

Spectral Verification of a Mesoscale Ensemble

Claire Vincent¹, Caroline Draxl¹, Gregor Giebel¹,
Pierre Pinson², Jess Jørgensen³ and Corinna Möhrlen³

¹Risø National Laboratory for Sustainable Energy, Technical University of Denmark

²Technical University of Denmark, Dpt. of Informatics and Mathematical Modelling

³WEPROG ApS

Corresponding Author email: clav@risoe.dtu.dk

Abstract

An adaptive spectral method is used to verify wind speed forecasts and simulations from members of the Multi-Scheme Ensemble Prediction System (MSEPS), set up for the Horns Rev offshore wind farm in the Danish North Sea. The 75 members of the MSEPS are run on the same model grid with a resolution of 5km. The members differ in their numerical formulation, mainly in the fast reacting atmospheric processes and in their initial conditions.

Temporal scales of 24 minutes to 10 hours are verified separately by using the Hilbert-Huang transform as a filter to extract envelopes of variance over a specific range of frequencies. This means that variability is being verified, but not phase, which may be a more appropriate way of assessing a model's ability to capture mesoscale fluctuations. Using the Hilbert-Huang transform, the forecast wind variability can also be expressed as a time evolving probability density function, which is useful for verification, and potentially as a warning tool for severe wind variability.

Introduction

Managing the variable power supply from wind farms is challenging from both technical and economic perspectives. While power fluctuations from groups of land based turbines may be smoothed by the spatial distribution of the wind farms [5], very large fluctuations in generation from offshore wind farms may occur as a result of a high concentration of offshore turbines within a small geographical area. For example, at the Horns Rev offshore wind farm near the west coast of Denmark, fluctuations in power production have been observed to be significantly larger than for onshore turbines [1].

Good wind and power forecasts have been well recognised as having an important role in planning for contingencies, in ensuring adequate reserve requirements, and in safely and economically managing the grid integration of wind power [6, 8]. Improving the predictability of severe wind fluctuations can, in particular, help Transmission System Operators (TSOs) who need to maintain balance in the electricity grid, and wind farm operators, who need to optimise the economic value of produced wind power. For example, understanding the time scales on which imbalance amongst scheduled units is most likely to occur could be a decision factor in the preallocation of reserve power.

A model which has been successfully used for forecasting wind power is the Multi-Scheme Ensemble Prediction System (MSEPS), run by the Danish company WEPROG, where 75 mesoscale numerical weather prediction (NWP) models with varying condensation schemes, advection schemes and diffusion schemes are used to create a probabilistic prediction of the wind and power [12,13,14]. The benefit of probabilistic forecasts over deterministic forecasts is that TSOs can take uncertainty into account in their decision-making processes, as discussed in [2]. It has also been shown that there are economic incentives to wind power producers including uncertainty information in their market-bidding strategies [17]. Advanced statistical techniques have been applied to map the MSEPS wind forecasts to power forecasts and to re-calibrate the distribution of power forecasts to a realistic probability density function of expected power output [18]. In this work, the ability of the MSEPS to explicitly forecast wind variability on shorter temporal scales than normally considered is explored.

Forecasts from prediction systems such as MSEPS are traditionally provided to end-users with hourly time steps, but the actual integration time step of the model is much less than an hour. Meteorological variables are computed with a dynamical time step of around 30 seconds, and the

impact of various physical processes is incorporated at intervals relevant to the setup of the ensemble member. For this work, MSEPS variables have been saved with a temporal resolution of around 6 minutes, allowing a study of the shortest time scales at which useful information exists in the model. It is not expected that the model will contain variance with periods as short as 6 minutes, since the spatial variance of the model (and therefore the temporal variance, through Taylor's hypothesis [22]) is limited by the grid spacing, in this case 5km.

In this work, we aim to assess the temporal scales on which realistic information exists in the members of the MSEPS, with particular reference to providing warnings about severe wind variability for wind and power forecasting. The variance in the model wind speed time series is analysed using an adaptive spectral method called the Hilbert-Huang transform, which is ideal for this type of study due to its non-parametric decomposition of the data, and its strengths in describing the spectral behaviour of non-stationary time series [23]. By using the Hilbert-Huang transform as a filter, time series of the envelope of the total variability within a particular frequency range can be calculated, which provides both a new verification diagnostic, and an efficient way to extract variability forecasts from the model. The analysis is applied to a special set of 10 month continuous time series of wind simulations, such that the differential effects of the 75 different MSEPS model formulations on the highest resolution variance can be explored.

The quality of the mesoscale variability forecasts is further explored by creating probabilistic forecasts of wind variability, based on 24 hour MSEPS forecasts, and comparing them with observed variability.

The structure of this paper is as follows: first, spectral verification of numerical weather prediction models is discussed, then the model and observations that are used in this study are described. After briefly introducing the methodology of the Hilbert-Huang transform as a verification tool, results are presented both for the scale dependent verification of the MSEPS model, and for the probabilistic forecasts of wind variability.

Spectral verification of mesoscale forecasts

As discussed by [11], many traditional verification scores incorporate errors on different scales of atmospheric motion into a single score. This means that scores such as the root mean square error (RMSE), bias or mean absolute error (MAE) include errors from different physical processes with different length and time scales. Contributing errors may arise from planetary, synoptic or diurnal scales, as well as insufficient model parameterisation of smaller scale physical processes.

For assessing a model's skill in forecasting these smaller scale processes, a filtering process which can isolate a particular range of scales is required. [11] followed this approach in verifying forecasts from the ECMWF ensemble system, where a spatial filter was used to separate planetary, synoptic and sub-synoptic contributions to the forecast error and ensemble dispersiveness. Using this methodology, it was possible to diagnose short range over dispersiveness of the ECMWF ensemble system as a synoptically driven feature of the model. Since the ECMWF ensemble system uses a spectral model, filtering the model fields according to a particular set of spatial scales was achieved by applying total or zonal wavenumber filters. For models which are not formulated in spectral space, a transformation which allows consideration of the contribution of different scale to overall variance is needed. For example, [3] and [4] used a spatial two-dimensional filter to conclusively demonstrate enhanced variance on spatial scales of 250-550 km of a local area model compared with a global analysis.

Although the approach of [4] was effective for assessing the variance in the model over a large spatial area at a particular time, the data in this study is time series data, so we require a temporal filter. For example, a relevant strategy for scale separation in this case could be band pass filtering the time series, such as in [7], where band pass filtering was used in a climatological study of subsynoptic, synoptic, slow synoptic and low frequency contributions to variation in North-Atlantic processes.

However, since the focus here is on mesoscale variability in the planetary boundary layer rather than synoptic scale variability, we are not necessarily interested in precise timing of peaks and troughs within an episode of severe wind variability, but in the phenomenological existence of severe wind variability.

Verifying a band pass filtered forecast against a band pass filtered observation means that a good variability forecast with a small timing error will be strongly penalised, since if the fluctuations are out of phase then the forecast will be wrong most of the time. Instead, we have adopted the strategy of creating a time-evolving spectral representation of the forecast and observed time series, using an adaptive, spectral method called the ‘Hilbert-Huang Transform’ [10]. By integrating the time evolving spectrum over the frequency range of interest, a time series of total variability is obtained. This time series gives the upper bound on the total variability that would be observed if all oscillatory components were in phase [23]. We argue that verifying this time series of total variability is an effective strategy for mesoscale scale-dependent verification, where the existence of fluctuations is a more important measure of forecast skill than their precise timing.

Model specifications and description of observations

The Multi-Scheme Ensemble Prediction System (MSEPS) that has been used in this study is a limited area ensemble prediction system using 75 different NWP formulations of various physical processes. In the multi-scheme approach, perturbations can be added effectively in the physical formulation of convection, cloud and microphysics, horizontal and vertical diffusion and surface roughness [13, 20]. In the MSEPS, the ‘schemes’ differ in their formulation of the fast meteorological processes: dynamical advection, vertical mixing and condensation. In this way, uncertainty in surface processes is targeted.

By varying the formulations of those processes in the individual NWP models that are most relevant for the simulation of fronts and the friction between the atmosphere and earth’s surface, it is possible to simulate the physical uncertainty in short-range weather forecast at all times of the forecast horizon. All derived products such as wind power forecasts then automatically inherit this uncertainty. Another requirement to be met by an ensemble forecasting system is that the spread of the ensemble should be sufficient to cover uncertainties in the forecast which are due to inaccuracies in the initial conditions [21]. In WEPROG’s MSEPS, this is achieved by using slightly different initial conditions for each ensemble member.

The MSEPS was configured to create one long forecast time series for each ensemble member. This is comparable with a climate simulation, but each member was forced with lateral boundary conditions in hourly resolution (with age of 1 to 6 hours). In this way, the ensemble members cannot deviate too much from the true state. The lateral boundary conditions were smoothed so that they did not trigger any variability on time scales less than 3 hours. Although these long forecasts are different to the normal forecasts produced by the MSEPS (which are initialised every 6 hours with a new set of initial conditions), they still maintain nearly correct timing of major features, and as such can be compared to observations with some acknowledged degradation in accuracy. Therefore, the scores that are presented in this paper should not be taken as representative of the normal model behaviour.

As well as the long 10 month time series, 24 hour forecasts, initialised every 6 hours, were used for studying probabilistic variability forecasts. They are run with the same model setup as the long time series.

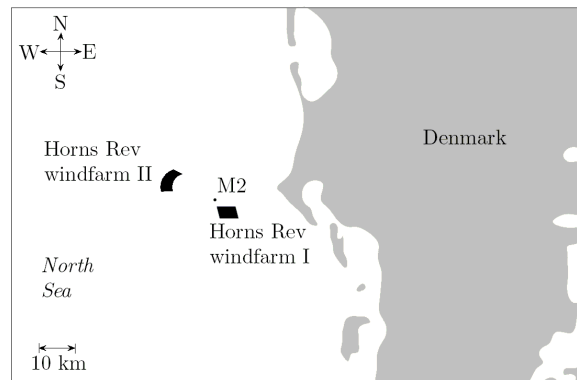


Figure 1. Map of the Horns Rev wind farm, relative to the coastline of western Denmark. The Horns Rev phase II wind farm (under construction) is also shown.

The observations used to verify the model are from a sonic anemometer mounted on a meteorological mast to the North east of the Horns Rev wind farm, at a height of 50 m. The location of the mast, relative to the wind farm, is given in Figure 1. The measurements, which were taken at a frequency of 12Hz, were averaged to a time resolution of 6 minutes for reasonable comparison with model data. Further description of the site is given in [16].

Methodology – The Hilbert-Huang transform as a verification tool

The Hilbert-Huang transform was first described by [10] who systematically compared it to the wavelet transform and illustrated its potential in capturing the time evolving frequency information in time series data including wind speeds, wave height and earthquake vibrations. Recent progress in the methodology and application of the Hilbert-Huang transform are given in [9].

The Hilbert-Huang transform is based on an empirical decomposition of time series into a set of basis functions, and an application of the Hilbert transform to calculate instantaneous amplitude and frequencies of each component. Obvious advantages of the method are its adaptivity and flexibility to capture non-stationary and non-linear behaviour in the time series. It differs from a Fourier transform because it has a local basis, rather than a set of global harmonics which are fitted to the whole time series. It differs from the wavelet transform because no *a priori* decision about a suitable wavelet shape is required. Disadvantages of the method are the empirical nature of the decomposition, that it is hard to show that the basis functions are orthogonal, and that only periods longer than 4 times the measurement resolution are resolved.

The time series is decomposed into adaptive basis functions according to the following procedure: the component with the fastest oscillations is extracted first, by recursively subtracting the mean of the time series (where the mean is defined as the average of two cubic splines fitted through the local maxima and local minima) until the remaining signal contains information about only a single, time dependent frequency. This component forms the first 'Intrinsic Mode Function' (IMF), and is subtracted from the time series. The process of extracting the fastest oscillations is repeated for subsequent modes of increasing time scale. The decomposed time series, $U(t)$, can then be expressed in terms of its IMFs, $x_i(t)$, as shown in equation 1.

$$U(t) = \sum_{i=1}^N x_i(t) + \varepsilon(t) \quad (1)$$

where $\varepsilon(t)$ is the low frequency trend that remains after the higher frequency components have been extracted.

The Hilbert transform is then used to calculate the instantaneous frequency of each component, and the information from each IMF is finally combined to create the Hilbert spectrum: a two dimensional map of amplitude as a function of frequency and time, $H(\omega, t)$. By summing the two dimensional spectrum, $H(\omega, t)$, between two frequencies ω_1 and ω_2 , a scalar time series, $H(t)$ of total amplitude of variability for all scales of motion between ω_1 and ω_2 can be calculated, as given in equation 2.

$$H(t) = \sum_{\omega=\omega_1}^{\omega_2} H(\omega, t) \quad (2)$$

Examples of Hilbert spectra for a simulated and observed 10 day time series of wind speed are shown in Figure 2. By verifying scalar variability time series, we are actually comparing different horizontal slices of the Hilbert spectrum.

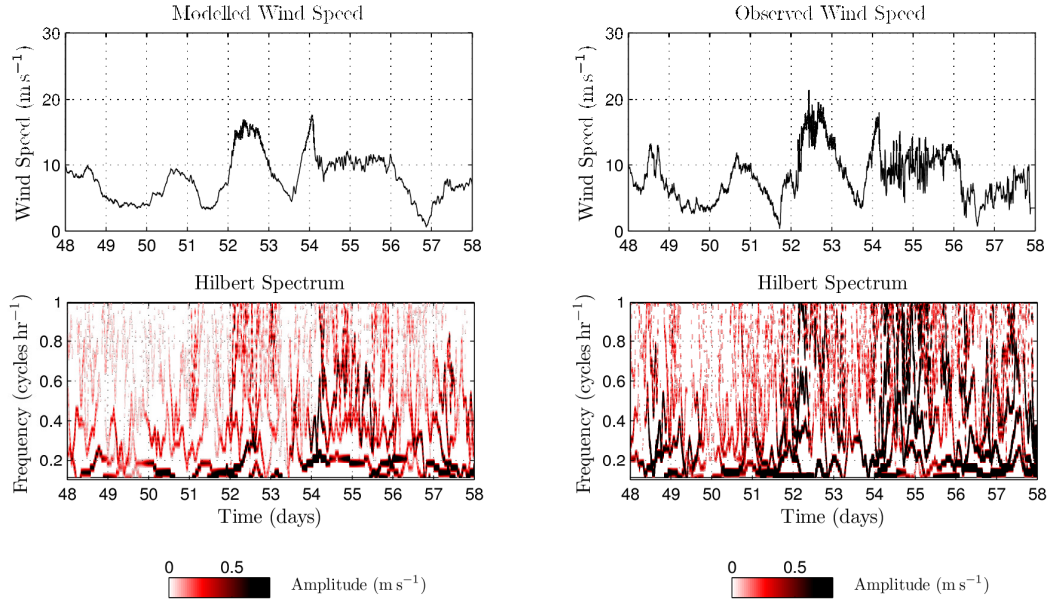


Figure 2. Hilbert spectrum for simulated (left) and observed (right) time series of wind speed.

Results

1. Scale dependent verification scores

10 month continuous simulation time series were verified against observations on six different time scales: 24 min – 3 hr, 24 min – 2hr, 1 – 2hr, 2 – 4 hr, 4 – 8 hr, and 3 – 10 hr. The lower limit of 24 minutes is dictated by the physical time step of the model, which is close to 6 minutes. To resolve variability using the Hilbert-Huang transform, four time steps are required (ie $4\Delta t$, analogous to the $2\Delta t$ Nyquist resolution in a Fourier transform). In the shortest time scales studied, the observed wind field may be influenced by features such as gravity waves, convection and thunderstorms, while the longest time scales may include large thunderstorms, cloud clusters and low-level jets, according to suggested characteristic time scales of [15].

The benefit of using long time series was that the average performance of the system could be calculated on the longer time scales without any necessity to join individual forecasts. For example, studying variability with a period of 10 hours using individual 24 hour forecasts will not give good representivity. The other option of ‘stitching’ individual forecasts together to make a long time series causes the difficult problem of introducing variance with the discontinuities at each join.

The RMSE for the 75 members over the 6 time scales are shown in Figure 3. Several members are missing due to incomplete data. There are patterns in the scores that are due to the different combinations of vertical diffusion, condensation and dynamical schemes in the members. It is seen that members 5,10,15,20 . . . (corresponding to the fifth condensation scheme, CS5) have a particularly good score for the highest resolution variability, while many members in the sequence 3,8,13,18 . . . (corresponding to CS3) have good scores on scales of 4-8 hours. Other scores, including mean absolute error and bias, showed similar patterns.

The RMSE for six time scales is summarised in Figure 4, where the three panels show average scores for each time scale, grouped according to dynamical scheme, condensation scheme and vertical diffusion scheme respectively. Each result was normalised by the averaged observed variance on its respective time scale. It is seen that for the fastest timescales, the ratio of RMSE to average observed variance is close to 1, indicating that the variance in the model has little skill. As the time scales increase, the ratio of error to observed variance decreases.

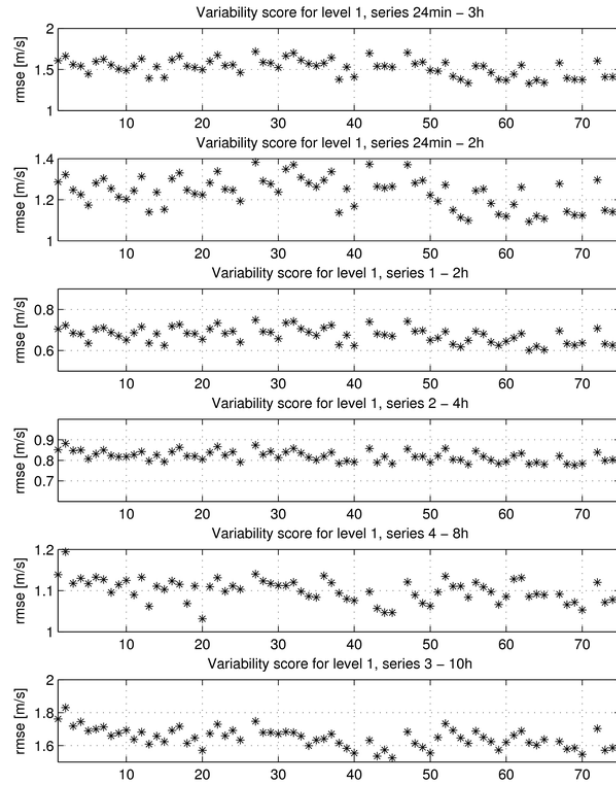


Figure 3. RMSE for the variability scores for the 75 ensemble members, for time scales of 24min-3hr, 24min- 2hr, 1-2hr, 2-4hr, 4-8hr and 3-10hr.

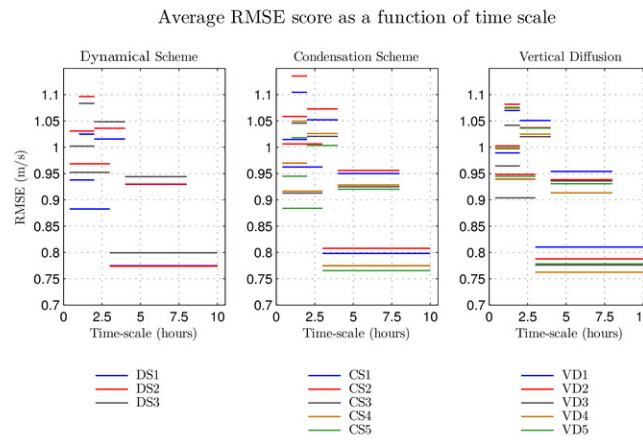


Figure 4. Average RMSE as a function of timescale and parameterisation.

In Figure 5 to Figure 7, scatter plots of variability score against wind speed score are shown, colour coded by model parameterisation. This analysis is intended to address the question of whether a good variability simulation and a good wind speed simulation are possible simultaneously. As discussed above, correct variance but incorrect timing will lead to a poor score for the wind speed forecast. For high frequency mesoscale features, precise timing of fluctuations may be strongly governed by random or chaotic processes with low predictability anyway.

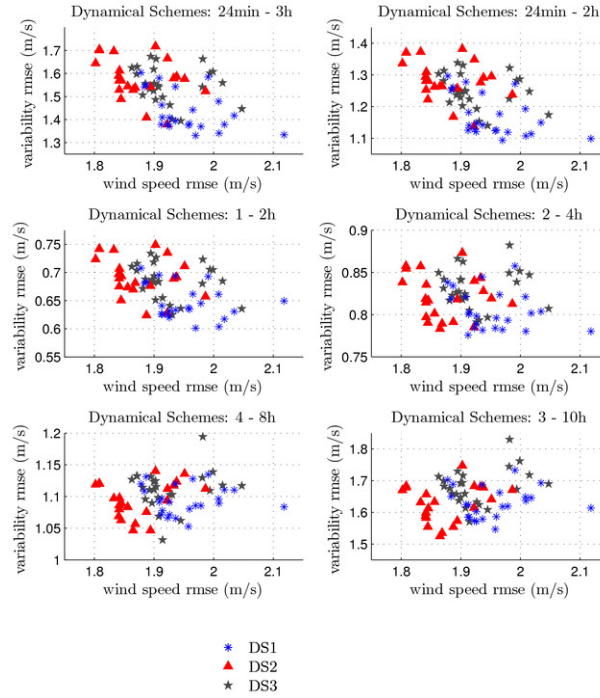


Figure 5. RMSE of variability time series (model level 1), against RMSE of wind speed (model level 1) for 6 time scales. The ensemble members are coded by dynamical scheme (DS).

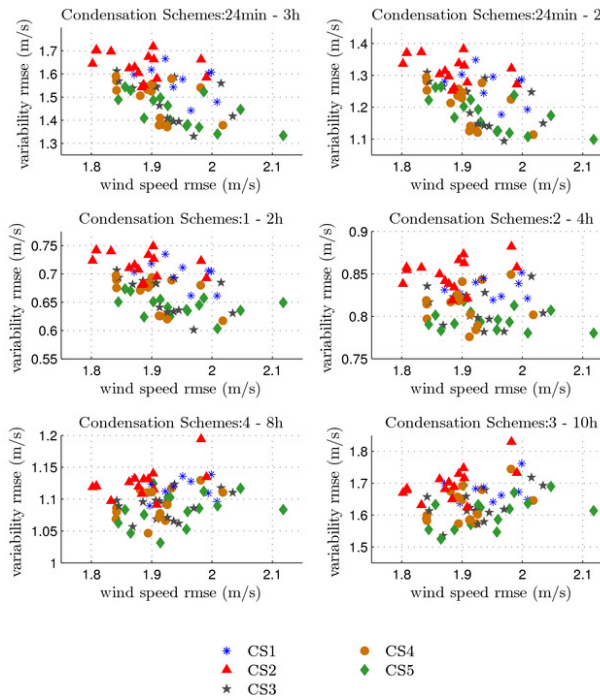


Figure 6. As for figure 5, but coded by condensation scheme (CS).

The plots in figure 5 to figure 7 differentiate the strengths of the ensemble members. For example, members with the second condensation scheme (CS2) tend to have poor variability scores, but rather good wind speed scores, and members with CS5 tend to be good all-rounders with moderate scores for both wind speed and variability. The results coded by dynamical scheme also show groups of members performing differently, while the results coded by diffusion scheme show less pronounced grouping. The condensation scheme is significant because it is linked with the triggering of convective cells, which should certainly influence the boundary layer wind fields.

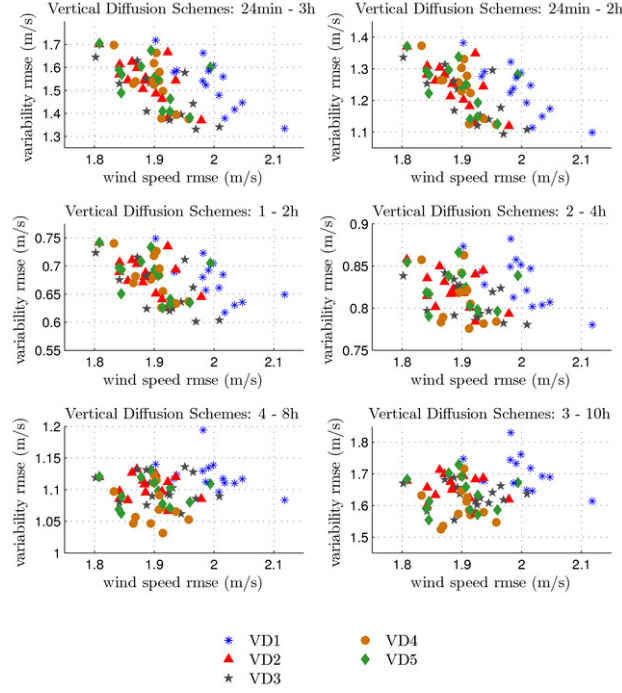


Figure 7. As for figure 5, but coded by vertical diffusion scheme (VD).

2. Probabilistic Forecasts of wind variability

In a real-time setting, model forecasts of only a few days are available, so they cannot be used to study longer period variability. However, it is possible to calculate spectral information for shorter time scales from these forecasts.

Here, variability forecasts (derived, as before, from the Hilbert-Huang transformed wind speed forecasts) are calculated for the 75 MSEPS members for 24 hour forecasts in January 2004 for the time scale of 24 mins – 3 hours. As before, the time resolution of the forecasts is approximately 6 minutes. Together, the 75 variability time series form a probabilistic forecast, which is compared with the observed variability. Full verification of the probabilistic variability time series is not performed within the scope of this work.

The shortness of the time series (240 points) creates some concern about end effects of the Hilbert-Huang transform. It is not satisfactory to simply throw away some of the end data, since this shortens the length of useable forecast. Proper treatment of end effects in the Hilbert-Huang transform is an ongoing research problem [9]. However, for the purpose of this work, a simple autoregressive moving average model was fitted to each end of each IMF. This allowed calculation of a smooth variability time series right to the end of the available time series, although there are almost certainly errors at the very ends of the time series arising from this treatment.

The forecast variability time series showed generally suppressed amplitude compared with the observed variability, so a linear correction was applied to each ensemble member to amplify any existing variance in the forecasts.

$$V_{\text{corrected}}^n = a^n V_{\text{original}}^n \quad (3)$$

where $V_{\text{corrected}}$ and V_{original} are the variability, n is the member and a is a correction factor.

The correction factor, a was calculated from a linear regression between all forecasts and observations for the month of January 2004. It is obviously not ideal to use the same data set for training and testing the model, and this would need to be addressed in a larger study.

Finally, a gamma distribution was fitted to the 75 corrected forecasts for each time step using a maximum likelihood estimation, and the distributions were contoured as shown in Figure 8 (left). The area under each gamma distribution equates to unity, so Figure 8 (left) actually shows a time-

evolving probability density function of variability. Corresponding forecast and observed wind speed time series are also shown. In these cases, and in others not shown here, many of the important changes in variability conditions are captured by at least the tail of the distribution, indicating some skill on the timescales 24 mins – 3 hours. Objective verification of these forecasts is a complex task in itself.

There are several better ways of fitting probability density functions to ensemble forecasts which correct the distribution based on observed data (for example, in [18] or [19]); in a more in-depth study this would certainly be an important extra step.

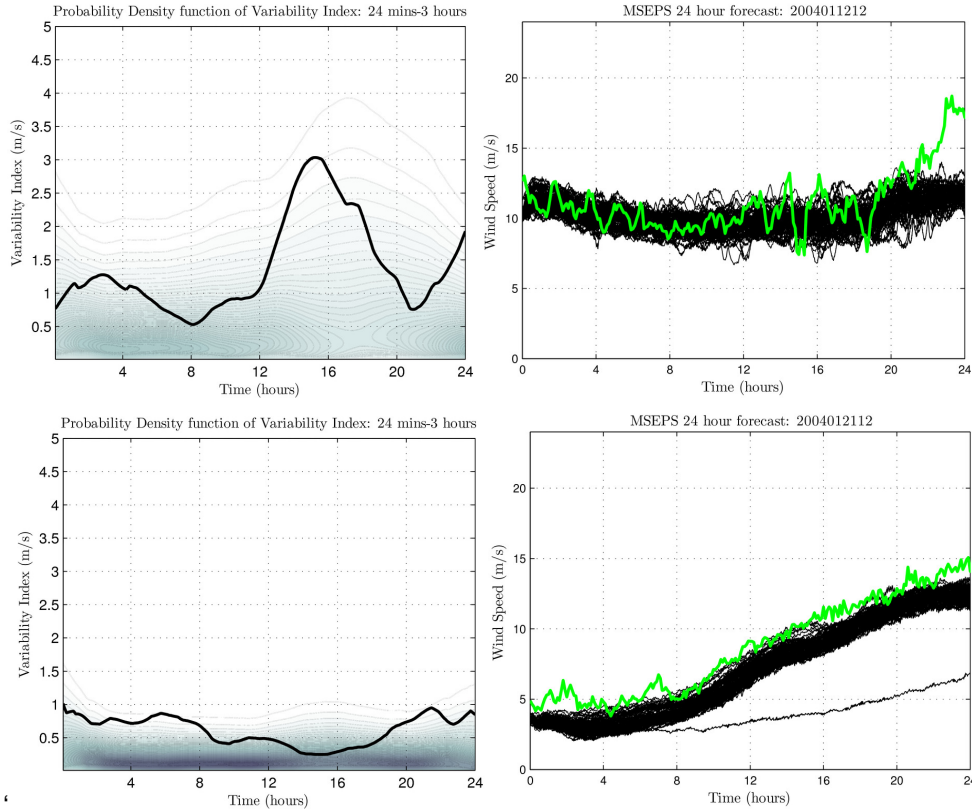


Figure 8. Probabilistic variability forecasts (grey contours, left) based on MSEPS forecasts for (black lines, left). The observed variability time series (black line, left) and observed wind speed time series (green line, right) are also shown. Dates shown are 2004011212 (top) and 2004012112 (bottom).

Conclusions

Mesoscale verification will always be complicated by competing demands of phase and amplitude, and by contributing errors on many time scales. While phase errors are very important for slow components such as the synoptic cycle, phase errors may be less important for short time scales where the existence of mesoscale variability at the right time is a better measure of model skill than the precise timing of peaks and troughs.

In the validation of mesoscale models, spectral verification is a good strategy because it allows us to consider model performance on particular time-scales. The Hilbert-Huang transform was especially suited to this task, because it gives a time evolving spectral representation. It is better than a band pass filter for this application, because it represents the envelope of the fluctuations independent of phase. It has been shown to have fast adaptivity to sudden changes in the statistical properties of the time series, and is therefore ideal for analysing wind speed.

Although it was difficult to draw strong conclusions from the results, it was clear that the accuracy of the model variance increases with increasing time-scale. Ensemble members had different strengths and weaknesses with regard to their relative skill in forecasting the mean wind, and in forecasting wind variability. Probabilistic forecasts of variability suggested that many of the observed changes in variability conditions may be captured by at least the tail of the ensemble distribution.

References

- [1] Akhmatov, A., Rasmussen, C., Eriksen, P.B. and Pedersen, J., "Technical aspects of status and expected future trends for wind power in Denmark", *Wind Energy*, 2007, 10, 31-49
- [2] Doherty, R., and O'Malley, M., "A new approach to quantify reserve demand in systems with significant installed wind capacity", *IEEE Trans. Power. Syst.*, 2005, 2, 587-595
- [3] Fesor, F. and von Storch, H., "A spatial two-dimensional filter for limited area model evaluation purposes", *Mon. Wea. Rev.*, 2005, 133, 1774-1786
- [4] Fesor, F., "Enhanced detectability of added value in limited-area model results separated into different spatial scales", *Mon. Wea. Rev.*, 2006, 134, 2180-2190
- [5] Focken, U., Lange, M., Mönnich, K., Wadl, H.-P., Beyer, H.G., and Luig, A., "Short-term prediction of the aggregated power output of wind farms - a statistical analysis of the reduction of the prediction error by spatial smoothing effects", *J. Wind Eng. Ind. Aerodyn.*, 2002, 90, 231-246
- [6] Giebel, G., Brownsword, R., and Kariniotakis, G., "The state-of-the-art in short-term prediction of wind power - a literature overview". Deliverable report d1.1, Project ANEMOS, 2003. [Available from <http://anemos.cma.fr/>].
- [7] Gulev, S., Jung, T., and Ruprecht, E., "Climatology and Interannual Variability in the Intensity of Synoptic-Scale Processes in the North Atlantic from the NCEP-NCAR Reanalysis Data", *J. Clim.*, 2002, 15, 809-828
- [8] Henderson, A.R., Morgan, C., Smith, B., Sørensen, H.C., Barthelmie, R.J., and Boesmans, B., "Offshore wind energy in Europe - a review of the state-of-the-art", *Wind Energy*, 2003, 3, 35-52
- [9] Huang, N., and Wu, Z., "A review on Hilbert-Huang transform: Method and its applications to geophysical studies", *Rev. Geophys.*, 2008, 46, DOI: 10.1029/2007RG000228.
- [10] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., and Liu, H.H., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proc. Roy. Soc. A.*, 1998, 454, 903-995
- [11] Jung, T., and Leutbecher, M., "Scale-dependent verification of ensemble forecasts", *Quart. J. Roy. Meteor. Soc.*, 2008, 134, 973-984
- [12] Möhrlen, C., Jørgensen, J., Pinson, P., Madsen, H., and Kristoffersen J.R., "HREnsembleHR – High Resolution Ensemble for Horns Rev: A project overview", Proc. European offshore wind energy conference, Berlin 2007.
- [13] Möhrlen, C., "Uncertainty in wind energy forecasting", Ph.D. dissertation, University College Cork, Ireland, 2004.
- [14] Möhrlen C., Jørgensen, J., "Forecasting wind power in high wind penetration markets using multi-scheme ensemble prediction methods", Proc. German Wind Energy Conference DEWEK, Bremen, Nov. 2006.
- [15] Orlanski, I., "A rational subdivision of scales for atmospheric processes", 1975, 56, 527-530
- [16] Peña, A., Hasager, C. B., Gryning, S.-E., Courtney, M., Antoniou, I., and Mikkelsen, T., "Offshore wind profiling using light detection and ranging measurements", *Wind Energy*, 2008, 12, 105-124
- [17] Pinson, P., Chevallier, C., and Kariniotakis, G., "Trading wind generation with short-term probabilistic forecasts of wind power", *IEEE Trans. Pow. Sys.*, 2007, 22, 1148-1156
- [18] Pinson, P., and Madsen, H., "Ensemble-based probabilistic forecasting at Horns Rev", *Wind Energy*, 2009, 12, 137-155
- [19] Slughter, J.M., Gneiting, T., and Raftery, A.E., "Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging", <http://www.stat.washington.edu/research/reports/2008/tr544.pdf>. University of Washington, Department of Statistics, Technical Report no. 544.
- [20] Stensrud, D.J., Brooks, H.E., Du, J., Tracton M.S., and Rogers, E., "Using Ensembles for Short-Range Forecasting", *Mon. Wea. Rev.*, 1999, 127, 433-446.
- [21] Strauss, B., and Lanzinger, A., "Verification of the Ensemble Prediction System (EPS)", *ECMWF Newsletter*, 1996, 72, 9-13
- [22] Stull, R.B., *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, 1988.
- [23] Vincent, C., Giebel, G., Pinson, P., and Madsen, H., "Resolving non-stationary spectral information in wind speed time series using the Hilbert-Huang transform", 2008, submitted

Acknowledgement

This work was supported by the Danish Public Service Obligation (PSO) fund project 'HREnsemble – High Resolution ENSEMBLES for Horns Rev' (under contract PSO-6382), which is gratefully acknowledged. The sonic anemometer data used for verification was downloaded from the Danish database <http://www.winddata.com>.