Norwegian Meteorological Institute

METreport

No. 7/2017
ISSN 2387-4201
Meteorology

Do regional weather models contribute to better wind power forecasts?

A few Norwegian case studies

John Bjørnar Bremnes and Gregor Giebel[1]

[1]Technical University of Denmark

Norwegian Meteorological Institute

METreport

Title	Date
Do regional weather models contribute to better wind power forecasts? A few Norwegian case studies	March 24, 2017
Section	**Report no.**
	07/2017
Authors	**Classification**
John Bjørnar Bremnes and Gregor Giebel	● Free ○ Restricted
Client	**Client's reference**
Nordic Top-level Research Initiative	

Abstract

In most operational wind power forecasting systems statistical methods are applied to map wind forecasts from numerical weather prediction (NWP) models into wind power forecasts. NWP models are complex mathematical models of the atmosphere that divide the earth's surface into a grid. The spatial resolution of this grid determines how accurate meteorological processes can be modeled and thereby also limits forecast quality. In this study, two global and four regional operational NWP models with spatial horizontal resolutions ranging from 1 to 32 km were applied to make wind power forecasts up to 66 hours ahead for one offshore and two onshore Norwegian wind farms. A statistical meta-Gaussian method was applied to generate both probabilistic and deterministic wind power forecasts based on the NWP model wind forecasts. The experiments showed that the regional NWP models with higher resolution did not result in better wind power forecasts for these wind farms. In fact, the best wind power forecasts were obtained using one of the coarsest global NWP models.

Keywords

wind power forecasting, numerical weather prediction, spatial resolution

Disciplinary signature	Responsible signature
John Bjørnar Bremnes	Jørn Kristiansen

Abstract

In most operational wind power forecasting systems statistical methods are applied to map wind forecasts from numerical weather prediction (NWP) models into wind power forecasts. NWP models are complex mathematical models of the atmosphere that divide the earth's surface into a grid. The spatial resolution of this grid determines how accurate meteorological processes can be modeled and thereby also limits forecast quality. In this study, two global and four regional operational NWP models with spatial horizontal resolutions ranging from 1 to 32 km were applied to make wind power forecasts up to 66 hours ahead for one offshore and two onshore Norwegian wind farms. A statistical meta-Gaussian method was applied to generate both probabilistic and deterministic wind power forecasts based on the NWP model wind forecasts. The experiments showed that the regional NWP models with higher resolution did not result in better wind power forecasts for these wind farms. In fact, the best wind power forecasts were obtained using one of the coarsest global NWP models.

Contents

1 Introduction

[2] Since the turn of the century there has been a tremendous growth in wind power production worldwide, and in several countries wind power now accounts for a significant share of the total energy production. In contrast to traditional energy sources wind power is produced instantaneously as the wind blows and due to its uncertain and fluctuating nature additional challenges for grid operators and energy producers have emerged. To alleviate this, wind power forecasting has become a valuable tool for planning and decision making.

For forecast horizons beyond three to six hours, say, the best wind power forecasts are generally obtained using statistical methods with input from numerical weather prediction (NWP) models (*Giebel et al.*, 2011). Consequently, the skill of wind power forecasts depends both on the quality of statistical and NWP models. In this article, the attention is primarily paid to the role of the latter in such forecast systems.

NWP models are complex mathematical models of the atmosphere derived from the fundamental laws of fluid dynamics and thermodynamics (*Coiffier*, 2011; *Pielke*, 2013). The models are often defined on a grid in space and time and their resolutions determine how accurate physical processes can be described. Unfortunately, the resolution is limited by the availability of super computing resources, and, thereby, also forecast accuracy. As NWP models require an estimate of the state of the atmosphere at the initialization time, forecast accuracy will also be dependent on the quality of this estimate which further depends on the methodology and the data being assimilated. Due to the atmospheric motions global NWP models covering the entire earth are needed for all weather forecasting purposes. The spatial horizontal resolutions of these currently ranges from about 16 to 30 km[3]. Hence, small scale features in the wind flow within and around wind farms are not resolved with these models. Nested within the global models are regional (limited area) models defined for smaller regions with grid resolutions usually from 1 to 4 km. These rely on data from a global model at their boundaries and are applied for weather forecasting up to two or three days ahead. For special applications NWP models with even higher resolutions are run, but their domains are often quite small and forecast quality can be strongly influenced by the input at the boundaries.

The skill of operational NWP models is continuously being monitored by national me-

[2] The main work behind this report was carried out in 2012.

[3] http://srnwp.met.hu/C_SRNWP_project/Eumetnet_List.html

teorological services and others; e.g. *Richardson et al.* (2012, page 32) showed that the forecast error (standard deviation of error) of 10 meter wind speed forecasts has roughly been reduced by 30% the last 20 years. In the context of wind energy there are several studies related to the resolution of NWP models. *Bedard et al.* (2013) applied Environment Canada's NWP model at 2.5 km horizontal resolution and concluded that the model does not have a sufficiently refined grid to properly represent the meteorological phenomena over a complex Canadian coastal site. *Alessandrini et al.* (2013) compared wind power forecasting skill based on a global and a regional ensemble prediction system and found that the regional with the highest spatial resolution had slightly better skill for a site in southern Italy. *Möhrlen* (2004) studied the performance of deterministic wind power forecasts generated within NWP models run at horizontal resolutions from 1.4 to 30 km and concluded that higher resolution did not reduce the forecast errors at five Irish wind farms. *Dobschinski* (2014) generated wind power forecasts for a large collection of German wind farms using 20 NWP models with horizontal resolutions from about 6 to 22 km. A general trend in favor of the NWP models with the highest resolution was noticed, although exceptions clearly were present. *Draxl et al.* (2014) studied seven planetary boundary layer parametrization schemes in the WRF model and evaluated wind forecasts at various heights at a Danish coastal site. Which parametrization scheme was best, depended on the meteorological condition. Concerning data assimilation there has to our knowledge not been any study on its direct impact on wind power forecasting skill. For regional NWP models it would be possible to compare wind power forecasts based on NWP models with and without data assimilation, while for global models only variation in the data assimilation method and input data would be possible as all global models somehow need a data assimilation system. In summary, the literature on NWP models' impact on wind power forecasting skill is somewhat sparse. Of particular interest would be more studies involving high resolution NWP modeling for various types of terrain around the world. The present article addresses the latter, but is limited to a few Norwegian wind farms.

There are several ways to make wind power forecasts. One alternative would be to just apply given wind power curves to NWP wind speed forecasts, but that approach would be highly sensitive to possible biases in the wind speed. Hence, the underlying predictive information in NWP models is not fully employed and the outcome of an inter-comparison of NWP models using this strategy would to some degree be random. Instead in this study probabilistic wind power forecasts are made using a statistical method with separate input from six operational NWP models, two global and four regional, with spatial horizontal resolutions ranging from 1 to 32 km. From the probabilistic forecasts deterministic point forecasts are also derived and the evaluation of the forecasts is carried out for both types. Our main objective is to assess whether high resolution regional NWP models improve wind power forecasts. It should be noted that from the results it is not possible to attribute differences in quality to certain components or features of the NWP models – at least not with certainty. Our study resembles that of *Dobschinski* (2014), though with several differences. First, the orography of the three coastal Norwegian wind farms under study are quite different from the German. Second, NWP models of finer spatial resolution are considered. Third, both deterministic and probabilistic forecasts are made and evaluated, and, fourth, forecasts are made directly for the wind farms and for each wind turbine before aggregation to wind farm level. The latter may be in favor high resolution models that to some extent are capable to model variations in wind within wind farms.

The article is organized as follows. In the next section the wind farms, measurement data and NWP models are presented. Section 3 describes very briefly the statistical meta-Gaussian method applied and the setup of the experiments. The forecast validation approach is described in section 4. Section 5 contains the results, while discussions and conclusions are in the final section.

2 Data

2.1 Wind farm measurement data

Data from three wind farms along the Norwegian coastline were considered in this study:

- the single offshore floating wind turbine HyWind about 10 km southwest of Karmøy. The turbine (2.3 MW) was installed in autumn 2009 and is primarily dedicated for research and development.

- the Hitra wind farm with 24 turbines and a total capacity of 55.2 MW. The wind farm is located on Eidsfjellet, a hill about 300 m above sea level on the island of Hitra.

- the Smøla wind farm with 68 turbines and a total capacity of 150.4 MW. The wind farm is located in flat and open terrain about 10 to 40 m above sea level on the island of Smøla.

For all wind turbines hourly energy production data were made available and to some extent also each turbine's availability. In addition wind speed measurements at a central turbine of each wind farm were available. The production data was manually examined and suspicious cases were removed before further use.

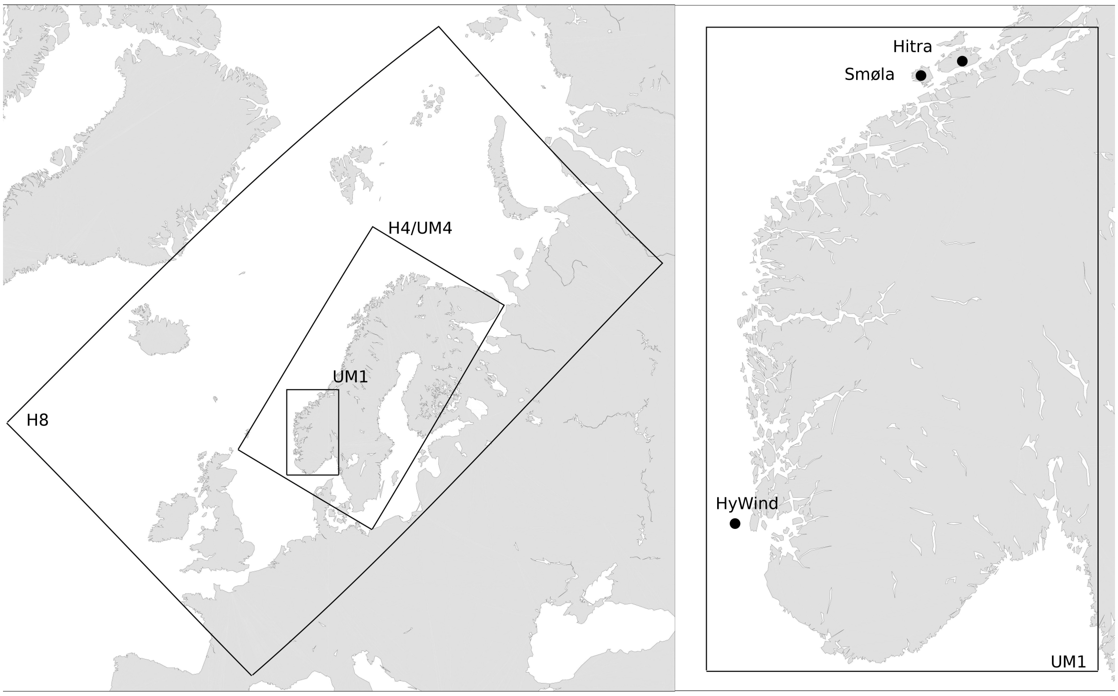

Figure 1: The domains of the NWP models (EC16 and EC32 are global) and the locations of the wind farms Hitra, Smøla and HyWind.

2.2 Wind speed measurements at synoptic stations

For wind speed validation 10-minute average wind speed measurements at 10 meter height were also made available at 31 synoptic stations along the Norwegian coastline. The data covered a period of eight months from February to September 2011.

2.3 Numerical weather prediction models

Table 1: List of numerical weather prediction models with information about model version number, data-assimilation method, approximate spatial horizontal resolution, model applied at the boundary of the domain, and lead times.

Model	Model system	Version	Data assimilation	Spatial resolution	Boundary model	Lead times
UM1	Unified Model	7.3	No	1 km	UM4	+3, +6, ..., +21h
UM4	Unified Model	7.3	No	4 km	H8	+3, +6, ..., +66h
H4	HIRLAM	7.1.3	No	4 km	H8	+3, +6, ..., +66h
H8	HIRLAM	7.1.3	3D-VAR	8 km	EC16	+3, +6, ..., +66h
EC16	ECMWF IFS	36r1 - 37r2	4D-VAR	16 km	-	+3, +6, ..., +66h
EC32	ECMWF IFS	36r1 - 37r2	4D-VAR	32 km	-	+6, +12, ..., +66h

Six operational NWP models were considered in the study and wind forecasts from these were bi-linearly interpolated to the location of each wind turbine. The details of the NWP models are given in table 1. All forecasts were initiated at 00 UTC. The UM1, UM4, H4 and H8 models were all run at the Norwegian Meteorological Institute as part of the operational weather forecasting service, while the EC16 and EC32 models were run at the European Center for Medium Range Weather Forecasting (ECMWF). The EC32 is the control run of their ensemble prediction system. Only forecasts of wind speed and wind direction were applied. In order to make them consistent with the hourly measurement data, hourly mean wind speed forecasts were computed by averaging the wind speed at the start and end of each hourly period. For EC16 and EC32 hourly temporal resolution was not available in the raw data, consequently the forecasts at the start of the hour were obtained by linear interpolation in time. Hourly averaging was not applied to wind direction. Instead the direction at the end of the hour were applied directly. Further, wind forecasts at hub heights were not available for all models. It was therefore decided to use wind forecasts at 10 meter above ground for all models. A sensitivity analysis to the choice is reported in section 6 of the appendix where wind forecasts at selected model levels from the UM1 model were used to generate wind power forecasts at a turbine in the wind farm at Smøla.

Table 2 shows the time period and number of data cases for each wind farm after merging the NWP forecasts and the energy production data. For HyWind wind forecasts were only available from the EC32, EC16, and H4 models mainly due to the extended data period for this turbine. The HyWind turbine was for several parts of the period used for experimental projects and these data were consequently excluded. In order to ensure a fair inter-comparison in the experiments to follow it was required that for each wind farm exactly the same cases in time were available for all NWP models.

Table 2: List of wind farms, data period and average number of cases for each lead time

	Period		
Wind farm	Start	End	No. cases
HyWind	2010/01/10	2011/12/05	325
Hitra	2011/01/15	2011/10/02	260
Smøla	2011/02/25	2011/10/02	220

For validation of wind speed at the 31 synoptic stations the EC16, H8, H4, and UM4 models were also bi-linearly interpolated to the these measurement stations. No averaging in time was made.

3 Methodology

3.1 Statistical forecasting method

During the last decade several statistical methods have been developed for making probabilistic forecasts of wind power based on input from numerical weather prediction models, e.g. *Bremnes* (2004); *Nielsen et al.* (2006); *Pinson* (2012); *Bessa et al.* (2012); *Jeon and Taylor* (2012). In this article a meta-Gaussian method which provides forecasts in terms of probability distributions was applied. The basic idea of the method is to transform all variables to a joint multivariate Gaussian distribution and derive the conditional distribution of wind power from this on the transformed scale. Then the conditional distribution is transformed back to the original scale and applied as wind power forecast. The method was originally developed for precipitation forecasting, but has also been applied for operational wind power forecasting for 7-8 years at the Norwegian Meteorological Institute and in other wind power forecasting studies. Our experience is that the forecast quality is slightly worse compared to local quantile regression. Further details of the meta-Gaussian method are given in the appendix.

3.2 Wind power forecasting experiments

There are mainly two strategies for making forecasts for the total energy production of a wind farm. Depending on the data at hand, the total energy production can either be forecast directly or by aggregating forecasts for each turbine. Both were applied here as described below. Since only 10 meter wind forecasts were available for all models, a brief sensitivity study comparing wind forecasts at different heights and their impact on wind power forecasts skill was also carried out. The setup of this is described at the end of the section, while the results are relegated to the appendix.

3.2.1 Direct forecasting approach

In order to forecast the total energy production of a given wind farm directly, the wind speed forecasts were first averaged over all turbines, while the wind direction forecasts were taken at a central turbine. The wind speed averaging had insignificant effect on the wind speed forecasts for the coarse resolution models, but for the UM1 model who had several grid points within the Hitra and Smøla wind farms an average seemed most appropriate. Based on these data separate meta-Gaussian models were fitted for each wind farm, NWP model and lead time. The statistical meta-Gaussian models were trained using the 60 last cases in a rolling prediction framework. That is, for each prediction to be made data for the preceding 60 days/cases were used for training.

3.2.2 Forecast aggregation approach

Within a wind farm turbulence and complex wake effects may affect turbines differently. Thus, there may be systematic heterogeneities that can not be captured by modeling the total energy production of a wind farm directly. In large wind farms high resolution NWP models may to some extent also be able to forecast spatial variability of wind speed; for example, within the Smøla wind farm the UM1 model had approximately 18 grid points. Separate meta-Gaussian models were consequently set up for each wind turbine, NWP model and lead time. As for the first study, the statistical models were fitted using the 60 last cases for training in a rolling prediction framework. Before validation, the predicted 50 percentiles were computed and added up for each wind farm. Probabilistic forecasts were not made for the wind farm totals with this approach due to complexities in adding up interdependent probabilistic forecasts from turbines.

3.2.3 Wind power forecast skill versus vertical wind level

For wind power forecasting it would be sensible to apply wind forecasts at hub height, but these are not necessarily available and one would often have to rely on wind forecasts at 10 meter above ground. This was also the case here, except for the UM1 model where wind forecasts also were available for all model levels (20 in total). A selection of these and the 10 meter wind were employed to investigate how skill of wind power forecasts depended on the height of the wind forecast. Data for a 4-month period, July to October 2011, at a central turbine of the Smøla wind farm were applied in a two-fold cross-validation procedure to make wind power predictions for the whole period. The predictions were made using the meta-Gaussian approach with wind speed and direction forecasts separately from each level as input.

4 Validation approach

Since the main objective of the study was to compare the overall skill of various forecast models the focus was only on summarizing scores. The most common metric for forecasts in terms of probability distributions is the continuous ranked probability score (CRPS) which quantifies the difference between the cumulative forecast distribution and the measurement (*Matheson and Winkler*, 1976; *Hersbach*, 2000). The CRPS was calculated by means of simulation using the identity established in *Gneiting and Raftery* (2007) and averaged over the evaluation period. The CRPS is negatively oriented, that is, the lower the better. From the probabilistic wind power forecasts the 50 percentile (median) was derived and applied as deterministic point forecasts. The 50 percentile is optimal with respect to absolute error and, consequently, the mean absolute error (MAE) was the obvious measure for these forecasts. Prior to validation the power production forecasts were normalized by the total power production capacity.

The skill of the NWP wind speed forecasts was also evaluated using common statistics like the mean error (ME), mean absolute error, standard deviation of error (SDE), and the ratio between the standard deviations of forecasts and measurements (SDR). The latter was included to measure forecast variability compared to the actual variability. From a physical modeling perspective its optimum value is 1. Since the statistical method will account for simple biases, SDE was possibly the most interesting metric as it measures the error without the bias.

5 Results and discussion

In this section the outcome of the main wind power forecasting experiments is described followed by a brief evaluation of the NWP wind speed forecasts at the end of the section.

5.1 Direct forecasting approach

The results of the direct forecasting approach are summarized in figure 2 using CRPS and MAE of the 50 percentile. As expected, the forecast quality decreased with increasing lead time for all models. Further, the ranking of the models was more or less the same for both scores and each site. Overall the global models EC16 and EC32 with the coarsest spatial horizontal resolutions clearly had the best performance except for the first few hours. The scores were further investigated for pairs of the models for Hitra and Smøla by averaging the scores over common lead times. These revealed that UM4 was 1.1 % and 3.2 % better than UM1 in terms of MAE at Hitra and Smøla, respectively. The numbers for H8 versus H4 were 1.1 % and -0.2 %, while for EC32 versus EC16 they were 0.9 % and 1.0 %. The figures were similar for the CRPS. Thus, the results seemed to be in favor of the models with the coarsest resolutions. These have in general smoother wind fields both in time and space than those with higher resolution. A possible explanation can be that the spatial and temporal averaging of wind speed before it is applied in statistical methods has a significant impact on the wind power forecasting quality. It should also be added that there are minor differences in the model configurations within the model groups, especially for the UM and Hirlam models, that also may have affected the results.

5.2 Forecast aggregation approach

The MAE scores for the forecast aggregation approach are given in figure 3; the results for the single turbine at HyWind are omitted as they were the same as in the previous subsection. The mean absolute errors showed more or less the same characteristics as in the former section with EC16 and EC32 producing the best scores. It can also be noticed that this approach gave better scores than modeling the energy production of wind farms directly. The latter is likely due to that with this approach the relations between NWP models and power production data are allowed to vary between turbines. Results for pairs of models were also evaluated for this approach. UM4 was 0.1 % and 3.6 % better than UM1 at Hitra and Smøla, respectively. The corresponding figures for H8 versus H4 were

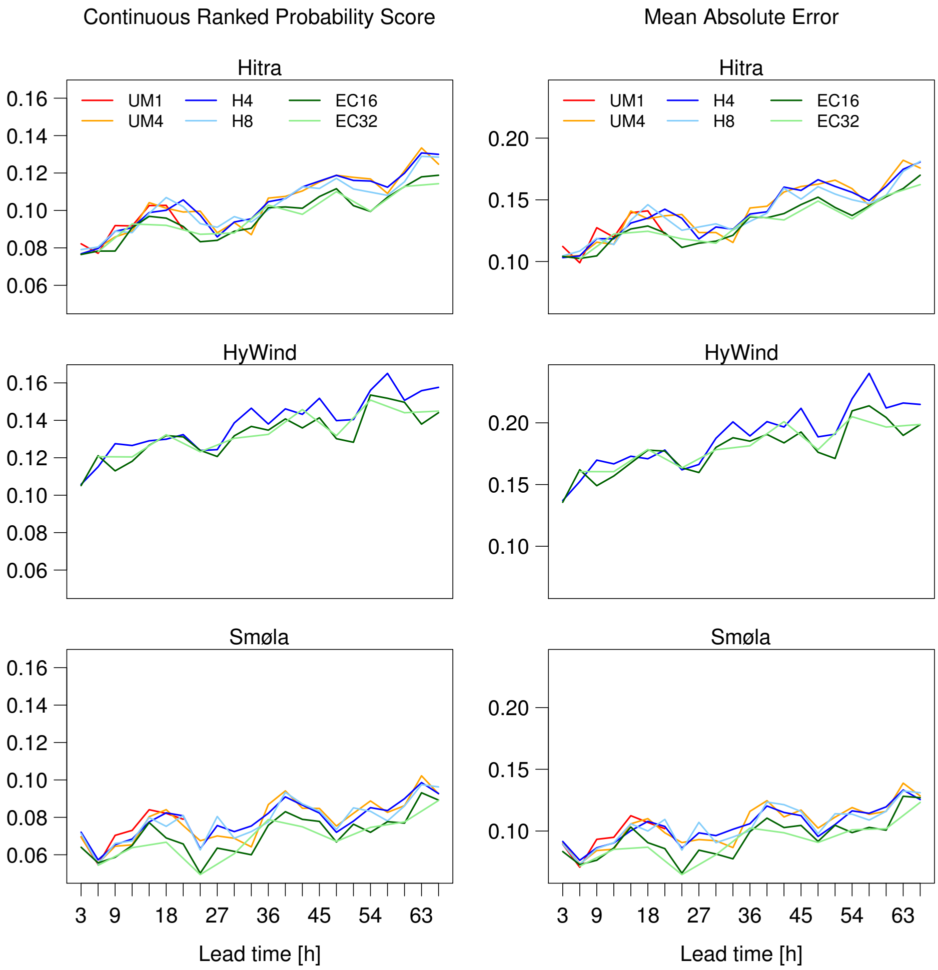

Figure 2: Skill of wind energy forecasts in terms of the continuous ranked probability score (left) and the mean absolute error of the 50 percentile (right) by forecasting the energy production directly.

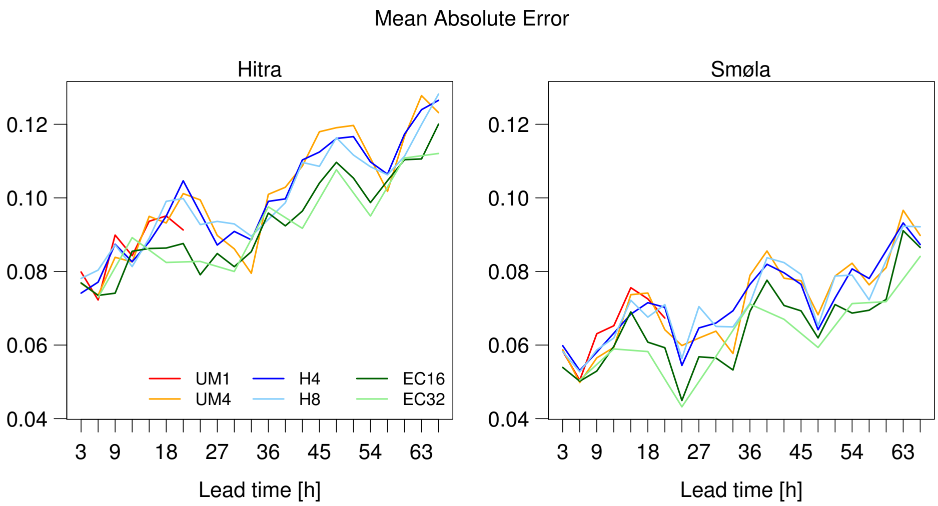

Figure 3: Skill of wind power forecasts in terms of the mean absolute error of the 50 percentile by forecasting the energy production for each turbine first and then aggregating to wind farm level.

0.6 % and -0.2 %, while for EC32 versus they were 1.4 % and 1.3 %. In summary, these results are very similar to those of the direct forecasting approach.

5.3 Validation of wind speed forecasts

In order to investigate the underlying causes of the results in the previous two subsections, a validation of the wind forecasts was carried out. For Hitra and Smøla hourly averaged 10 meter wind speed forecasts at a central turbine were compared to 10-minute average wind speed measurements at the nacelle and similarly for HyWind except that the measurements were 20-minute averages. For the former two wind farms measurements were only available until 1 July 2011, resulting in 167 and 126 cases on average for each lead time at the Hitra and Smøla wind farms, respectively, while for HyWind the period was as before (325 cases). The mean error, standard deviation of error, mean absolute error, and the ratio between the standard deviations of forecast and measurement were calculated and the outcome is reported in table 3.

At Hitra all NWP models underestimated the wind speed as expected due to height differences in the NWP wind and measurement. EC16 and EC32 were least biased and had the lowest standard deviation of errors (SDE). Further, the variability in wind speed was too low for all models, mainly due to the lack of strong wind speeds. The latter was

Table 3: Validation of wind speed forecasts at Hitra, Smøla and HyWind with regard to mean error (ME), standard deviation of error (SDE), mean absolute error (MAE) and ratio between standard deviations of forecast and measurement (SDR). The statistics are averaged over the lead times +6, +12, ..., +66 hours.

Hitra

	ME	SDE	MAE	SDR
UM4	-3.79	3.43	4.04	0.55
H4	-3.79	3.51	4.05	0.53
H8	-3.50	3.39	3.82	0.56
EC16	-2.63	3.13	3.19	0.68
EC32	-2.24	3.25	3.01	0.62

Smøla

	ME	SDE	MAE	SDR
UM4	-2.22	2.71	2.70	0.72
H4	-2.50	2.83	2.88	0.64
H8	-1.75	2.70	2.43	0.76
EC16	-0.14	2.61	1.95	0.94
EC32	-0.36	2.61	1.96	0.89

HyWind

	ME	SDE	MAE	SDR
H4	-0.52	2.67	2.13	0.83
EC16	-0.81	2.62	2.16	0.79
EC32	-0.78	2.65	2.17	0.81

least pronounced for the EC16 and EC32 models. At Smøla the biases for EC16 and EC32 were close to zero, but the SDEs were almost the same as for the other models. With respect to variability (SDR) EC16 and EC32 were clearly in better agreement with the measurements than the other models. In summary, the EC16 and EC32 models produced the best wind speed forecasts for Hitra and Smøla. For HyWind there were only minor differences between the NWP models.

The validation results for wind speed forecasting agreed well with that of wind power. Thus, the quality of wind speed forecasts is to some extent indicative of wind power forecasting skill, as expected. Using this, a second validation exercise on wind speed was conducted to explore the potential skill of wind power forecasts at other sites. Wind speed forecasts at 10 meter for the NWP models EC16, H8, H4, and UM4 at 31 locations along the Norwegian coastline were compared to 10-minute average measurements at 10 meter for a period of eight months. The results are summarized in terms of boxplots over the sites, see figure 4. EC16 was superior with respect to SDE and partly also to MAE. In fact, at 22 of the 31 stations EC16 had the best SDE. However, the variation (SDR) of EC16 was generally less than for the other models, which is in contrast to the results at Smøla and Hitra. Despite the latter, it seems reasonable to anticipate that the main results at the three wind farms also could apply for other wind farms along the Norwegian coastline.

6 Concluding remarks

The main objective of the study was to examine whether use of high resolution regional numerical weather prediction models would improve hourly wind power forecasts at wind farm level compared to using global NWP models with coarser resolution. For the three Norwegian wind farms considered here this was not the case. On the contrary, the best forecasts were produced by input from the two global NWP models. It is hard to generalize the results to other sites, but the quality of the global EC16 model's wind speed forecasts was better than for the high resolution models at most of the Norwegian coastline. Further inland, though, the regional NWP models UM and Hirlam generally verify better than EC16 with respect to wind speed (*Bremnes and Homleid*, 2011) and would therefore probably also yield better wind power forecasts. There are, however, no wind farms in these areas.

Wind fields generated by high resolution NWP models are arguably more realistic physically than those from coarser models, especially in complex terrain, but higher fluc-

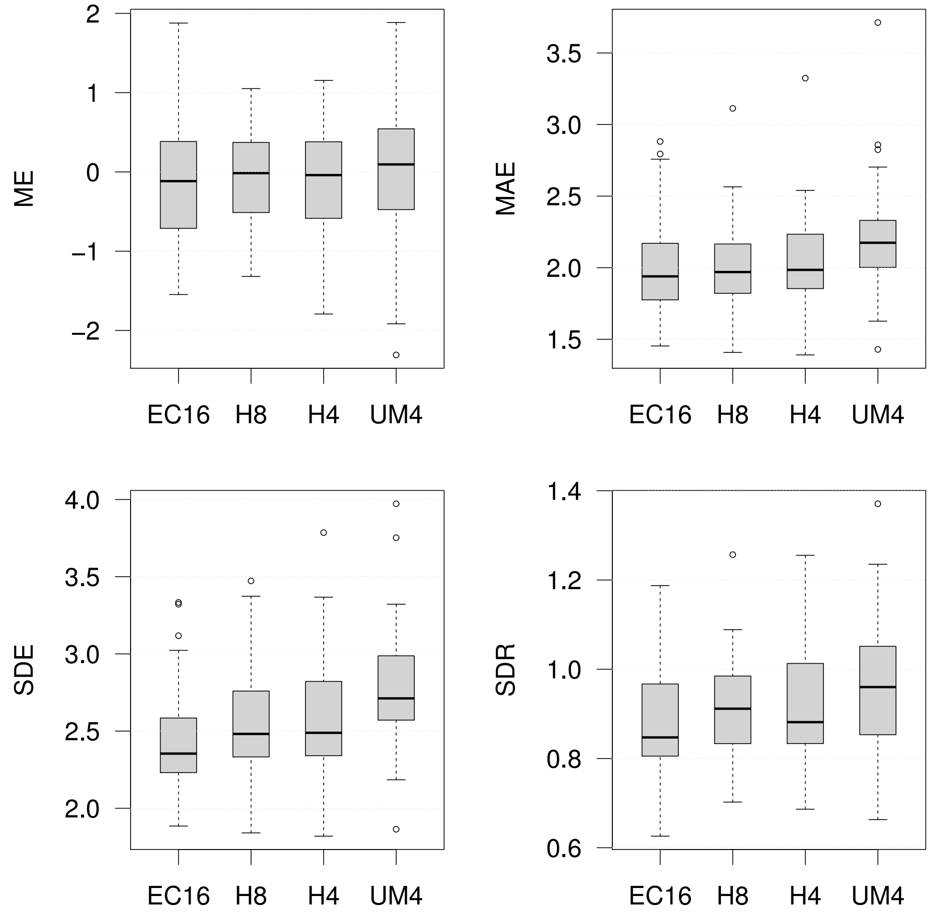

Figure 4: Boxplots of mean error (ME), mean absolute error (MAE), standard deviation of error (SDE) and ratio between standard deviations of forecast and measurement (SDR) for 10-meter wind speed forecasts over 31 Norwegian coastal synop stations. The statistics are averaged over the lead times +6, +12, ..., +66 hours.

tuations also tend to imply higher phase shift errors (*Möhrlen*, 2001; *Jørgensen et al.*, 2002; *Rife and Davis*, 2005). For hourly wind power forecasting the timing and positioning in space are crucial. The dilemma resembles that of convective precipitation forecasting where several studies, e.g. *Roberts and Lean* (2008); *Rossa et al.* (2008), have shown that even though high resolution models generate realistic precipitation patterns, they are often penalized strongly when validating against measurements. By smoothing high resolution forecast data in space and/or time better scores can be obtained (*Roberts and Lean*, 2008). In this study the spatial and temporal averaging was only over the wind farms and one hour, respectively, in addition to the smoothing caused by the bi-linear interpolation to the turbine locations. For future work it may therefore be worthwhile to investigate whether further smoothing could be advantageous to high resolution NWP forecasts before using them for wind power forecasting.

In this article we have only investigated skill of wind power forecasts using one NWP model at a time. Methods for combining forecasts from several NWP models have been proposed by *Nielsen et al.* (2007); *von Bremen* (2007) amongst other. These have also demonstrated improved forecast quality by using several NWP forecasts simultaneously. Hence, even if a high resolution model may not be better than a coarser model, it may contribute to added skill in combination with other models. Lastly, it should be emphasized that only the case of hourly wind power production forecasting has been evaluated. For ramp forecasting, for instance, wind fields from high resolution models may have been more decisive.

Acknowledgements

The work was carried out under the project "Improved forecast of wind, waves and icing" (ICEWIND) funded by the Nordic Top-level Research Initiative, and national and private institutions. The measurement data were kindly provided by Statkraft ASA (Mr. Arild Tanem) and Statoil ASA (Dr. Nenad Keseric).

Appendix

Meta-Gaussian forecasting method

The first and most essential step of the meta-Gaussian method is to transform the data to multivariate Gaussian. In general, it is very difficult to find an appropriate multivariate transformation. Instead one usually resorts to the simpler task of transforming each variable, in our case wind power measurement, wind speed forecast and wind direction forecast, separately to the standard Gaussian distribution and assume that the joint distribution of these is multivariate Gaussian. One option is to estimate the cumulative distribution function of each variable and apply the probability integral transform *Angus* (1994), but here the more direct approach of estimating the relation between quantiles is taken, also known as quantile mapping. Evenly distributed quantiles of the standard Gaussian distribution are then paired with corresponding quantiles of the variable to be transformed. By using these pairs the relation, and hence the transformation, can be estimated. Local linear least square regression was here applied *Cleveland and Devlin* (1988), but other regression approaches could have been applied as well. For the wind power variable cases with zero and nominal production were removed before estimation of the transformations in order to deal with the mixed discrete-continuous nature of this variable.

The next step is to create training data in the Gaussian space. For forecasted wind speed and wind direction the estimated transformations are applied directly, while for the wind power variable special attention is required for zero and nominal production values. Assuming that wind power production is standardized to the interval $[0,1]$) let p_{low} and p_{high} denote the fractions of wind power measurements equal to zero and one, respectively. Then for these cases, the transformed values are defined by $\Phi^{-1}(U)$ where Φ^{-1} is the inverse cumulative distribution function of the standard Gaussian distribution. U is drawn from the uniform distribution on either of the intervals $(0, p_{low}]$ and $[p_{high}, 1)$, depending on whether it is a zero or a one value to be transformed, respectively.

In the Gaussian space the transformed data are now assumed to be multivariate Gaussian

$$\begin{bmatrix} Y^* \\ \mathbf{x}^* \end{bmatrix} \sim \mathscr{N}\left(\begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix}, \begin{bmatrix} \sigma_y^2 & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_x \end{bmatrix} \right) \tag{1}$$

where Y^* denote the wind power and $\mathbf{x}^*$ the vector of wind speed and wind direction forecast after the transformations. The mean and covariance parameters are estimated by their sample statistics. By using one of the many nice properties of the multivariate Gaussian the distribution of wind power given the wind forecast $Y^*|\mathbf{x}^*$ is also Gaussian with mean and variance

$$\begin{aligned}\mu(\mathbf{x}^*) &= \mu_y + \Sigma_{yx}\Sigma_x^{-1}(\mathbf{x}^* - \mu_x) \\ \sigma^2 &= \sigma_y^2 + \Sigma_{yx}\Sigma_x^{-1}\Sigma_{xy}\end{aligned} \tag{2}$$

respectively, see e.g. *Johnson and Wichern* (1992).

Finally, it remains to transform back to the original units. Let h be the estimated function transforming wind power to a standard Gaussian variable. Since it is strictly monotone its inverse exists, but it is only well defined on the interval $[h(0), h(1)]$. For values below $h(0)$ the inverse is set to 0 and above $h(1)$ to 1. It follows that the cumulative distribution of wind power given the predictor variables are

$$F_{Y|\mathbf{x}}(y) = \begin{cases} \Phi\left(\frac{h(y)-\mu(\mathbf{x}^*)}{\sigma^2}\right) & y \in [0,1) \\ 1 & y = 1. \end{cases} \tag{3}$$

where $\mathbf{x}^*$ is the transformation of predictor vector $\mathbf{x}$, and $\mu(\mathbf{x}^*)$ and σ^2 are given by (2). Similarly, the τ quantile is given by $h^{-1}(\mu(\mathbf{x}^*) + \sigma^2\Phi^{-1}(\tau))$.

The success of the meta-Gaussian method clearly relies on the multivariate Gaussian assumption, but to some degree also estimation uncertainty. In order to reduce the latter, and thereby also improving forecast quality, a bootstrap procedure was introduced at two stages. First, the estimation of the transformations was repeated ten times, each time by re-sampling half of the data without replacement. The mean of the transformations was used further on. Second, a similar re-sampling scheme was applied to the transformed training data to estimate the parameters of the multivariate Gaussian distribution.

Wind power forecast skill versus vertical wind level

Wind power forecasts using wind forecasts at various model levels were generated as described in section 3.2.3. In figure 5 the results are shown in terms of the CRPS and MAE of the 50 percentile for selected levels. Not surprisingly, wind forecasts at 500 meter gave the worst wind power forecasts, though for some lead times the scores were actually as good as for most other levels. The relative poor scores for the first 12 to 18 hours may be due to that the initial state is less accurate at higher heights due to the data assimilation process. The fact that the boundaries are quite close to the site may also have some influence. The lower levels gave more similar scores, but overall using wind at 10 meter turned out to give the best wind power forecasts. It should be added that the

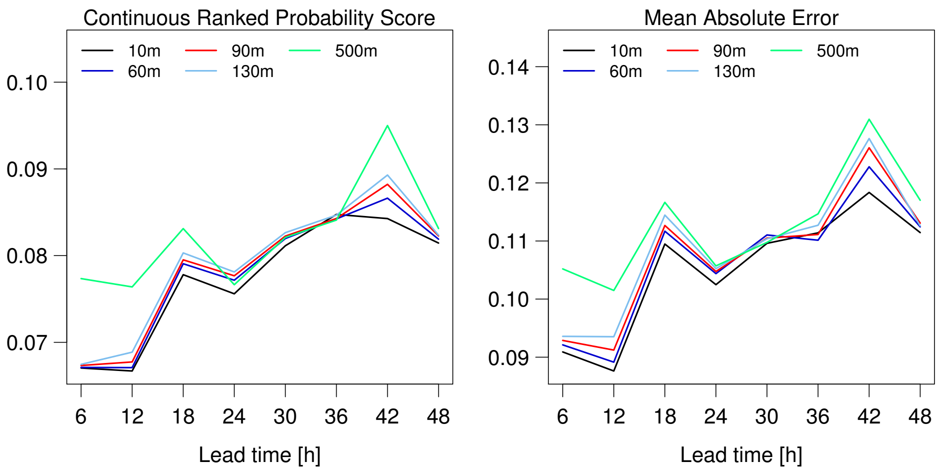

Figure 5: Skill of wind power forecasts at a central turbine of the Smøla wind farm using wind extracted from a few selected model levels of the UM1 model. The continuous ranked probability score is shown to the left and the mean absolute error of the 50 percentile to the right.

10 meter wind is a post-processed field in NWP systems, and that due to the availability of measurement data wind at the surface likely receives more attention than wind at for example hub heights. Further, simple biases in wind speeds are not an issue since these are generally well handled by statistical methods.

The quality of the wind forecasts at this site was briefly evaluated by comparing hourly averaged wind speed forecasts against 10-minute average wind speed measurements at the nacelle. The scores for the various wind levels are given in figure 6 showing the mean error and the mean absolute error. Clearly, the 10 meter wind speed forecasts underestimated the wind speed at hub height (as it should) by around 3 m/s and were inferior to those at 60 meter height. The wind speed forecasts at 60 meter were best both in terms of the mean error and the mean absolute error. In a pure physical based wind power forecasting system the wind at 60 meter would consequently be far better than at 10 meter, but apparently this was not the case in a combined physical and statistical approach.

To conclude it seemed to be of no disadvantage to use wind forecasts at 10 meter above ground at least for this site and model. This is in line with Joensen et. al. *Joensen et al.* (1999) who also concluded that the 10 meter was better for wind power forecasting than higher model levels in a Hirlam model. However, more recently Junk et. al. *Junk et al.* (2012) found that wind forecasts at hub height were favorable. In the article wind forecasts at 10 meter are applied throughout.

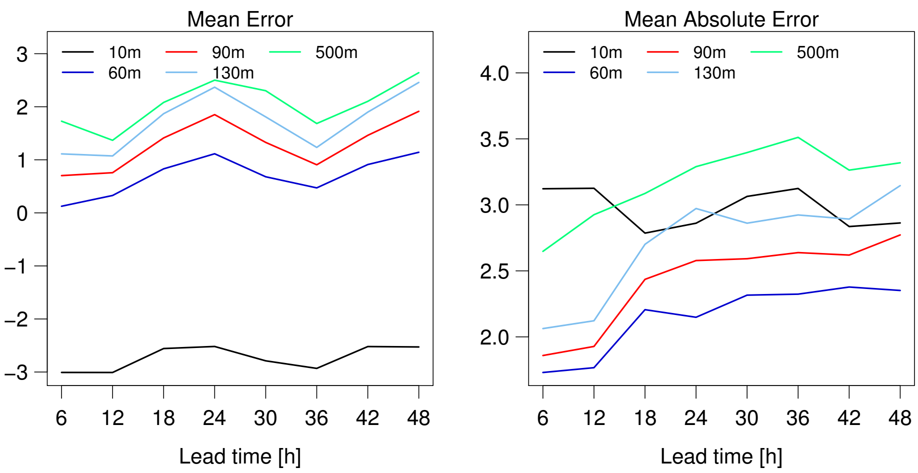

Figure 6: Validation of wind speed forecasts for various levels of the UM1 model against measurements at the nacelle of a central turbine of the Smøla wind farm. The mean error is shown to the left and mean absolute error to the right.

References

Alessandrini, S., S. Sperati, and P. Pinson (2013), A comparison between the ECMWF and COSMO ensemble prediction systems applied to short-term wind power forecasting on real data, *Appl. Energ.*, *107*, 271–280.

Angus, J. E. (1994), The probability integral transform and related results, *SIAM Rev.*, *36*(4), 652–654.

Bedard, J., W. Yu, Y. Gagnon, and C. Masson (2013), Development of a geophysic model output statistics module for improving short-term numerical wind predictions over complex sites, *Wind Energy*, *16*, 1131–1147.

Bessa, R. J., V. Miranda, A. Botterud, Z. Zhou, and J. Wang (2012), Time-adaptive quantile-copula for wind power probabilistic forecasting, *Renewable Energy*, *40*(1), 29–39.

Bremnes, J. B. (2004), Probabilistic wind power forecasts using local quantile regression, *Wind Energy*, *7*(1), 47–54.

Bremnes, J. B., and M. Homleid (2011), Verification of operational Numerical Weather Prediction models. March to May 2011, *Note 7/2011*, Norwegian Meteorological Institute.

Cleveland, W. S., and S. J. Devlin (1988), Locally-weighted regression: An approach to regression analysis by local fitting, *J. Am. Statist. Ass.*, *83*, 590–610.

Coiffier, J. (2011), *Fundamentals of Numerical Weather Prediction*, Cambridge University Press.

Dobschinski, J. (2014), "how good is my forecast? comparability of wind power forecast errors", in *Proc. of 13th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants*, Berlin, Germany.

Draxl, C., A. N. Hahmann, A. Pena, and G. Giebel (2014), Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes, *Wind Energy*, *14*, 39–55.

Giebel, G., R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl (2011), The state-of-the-art in short-term prediction of wind power. a literature overview. 2nd edition., *Tech. rep.*, Risø National Laboratory.

Gneiting, T., and A. E. Raftery (2007), Strictly proper scoring rules, prediction, and estimation, *J. Am. Statist. Ass.*, *102*, 359–378.

Hersbach, H. (2000), Decomposition of the continuous ranked probability score for ensemble prediction systems, *Wea. Forecasting*, *15*, 559–570.

Jeon, J., and J. W. Taylor (2012), Using conditional kernel density estimation for wind power density forecasting, *J. Am. Statist. Ass.*, *107*(497), 66–79.

Joensen, A., G. Giebel, L. Landberg, H. Madsen, and H. A. Nielsen (1999), Model output statistics applied to wind power prediction, in *Proceedings of the European Wind Energy Conference*, Nice, France.

Johnson, R. A., and D. W. Wichern (1992), *Applied multivariate statistical analysis*, 3rd ed., Prentice-Hall International.

Jørgensen, J., C. Möhrlen, and E. McKeogh (2002), A new generation operational on- and offshore numerical prediction system, in *World Wind Energy Conference*, Berlin, Germany.

Junk, C., L. von Bremen, and D. Heinemann (2012), Assessment of probabilistic wind power forecasts for wind farms in Northern Ireland, in *Proc. of EWEA Annual Event*, Copenhagen, Denmark.

Matheson, J. E., and R. L. Winkler (1976), Scoring rules for continuous probability distributions, *Management Science*, *22*, 1087–1096.

Möhrlen, C. (2001), On the benefits of and approaches to wind energy forecasting, in *Irish Wind Energy Association Annual Conference. Towards 500MW.*

Möhrlen, C. (2004), Uncertainty in wind energy forecasting, Ph.D. thesis, University College Cork.

Nielsen, H. A., H. Madsen, and T. S. Nielsen (2006), Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts, *Wind Energy*, *9*(1-2), 95–108.

Nielsen, H. A., T. S. Nielsen, H. Madsen, M. J. S. I. Pindado, and I. Marti (2007), Optimal combination of wind power forecasts, *Wind Energy*, *10*, 471–482.

Pielke, R. A., Sr. (2013), *Mesoscale Meteorological Modeling*, International Geophysics, 3rd ed., Academic Press.

Pinson, P. (2012), Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions, *J. R. Statist. Soc., Ser. C*, *61*, 555–576.

Richardson, D. S., J. Bidlot, L. Ferranti, A. Ghelli, T. Haiden, T. Hewson, M. Janousek, F. Prates, and F. Vitart (2012), Verification statistics and evaluations of ECMWF forecasts in 2011-2012, *Tech. rep.*, ECMWF.

Rife, D. L., and C. A. Davis (2005), Verification of temporal variations in mesoscale numerical wind forecasts, *Mon. Wea. Rev.*, *133*(11), 3368–3381.

Roberts, N. M., and H. W. Lean (2008), Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Mon. Wea. Rev.*, *136*, 78–97.

Rossa, A., P. Nurmi, and E. Ebert (2008), *Precipitation: Advances in Measurement, Estimation and Prediction*, chap. Overview of methods for the verification of quantitative precipitation forecasts, pp. 417–450, Springer Verlag.

von Bremen, L. (2007), Combination of deterministic and probabilistic meteorological models to enhance wind farm power forecasts, *Journal of Physics: Conference Series*, *75*.