presumably improve the performance of the combined forecasts marginally.

In order to reduce estimation error, it might be beneficial to restrict the mean and variance–covariance estimates to vary smoothly over the forecast horizons. In this way, neighbour horizons could borrow strength from each other. The most feasible way of implementing this seems to be to formulate an equivalent regression model with the weights as coefficients and then use an adaptive conditional parametric model[14] to estimate the weights as functions of the horizon.

In the paper, adaptive estimates of the mean and the variance–covariance of the individual h-step forecast errors are used to derive weights for combining forecasts h-step ahead and it is shown that this is a very effective approach. It is plausible that the mean and the variance–covariance depend on the overall weather situation. If this is expected to change quickly, as compared to the forgetting factor, it will be beneficial to model the weights directly as functions of some predictable index derived from the meteorological forecasts. It is, however, unclear whether such an index can be identified, see also Section 4·3 in the paper by Tambke *et al*.[4]

In this paper, linear bias-corrected combinations are considered. A more general approach is to use the individual forecasts as inputs to a general non-linear modelling procedure (e.g. a neural network); this is the idea behind *stacking*.[15] It is unknown if such a general modelling approach will result in further improvements. However, due to the relatively high level of uncertainty associated with wind power forecasting, it is somewhat doubtful if such a general approach is to be successful. This is further supported by the observation that regular re-calibration or time-adaptive estimation of the wind power forecast models is often required.

Acknowledgements

This work has been financially supported by the Danish PSO fund under contract 104922 (FU-4101), which is hereby acknowledged. The power data have kindly been provided by ACCIONA, Spain and Elsam Kraft A/S (now part of DONG Energy A/S), Denmark. HIRLAM forecasts for Alaiz (Spain) have been provided by the National Meteorological Institute (INM) of Spain and for Klim (Denmark) by the Danish Meteorological Institute (DMI). GFS forecasts have been obtained from the National Centers for Environmental Prediction

DOI: 10.1002/we

(USA). The forecasts from Deutscher Wetterdienst were obtained commercially via financial support by the Danish PSO fund under contract 101295 (FU-2101).

References

1. Bates JM, Granger CWJ. The combination of forecasts. *Operational Research Quarterly* 1969; **20**: 451–468.
2. de Menezes LM, Taylor JW. Review of guidelines for the use of combined forecasts. *European Journal of Operational Research* 2000; **120**: 190–204.
3. Zou H, Yang Y. Combining time series models for forecasting. *International Journal of Forecasting* 2004; **20**: 69–84.
4. Tambke J, Poppinga C, von Bremen L, Claveri L, Lange M, Focken U, Bye JAT, Wolff J. Advanced forecast systems for the grid integration of 25 GW offshore wind power in Germany. In *Proceedings of the European Wind Energy Conference & Exhibition*, Athens, Greece, 2006. The European Wind Energy Association. [Online]. Available: http://www.ewec2006proceedings.info/.
5. Sánchez I. Adaptive combination of forecast with application to wind energy forecast. *9th International Conference on Probabilistic Methods Applied to Power Systems*, KTH, Stockholm, Sweden, June 2006. Power Engineering Society, IEEE.
6. Nielsen TS, Madsen H, Christensen HS. WPPT—a tool for wind power prediction. *Proceedings of the Wind Power for the 21st Century Conference*, Kassel, 2000.
7. Madsen H, Nielsen HA, Nielsen TS. A tool for predicting the wind power production of off-shore wind plants. In *Proceedings of the Copenhagen Offshore Wind Conference & Exhibition*, Copenhagen, October 2005. Danish Wind Industry Association. [Online]. Available: http://www.windpower.org/en/core.htm.
8. Dudhia J, Gill D, Manning K, Wang W, Bruyere C. PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and Users' Guide. Pennsylvania State University/National Center for Atmospheric Research. [Online]. Available: http://www.mmm.ucar.edu/mm5/. (Accessed June 2002)
9. Marti I, Nielsen TS, Madsen H, Navarro J, Barquero CG. Prediction models in complex terrain. *Proceedings of the European Wind Energy Conference*, Copenhagen, Denmark, 2001.
10. Marti I, San Isidro MJ, Cabezón D, Loureiro Y, Villanueva J, Cantero E, Pérez I. Wind power prediction in complex terrain, from de synoptic scale to the local scale. *Proceedings of The Science of making Torque from Wind*, Delft, The Netherlands, April 2004.
11. Bowen AJ, Mortensen NG. Exploring the limits of WAsP: the wind atlas analysis and application program. In *Proceedings of the European Wind Energy Conference, May 1996*. Göteborg, Sweden, 1996; 584–587.
12. Granger CWJ, Ramanathan R. Improved methods of combining forecasts. *Journal of Forecasting* 1984; **3**: 197–204.
13. Ljung L. *System Identification, Theory for the User*. Prentice-Hall: Englewood Cliffs, NJ, 1987.
14. Nielsen HA, Nielsen TS, Joensen AK, Madsen H, Holst J. Tracking time-varying coefficient-functions. *International Journal of Adaptive Control and Signal Processing* 2000; **14**: 813–828.
15. Wolpert DH. Stacked generalization. *Neural Networks* 1992; **5**: 241–259.

DOI: 10.1002/we