

Reducing errors of wind speed forecasts by an optimal combination of post-processing methods

Conor P. Sweeney,* Peter Lynch and Paul Nolan

Department of Meteorology and Climate Centre, School of Mathematical Sciences, University College Dublin, Belfield, Ireland

ABSTRACT: Seven adaptive approaches to post-processing wind speed forecasts are discussed and compared. Forecasts of the wind speed over 48 h are run at horizontal resolutions of 7 and 3 km for a domain centred over Ireland. Forecast wind speeds over a 2 year period are compared to observed wind speeds at seven synoptic stations around Ireland and skill scores calculated. Two automatic methods for combining forecast streams are applied. The forecasts produced by the combined methods give bias and root mean squared errors that are better than the numerical weather prediction forecasts at all station locations. One of the combined forecast methods results in skill scores that are equal to or better than all of its component forecast streams. This method is straightforward to apply and should prove beneficial in operational wind forecasting. Copyright © 2011 Royal Meteorological Society

KEY WORDS adaptive post-processing; numerical weather prediction; Kalman filter; artificial neural network

Received 7 February 2011; Revised 19 July 2011; Accepted 4 August 2011

1. Introduction

Wind energy is a growing industry, and is supplying electricity to national grids worldwide. Wind energy cannot be generated on demand, in the manner of traditional electricity generation. Forecasting methods are required for the efficient management of wind energy. Reliable forecasts of wind energy production would reduce the costs of running the grid. They could also benefit wind farm operators by allowing a higher price to be obtained on the electricity spot market (Barthelmie *et al.*, 2008). The purpose of this paper is to assess methods of reducing the errors in wind speed forecasts. Adaptive post-processing methods that can easily be applied to wind forecasts produced by Numerical Weather Prediction (NWP) models are compared, and the advantage of combining forecast streams is investigated.

Ireland has a target of 40% renewable electricity generation by 2020. Most of this is expected to come from wind power. This will result in the Irish power system having one of the highest levels of wind penetration in the world by 2020. The highest penetration on the Irish power system to date was achieved on 5 April 2010. During this day, wind generation reached a peak of 1260 MW, which was 42% of the system demand at that time (EirGrid, 2010). Integrating this level of non-synchronous generation on the power system will materially affect the way the electricity system is operated. A key factor in managing this variability is the development of improved forecasting methods.

Reviews of wind power forecasting methods are available from different authors (Giebel, 2003; Costa *et al.*, 2008; Lei *et al.*, 2009), and only a selection of different methods are considered here. One common approach in wind speed forecasting uses recent meteorological data. This is usually available from anemometers on the site of the wind farm. The data may then be

analysed with different statistical models, such as autoregressive processes (Brown *et al.*, 1984; Torres *et al.*, 2005), and artificial intelligence techniques (Sfetsos, 2000; Cadenas and Rivera, 2010; Li and Shi, 2010). These techniques are only of use for forecasts of an hour or so into the future. This paper considers a forecast range of 48 h, and so it is important to use NWP model data.

All NWP forecasts contain errors, due in part to the quality of the data used to drive the model and in part to computational limitations in solving the governing equations with finite resolution. Some NWP errors are systematic and it is hoped that these can be reduced by applying statistical post-processing methods.

Model Output Statistics (MOS) is a commonly used statistical post-processing technique. It uses multiple linear regressions with forecast data and observations to attempt to remove forecast errors (Glahn and Lowry, 1972). However, MOS needs a long record of data for its training. This may cause difficulties when changing or updating the NWP model or the observing network.

It may be possible to equal or exceed MOS skill using adaptive short-term post-processing methods. Simple short-term bias-correction has been shown to produce ensemble mean forecasts of 2 m temperature, 2 m dew point temperature, and 10 m wind speed that are competitive with or better than those available from MOS (Stensrud and Yussouf, 2005). Another study showed that MOS outperforms post-processing with a Kalman filter or short-term bias-correction when model biases change dramatically, performs worse during quiescent cool season patterns, and that all three are comparable at other times (Cheng and Steenburgh, 2007).

Kalman filtering (Kalman, 1960) is also used as a post-processing tool for wind forecasting and has been shown to reduce systematic errors in a consistent manner (Crochet, 2004; Louka *et al.*, 2008). Artificial Neural Networks (ANNs) are another popular method for post-processing wind forecasts. They have shown good results when downscaling NWP wind

*Correspondence to: C. P. Sweeney, Department of Meteorology and Climate Centre, School of Mathematical Sciences, University College Dublin, Belfield, Ireland. E-mail: conor.sweeney@ucd.ie

speed, such as at a wind farm in Spain (Salcedo-Sanz *et al.*, 2009a).

Previous studies have shown that wind power forecasts produced by combining individual forecasts can perform better than any of the individual forecasts (Nielsen *et al.*, 2007). Two methods of combining forecast data are compared in this paper. One method uses ANNs to combine the forecasts, similar to Salcedo-Sanz *et al.* (2009b). The other method is a simple scheme using weights based on recent forecast skill, an updated version of the method used in Sweeney and Lynch (2011).

In this paper, seven different adaptive post-processing methods are applied to NWP data produced at two horizontal resolutions (7 and 3 km) to produce 48 h wind speed forecasts. The training period is limited to 30 days to allow forecast data streams to be updated or replaced without requiring a large lead-in period. Two different methods are then used to combine the forecast data. All 48 h wind speed forecasts are compared to the actual wind speeds observed at seven stations around Ireland over 2 years (June 2008 to June 2010). The skill scores considered are the bias and the Root Mean Square Error (RMSE).

Section 2 concerns the COSMO NWP model, the forecast verification, and the statistical post-processing methods, and methods used to combine the forecasts. The results of applying the different post-processing and forecast combination methods are given in Section 3. Conclusions are presented in Section 4.

2. Methodology

2.1. The COSMO model

The COSMO-Model Package is a regional numerical weather prediction system. It is based on the COSMO-Model, formerly known as the Lokal Modell (Steppeler *et al.*, 2003), a non-hydrostatic limited-area atmospheric prediction model. The COSMO interpolation program, INT2LM, interpolates data from different data sets to the limited-area rotated latitude-longitude grid of the COSMO-Model. Thus it provides the initial and boundary data necessary to run the COSMO-Model. The prediction model uses fully compressible hydrothermodynamical equations. A variety of different schemes are used for sub-grid-scale processes. More comprehensive information about COSMO is available on the project web-site (COSMO, 2010).

The deterministic forecast from the ECMWF Integrated Forecast System, IFS T799L91 (Untch *et al.*, 2006), supplied the data used to drive the COSMO model. This ECMWF forecast has a horizontal grid spacing equivalent to 25 km. The 0000 analysis and forecast (+00 to +48) data were retrieved each day. The boundary data were available at 3 h intervals.

The two areas used for the forecasts are shown in Figure 1. The 7 km forecast used a rotated lat./long. grid of 0.0625° , with 40 vertical levels and a time-step of 40 s. The time-integration scheme used was a three time-level Leapfrog scheme with time-split treatment of acoustic and gravity waves. The output of the 7 km forecast was used to drive the 3 km forecast (one-way nesting). The 3 km forecast used a rotated lat./long. grid of 0.025° , with 50 vertical levels and a time step of 25 s. The time-integration scheme used was a two time-level Runge-Kutta time-split scheme. Output data were saved every forecast hour from 00 to +48 h. The COSMO model version used was 4.11.

2.2. Forecast verification

The forecast models produced wind speed data for every hour over a 48 h period, on a grid covering Ireland. In order to

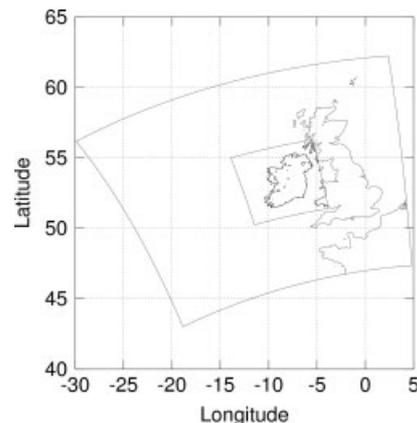


Figure 1. Domains used by COSMO. 7 km (large area) and 3 km (small area).

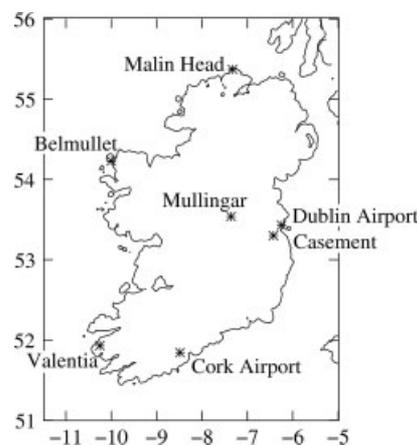


Figure 2. Synoptic stations used for forecast verification.

check these wind speeds they were compared to the actual wind speeds observed at different locations around Ireland. Met Éireann (the Irish National Meteorological Service) maintains an observation network covering Ireland, and they provided wind speed data for seven synoptic stations. These stations are in different locations around Ireland, as shown in Figure 2. The observed wind speed data are taken at a height of 10 m. The COSMO model has an option to output wind speed data at a height of 10 m and this was used. Bilinear interpolation was used to calculate the wind speed at the stations from the wind speed on the model grid. Bias and RMSE scores could then be calculated from the observed and forecast wind speeds.

2.3. Statistical post-processing methods

Seven different post-processing methods are considered in this paper. All of the methods are adaptive, in that they are trained each day on data from previous days. The number of previous days used to train the methods (called the width of the sliding window) has been limited to 30, so that the methods can be applied to new locations and/or NWP models in a relatively short time. The methods range from the very simple (bias-correction) to more advanced (Kalman filter and ANN). Methods are also included that target particular characteristics of wind speed forecast errors: the diurnal cycle, mean and standard deviation of errors, and wind direction.

2.3.1. Short-term bias-correction forecasts (STB)

Short-term bias-correction forecasts are calculated using a rolling window of 30 days. For each station, the mean of all wind speed forecast errors (forecast speed minus observed speed) over the previous 30 days is calculated. Note that the 48 h forecast from yesterday includes wind speeds for which there are no observations yet (FC + 24 to FC + 48), and so these are omitted from this and all other post-processing methods.

The bias-corrected forecast (STB) for each station is calculated by subtracting the mean error from the forecast speed. STB wind speeds are constrained to be greater than or equal to zero by setting any negative speeds to zero. This is also done for all other post-processing methods. The process is repeated for each station. A test was run to decide on the optimum number of days to use for the sliding window. STB forecasts were calculated for window sizes from 2 to 30 days for 1 year of wind speed data (January to December 2008). The overall root mean square error (RMSE) was calculated for each station and window size. The RMSE flattens out for most stations after 10 days, and the minimum mean RMSE for all stations occurs with a window size of 30 days. Therefore, the window size was set to equal 30 days for all STB forecasts.

2.3.2. Diurnal cycle forecast correction (DRL)

If the average forecast wind speed for each hour is plotted along with the corresponding average observed wind speed, it is possible to compare the diurnal cycle of the forecast winds to that of the observed winds. All forecasts in this study were started at midnight (0000), and are 48 h forecasts, so two diurnal cycles should be apparent. Figure 3 shows the observed wind speed at Dublin Airport, averaged over 2 years. Also shown is the average forecast wind speed of the 7 and 3 km forecasts. A bias can be seen, where the forecast wind speed is less than the observed wind speed at all forecast hours. If the mean observed speed at each hour is subtracted from the mean forecast speed for that hour a mean forecast error is obtained for each forecast hour. This is shown in the bottom part of the figure. A pattern can be seen in the error, which varies by a factor of two over a diurnal cycle. All of the forecast wind speeds are too low, but the error in the forecast wind speeds is larger in magnitude during the day than at night. An attempt can be made to correct this error by calculating the average error for each forecast hour over the past 30 days, err_{HH} , and applying

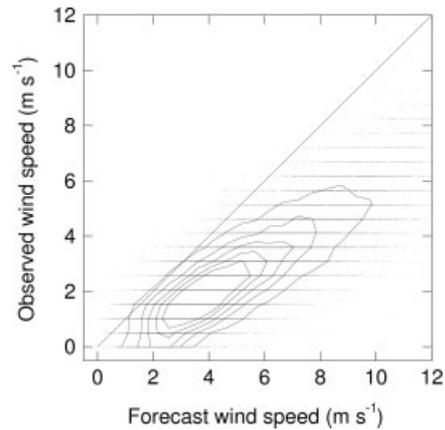


Figure 4. Scatterplot showing COSMO 7 km forecast wind speeds and corresponding observed wind speeds at Mullingar. Contours show population density of points, in levels of 100.

the result to the present forecast, FC_{HH} , for each forecast hour: $DRL_{HH} = FC_{HH} - err_{HH}$. This is done by the DRL forecast.

2.3.3. Linear least-square corrected forecast (LLS)

Linear regression attempts to find the relationship between the forecast wind speed and the observed wind speed. The linear expression that minimizes the least-square-error of the fit is found. The relationship between forecast wind speed and observed wind speed is different from station to station. Figure 4 plots the COSMO 7 km direct model output (DMO) forecast wind speeds (x-axis) and the corresponding observed wind speeds (y-axis) for the station at Mullingar over 2 years. A perfect fit (1:1) is shown as a solid line at 45°. Most of the points lie below this line showing that, for this station, the DMO forecast over predicts the wind speed.

If there is a linear relationship between the forecast and observed wind speeds, this can be used to correct the DMO forecast. Linear least-square corrected forecasts (LLS) are calculated using a rolling window of 29 days. For each station, the forecast wind speeds and corresponding observed wind speeds over the previous 29 days are used to calculate the slope (m) and intercept (c) of the linear least square error fit. The LLS forecast is then given by Equation (1):

$$(LLS \text{ speed}) = m(\text{forecast speed}) + c \quad (1)$$

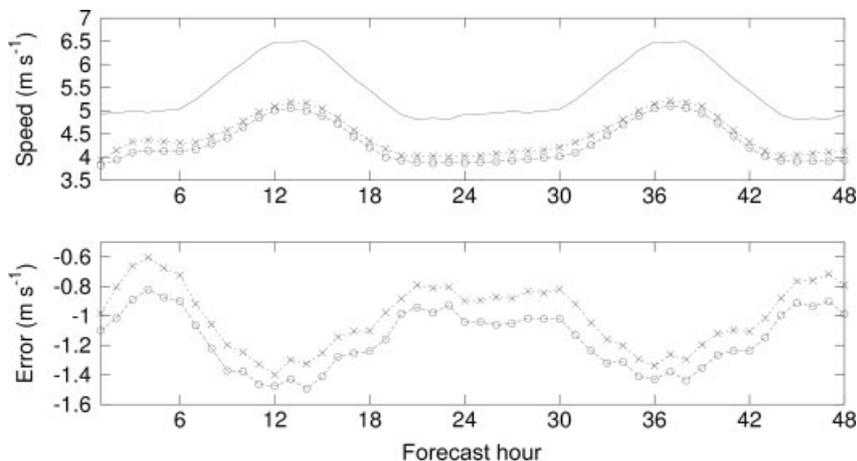


Figure 3. Hourly wind speed and forecast error for Dublin Airport, averaged over 2 years. OBS (—), COSMO 7 km (*), COSMO 3 km (○).

A test was run to decide on the best number of days to use for the sliding window. LLS forecasts were calculated for window sizes from 2 to 30 days for 1 year of wind speed data. The RMSE generally flattens out after 12 days, and the minimum mean RMSE for all stations occurs with a window size of 28/29 days. Therefore the window size was set to equal 29 days for all LLS forecasts.

2.3.4. Kalman filter forecasts (KAL)

The Kalman filter is a popular method for calculating the least-squares fit. It updates the fit as new data become available and uses the variance of recent errors and changes in the state vector in its calculations. A simple Kalman filter is used in this paper to generate a forecast (KAL). The Kalman filter is described in papers such as Crochet (2004) and Galanis and Anadrastakis (2002), and a description of the Kalman filter used here is given in Sweeney and Lynch (2011).

A test was run to decide on the best number of days to use for the sliding window. KAL forecasts were calculated for window sizes from 2 to 30 days for 1 year of wind speed data. The RMSE reached a minimum for all stations between 5 and 10 days, and then increased again. The window size was set to equal 8 days for all KAL forecasts.

2.3.5. Mean and variance corrected forecast (MAV)

Another method to improve a forecast is to process it so that it has the same mean and standard deviation as the observed data. Consider a set of forecast data (f_c) which is to be corrected so that it has the same mean (μ) and standard deviation (σ) as the observed data (obs). The standard deviation of f_c is corrected first:

$$f_{c2} = \left(\frac{\sigma_{obs}}{\sigma_{fc}} \right) f_c \tag{2}$$

The mean is then corrected:

$$f_{c3} = f_{c2} + (\mu_{obs} - \mu_{fc2}) \tag{3}$$

f_{c3} now has the same mean and standard deviation as obs . If the mean and standard deviation of the forecast and observed wind speeds are calculated over the past 30 days they can be used as an estimate of the true values and the current forecast can be corrected as above to produce the MAV forecast.

2.3.6. Directional-bias forecast (DIR)

It is possible that some locations will have wind speed errors that are related to the wind direction. This may happen, for example, due to nearby mountains that are not properly resolved by the forecast model, or incorrect model surface roughness for different wind directions.

To investigate this, the mean wind speed forecast error over 2 years is plotted, binned by wind direction with 30° bin width, for the different observation stations. Figure 5 shows the resulting plot for the 3 km COSMO forecast at Belmullet. It can be seen that there is an overall negative bias at this location (represented by the dashed line), but it is also apparent that the bias changes for different forecast wind directions, with

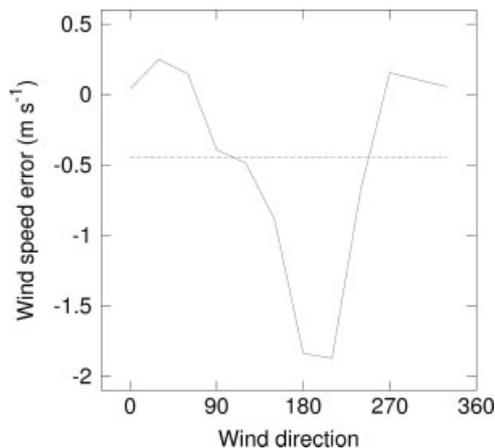


Figure 5. Mean wind speed error for the COSMO 3 km forecast at Belmullet with 30 degree wind direction bins.

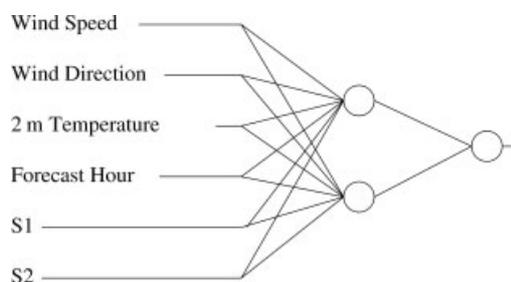


Figure 6. Schematic of ANN used, showing inputs and architecture.

southerly winds less accurately predicted than those from the north.

The DIR forecast uses a sliding window to bin wind speed errors by forecast wind direction. The present forecast wind direction is then used to select the error correction to apply to the present forecast wind speed, giving the DIR forecast wind speed. DIR test forecasts were calculated for window sizes from 2 to 30 days for 1 year of wind speed data. A 30 day window gave the lowest RMSE values, so the window width is set to 30 for all DIR forecasts.

Rare wind directions may not be represented in a 30° bin over the previous 30 days. To avoid this problem, all 30° bins are initialized to equal the overall 30 day mean error, and subsequently updated by errors for that bin where available. Wind directions are unreliable for low wind speeds, therefore the DIR methods corrects by overall 30 day mean error if wind speed is low.

2.3.7. Artificial neural network forecast (ANN)

Artificial neural networks are another popular method for producing wind forecasts. In this paper, the R Project for Statistical Computing program (R Development Core Team, 2010) was used with package nnet to generate wind speed forecasts. The nnet package iteratively minimizes the squared error criterion, including a penalty term, using a technique similar to, but more sophisticated than, standard backpropagation (Ripley, 1996; Venables and Ripley, 2002). A schematic diagram of the network is shown in Figure 6. The ANN used was a single-hidden-layer neural network. The input variables used were the wind speed, direction, and 2 m temperature from the forecast model, the forecast hour (HH = 01–48) and two solar cycle

variables, as used in Salcedo-Sanz *et al.* (2009b):

$$S_1 = \sin\left(\frac{HH2\pi}{24}\right) \quad S_2 = \cos\left(\frac{HH2\pi}{24}\right) \quad (4)$$

Two tests were run using 1 year of data to decide on the configuration of the ANN. First, the number of neurons in the hidden layer (*nhn*) was varied from 2 to 20, while using a sliding window width (*WNDW*) of 30 days. It was found that the lowest RMSE was produced with two nodes in the hidden layer, and this layout was used for all forecasts. A test was also run with *nhn* = 2, varying *WNDW* from 3 to 30. The overall RMSE reduced with increasing *WNDW* size, and so *WNDW* = 30 was used for all ANN forecasts. For each ANN forecast, the neural network was trained using inputs from the previous 30 days, and then generated the ANN forecast using that day's forecast data.

2.4. Combining forecasts

Statistical post-processing methods are effective at reducing the bias and RMSE of raw NWP model forecasts. However, this paper also considers whether a further improvement in forecast skill can be achieved by combining all available forecasts in an automatic and adaptive manner. To this end, two different methods of combining forecast streams to produce a new forecast are considered.

2.4.1. ANN-combined forecast (ANNCOM)

An ANN, similar to that in Section 2.3.7, is used to combine all forecast streams. NWP model output plus 7 post-processed forecasts at 2 resolutions result in 16 available forecasts. All 16 forecasts are used as inputs to the ANN. As before, tests were run using 1 year of data to decide on the optimal configuration for the ANN. The number of hidden neurons was varied from 1 to 20. The overall RMSE increased with *nhn*, and so *nhn* was set to 1. The window width was also varied from 2 to 30. The RMSE reduced sharply as *WNDW* increased from 2 to 14 days. RMSE continued to decrease, at a slower rate, with further increases to the window width. Therefore, data from the previous 30 days are used when training the ANN for each ANN-combined forecast, ANNCOM.

2.4.2. Mean square error-combined forecast (MSECOM)

The second method used to combine forecasts is a simple scheme using weights based on recent forecast skill. This is

done by taking the mean value of the squared wind speed errors over the previous 2 days for each forecast method at each station. This will result in one error value, err_i , for each of the available forecasts, fc_i . The forecasts are then assigned weights in proportion to the inverse of their error values, as described in Equation (5). These weights are normalized to sum to unity.

$$MSECOM = \sum_{i=1}^{16} w_i fc_i \quad \text{where } w_i = \frac{1/err_i}{\sum_{j=1}^{16} (1/err_j)} \quad (5)$$

When the forecasts are combined using these weights the mean square error-combined forecast (MSECOM) is obtained. Were $err_i = 0$ for any forecast fc_i , then that forecast would be automatically selected, but this was never found to happen. The method was tested for different values of window width, from 2 to 30. The overall RMSE did not change much as *WNDW* increased, but the lowest overall value was for *WNDW* = 2, and so this was used for all forecasts.

3. Results and discussion

A simple score often used to test forecast skill is the bias. The overall bias is calculated as the forecast wind speed minus the observed wind speed, averaged over all days (1 June 2008 to 31 May 2010) and forecast hours (+01 to +48) at each station. Figure 7 shows the wind speed bias of the COSMO forecasts at the two model resolutions used. The higher-resolution (3 km) bias is smaller than the 7 km bias at three stations, larger at another three stations, and similar at the remaining station.

All seven post-processing methods used in this paper reduce the wind speed bias at all stations and both forecast resolutions to under 0.1 m s^{-1} , with the exception of the DIR 7 km forecast, which results in a bias of 0.127 m s^{-1} at Mullingar. Therefore all methods are considered to be effective at reducing overall bias.

To investigate the skill of the forecasts further, Figure 8 shows the average wind speed bias for each forecast hour at Dublin Airport. The 7 km COSMO forecast shows an overall negative bias, as well as a diurnal signal in the error. The STB forecast applies a single bias-correction to all forecast hours for each forecast. It can be seen that the STB forecast has reduced the overall bias, but still contains a diurnal signal. The DRL forecast applies a different correction to each forecast hour, and

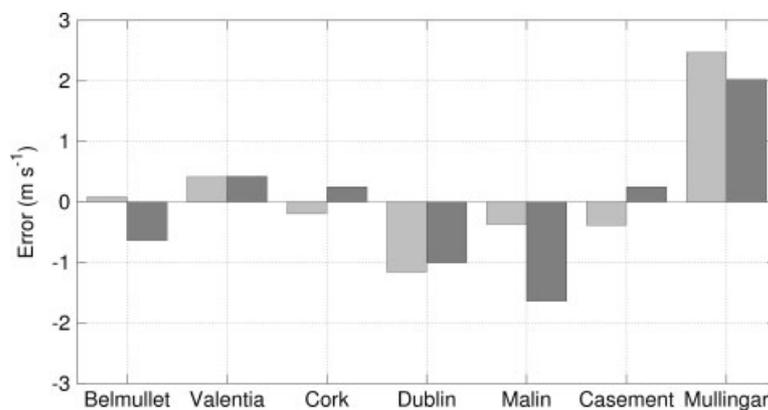


Figure 7. COSMO wind speed bias per station. —: 7 km; —: 3 km.

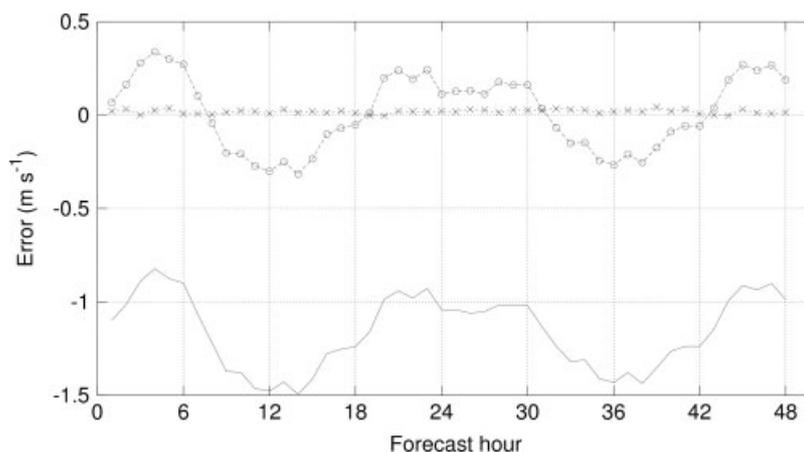


Figure 8. Wind speed bias by forecast hour at Dublin Airport. COSMO 7 km (—), STB (○), DRL 7 km (×).

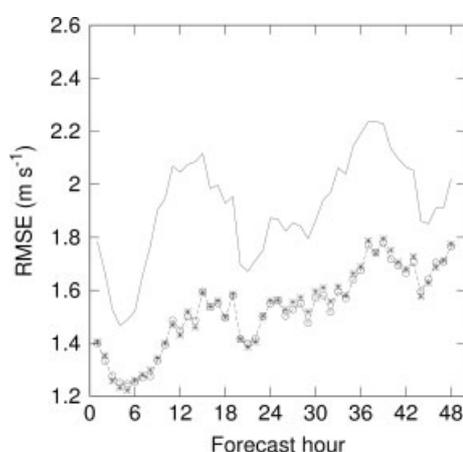


Figure 9. Wind speed RMSE by forecast hour at Dublin Airport. COSMO 7 km (—), DRL 7 km (×), STB 7 km (○).

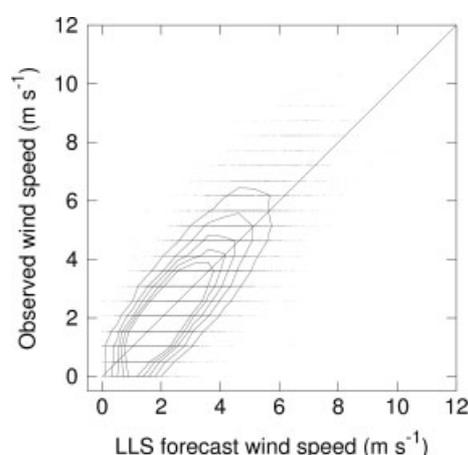


Figure 10. Scatterplot showing 7 km LLS forecast wind speeds and corresponding observed wind speeds at Mullingar. Contours show population density of points, in levels of 100.

it can be seen to do a good job of reducing not only the overall bias, but also the bias at each forecast hour.

The COSMO forecasts produce average hourly biases ranging from -1.972 to 2.656 m s^{-1} across all stations. The STB forecasts reduce these hourly biases to between -0.514 and 0.620 m s^{-1} , and the DRL forecasts reduce them further to between -0.041 and 0.118 m s^{-1} . The DRL forecast reduces the diurnal signal in the wind speed bias, as was hoped. Bias, however, is not a reliable measure of forecast skill on its own, as it may be hiding balanced negative and positive errors. To obtain another indicator of forecast skill, the RMSE of the forecast was calculated.

Figure 9 shows the RMSE, averaged over 2 years, for each forecast hour at Dublin Airport for the COSMO, STB and DRL forecasts. Although DRL has reduced the diurnal signal in the bias, it has not produced any improvement in the RMSE of the forecast compared to the simpler STB forecast. This is true for all stations.

The aim of the LLS forecast is to exploit the quasi-linear relationship between forecast and observed wind speeds. Figure 10 shows the LLS 7 km forecast and observed wind speeds at Mullingar over the 2 year period. It can be seen that the LLS forecast has corrected the DMO forecast (shown in Figure 4) so that it is in better agreement with the 1 : 1 line. This does result in an improved RMSE score. The RMSE scores for

the 7 km COSMO, STB and LLS forecasts at Mullingar are 2.873 , 1.389 and 0.981 m s^{-1} respectively.

The LLS forecast only produces an improvement over STB if there is a weak linear relationship to start with. The 7 km COSMO forecast for Casement, for example, has a strong linear relationship with observed wind speeds, and the LLS forecast does not improve the RMSE of the STB forecast for that case.

The KAL forecast, produced using the Kalman filter described in Section 2.3.4, also seeks to find the optimal linear relationship to correct forecast data. It does a good job, and produces data that are in close agreement with the 1 : 1 fit, but not quite as close as those of the LLS forecasts. The RMSE score of the LLS forecast is always better than that of the KAL forecast, for both resolutions at all stations.

The MAV forecast seeks to correct the distribution of forecast wind speeds so that their mean and variance agree with those of the observed wind speeds. Figure 11 shows the 3 km COSMO, MAV and observed wind speeds at Malin Head. The MAV forecast has successfully corrected the COSMO forecast so that it is in better agreement with the observed wind speed distribution. However, the RMSE of the MAV forecast is worse than that of the simple STB forecast at five of the seven stations.

The DIR forecast uses the forecast wind direction to apply a correction to the forecast wind speed. Figure 12 shows the wind speed error binned by forecast wind speed for the 3 km

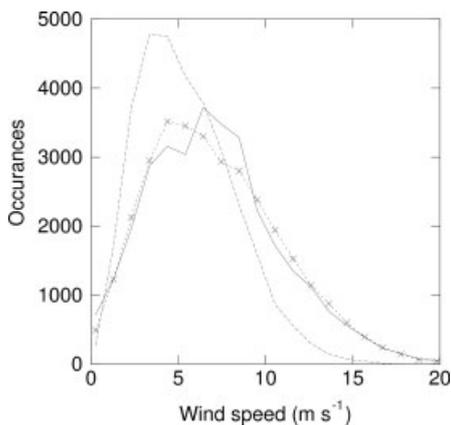


Figure 11. Observed (—), 3 km COSMO (---) and 3 km MAV (⋈) wind speed distributions at Malin Head.

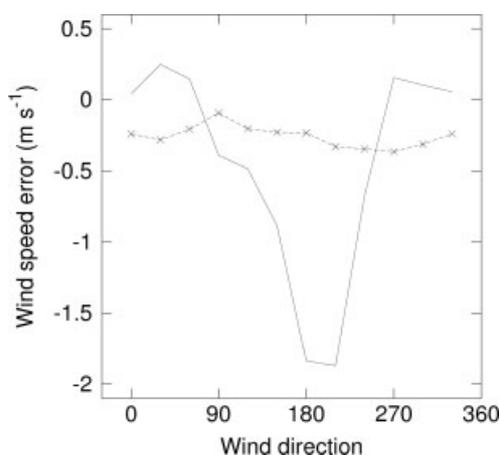


Figure 12. COSMO 3 km (—) and DIR 3 km (⋈) forecast wind speed error binned by forecast wind direction for Belmullet.

COSMO and DIR forecasts at Belmullet. It is clear that the DIR forecast has substantially reduced the dependence of wind speed error on forecast direction. This also results in a lower RMSE.

The RMSE for the 3 km COSMO, STB and DIR forecasts at Belmullet are 2.199, 2.114 and 2.006 m s^{-1} respectively. This improvement of DIR RMSE over STB RMSE happens at stations where there is a clear dependence of the wind speed forecast error on wind direction. Many wind farms are in hilly terrain, where the flow is strongly influenced by orography, and there is likely to be a strong dependence of model forecast error on wind direction.

The ANN forecasts were effective at reducing the RMSE at all stations to less than the COSMO forecasts, and improved on the STB RMSE for 8 of the 14 cases (seven stations at two resolutions).

Table 1 shows the overall RMSE of the wind speed forecasts for all post-processing methods, with the best RMSE scores in bold. There is no single post-processing method that produces the best score at all stations. Indeed, it is often the case that the best method for the 7 km forecast at a station is different to the best method for the 3 km forecast at the same station. The direct model output (COSMO) never produces the lowest RMSE. STB is best in one case, ANN is best in four cases, LLS is best in four cases, and DIR is best in five cases. The LLS forecast produces the best RMSE averaged over all cases.

The fact that no single method is consistently the best provides the motivation for the combined forecasts, ANNCOM and MSECOC. The RMSE scores for these forecasts are shown in Table 2, with the best scores in bold. There is a clear advantage in combining forecast streams. Both combined forecasts out-perform the COSMO forecast for all 14 test cases. ANNCOM gives better RMSE scores than its constituent forecast streams for 12 of the 14 test cases, but performs slightly worse than the Valentia 7 km LLS and the Dublin 7 km LLS forecasts. MSECOC, however, gives better RMSE scores than any of its constituent forecast streams for 13 of the 14 test cases and equals the RMSE skill in the 14th case.

4. Conclusions

Seven different post-processing methods have been applied to NWP output at seven locations around Ireland over a period of 2 years. All of the post-processing methods reduce the model

Table 1. Wind speed forecast RMSE (m s^{-1}) for seven post-processing methods.

Station	km	COSMO	STB	DRL	LLS	KAL	DST	DIR	ANN
Belmullet	7	1.954	1.953	1.964	1.956	2.005	2.032	1.915	1.947
	3	2.199	2.114	2.131	2.100	2.156	2.192	2.006	2.043
Valentia	7	1.648	1.600	1.622	1.580	1.647	1.645	1.630	1.615
	3	1.762	1.716	1.738	1.720	1.773	1.788	1.652	1.730
Cork Airport	7	1.530	1.483	1.500	1.436	1.503	1.506	1.478	1.480
	3	1.612	1.560	1.573	1.506	1.566	1.587	1.500	1.480
Dublin Airport	7	1.910	1.526	1.533	1.525	1.578	1.599	1.539	1.550
	3	1.883	1.602	1.605	1.606	1.663	1.683	1.607	1.615
Malin Head	7	2.219	2.149	2.164	2.127	2.188	2.231	2.113	2.147
	3	2.709	2.126	2.146	2.106	2.182	2.202	2.149	2.146
Casement	7	1.555	1.513	1.524	1.518	1.577	1.579	1.508	1.531
	3	1.681	1.668	1.651	1.579	1.603	1.660	1.599	1.564
Mullingar	7	2.873	1.389	1.398	0.981	0.994	1.019	1.408	0.969
	3	2.422	1.285	1.287	1.007	1.016	1.048	1.278	0.971
Average	7	1.956	1.659	1.672	1.589	1.642	1.659	1.656	1.606
	3	2.038	1.724	1.733	1.661	1.708	1.737	1.684	1.650

The best forecast in each row is in bold.

Table 2. COSMO and combined wind speed forecast RMSE (m s^{-1}). The best forecast in each row is in bold.

Station	COSMO 7 km	COSMO 3 km	ANNCOM	MSECOM
Belmullet	1.954	2.199	1.905	1.890
Valentia	1.648	1.762	1.588	1.554
Cork Airport	1.530	1.612	1.427	1.410
Dublin Airport	1.910	1.883	1.526	1.478
Malin Head	2.219	2.709	2.097	2.029
Casement	1.555	1.681	1.477	1.422
Mullingar	2.873	2.422	0.934	0.969
Average	1.956	2.038	1.565	1.536

bias. Average DMO bias over all stations varies from -1.64 to 2.47 m s^{-1} . All post-processing methods reduce this average bias to between -0.08 and 0.13 m s^{-1} . All of the methods are, therefore, considered to be effective at reducing model bias.

The adaptive post-processing methods are effective at reducing the errors for which they were designed. The STB forecast (and, indeed, every other post-processed forecast) reduces the overall model bias. The DRL forecast reduces the diurnal signal in forecast error. The LLS and KAL forecasts correct the linear relationship between forecast and observed wind speeds. The MAV forecast improves the match between the forecast and observed wind speed distributions. The DIR forecast reduces the dependence of forecast error on wind speed direction.

Bias scores can hide a balance between positive and negative errors, therefore the RMSE should also be considered. When comparing post-processing methods it has been found important to include a simple method in the comparison, such as the STB forecast used here. It is often difficult for other post-processing methods, even comparatively advanced ones, to improve on the RMSE scores achieved by the basic STB forecast. Although different methods are effective at reducing model RMSE at different locations and model resolutions, there is no single method that produces the best RMSE score for all cases.

Combining forecasts not only allows the performance of the method with the best RMSE to be automatically achieved at each station, but can also result in RMSE scores that are better than all of the available forecasts. The best overall forecast is produced by the MSECOM combined forecast: that gives better RMSE scores than any of its constituent forecast streams for 13 of the 14 test cases and equals the RMSE skill in the 14th case. The MSECOM forecast gives a 17% improvement in RMSE over the 7 km COSMO forecast, a 23% improvement in RMSE over the 3 km COSMO forecast, and keeps average bias below 0.1 m s^{-1} in all cases.

It is noted that the programming effort required to implement the post-processing schemes presented here is very small compared to that required to develop an NWP model. Moreover, the computational overhead is negligible compared to the computation required for the model integration. Therefore, the methods described in this study can yield substantial improvements in forecast accuracy at relatively small cost.

There are many different requirements that users may have from a forecast and many different ways of measuring skill. Warning systems for extreme events, for example, would require representation of outliers, and may be best served by probabilistic methods. This study considers deterministic

forecasts, and the bias and RMSE scores are used as indicators of forecast error. Future work will consider the benefit of post-processing and combining ensembles of forecasts to provide improved probabilistic forecasts.

Acknowledgements

This study is based upon work supported by the Science Foundation Ireland under Grant No. 09/RFP.1/MTH/2359. The authors also wish to acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support and Met Éireann for kindly supplying the observed wind speed data.

References

- Barthelmie R, Murray F, Pryor S. 2008. The economic benefit of short-term forecasting for wind energy in the UK electricity market. *Energy Policy* **36**(5): 1687–1696.
- Brown BG, Katz RW, Murphy AH. 1984. Time series models to simulate and forecast wind speed and wind power. *Journal of Applied Meteorology* **23**(8): 1184–1195.
- Cadenas E, Rivera W. 2010. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. *Renewable Energy* **35**(12): 2732–2738.
- Cheng WYY, Steenburgh WJ. 2007. Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States. *Weather and Forecasting* **22**(6): 1304–1318.
- COSMO. 2010. Cosmo Public area. The Consortium for Small-scale Modeling (COSMO). <http://www.cosmo-model.org> (accessed 1 July 2011).
- Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. 2008. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews* **12**(6): 1725–1744.
- Crochet P. 2004. Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland. *Meteorological Applications* **11**(2): 173–187.
- EirGrid. 2010. EirGrid Plc. Annual renewable report 2010. http://www.eirgrid.com/media/Annual_Renewable_Report_2010.pdf (accessed 1 July 2011).
- Galanis G, Anadranistakis M. 2002. A one-dimensional Kalman filter for the correction of near surface temperature forecasts. *Meteorological Applications* **9**(4): 437–441.
- Giebel G. 2003. The state-of-the-art in short-term prediction of wind power. Project ANEMOS. http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf (accessed 11 November 2010).
- Glahn HR, Lowry DA. 1972. The use of Model Output Statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology* **11**(8): 1203–1211.
- Kalman R. 1960. A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering* **82**: 35–45.
- Lei M, Shiyuan L, Chuanwen J, Hongling L, Yan Z. 2009. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews* **13**(4): 915–920.
- Li G, Shi J. 2010. On comparing three artificial neural networks for wind speed forecasting. *Applied Energy* **87**(7): 2313–2320.
- Louka P, Galanis G, Siebert N, Kariniotakis G, Katsafados P, Pytharoulis I, Kallos G. 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics* **96**(12): 2348–2362.
- Nielsen HA, Nielsen TS, Madsen H, Pindado M, Marti I. 2007. Optimal combination of wind power forecasts. *Wind Energy* **10**(5): 471–482.
- R Development Core Team. 2010. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna. <http://www.R-project.org> (accessed August 2011)
- Ripley BD. 1996. *Pattern Recognition and Neural Networks*. Cambridge University Press: Cambridge, UK.
- Salcedo-Sanz S, Pérez-Bellido A, Ortiz-García E, Portilla-Figuera A, Prieto L, Paredes D. 2009a. Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction. *Renewable Energy* **34**(6): 1451–1457.

- Salcedo-Sanz S, Pérez-Bellido A, Ortiz-García E, Portilla-Figueras A, Prieto L, Correoso F. 2009b. Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks. *Neurocomputing* **72**(4–6): 1336–1341.
- Sfetsos A. 2000. A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy* **21**(1): 23–35.
- Stensrud DJ, Yussouf N. 2005. Bias-corrected short-range ensemble forecasts of near surface variables. *Meteorological Applications* **12**(3): 217.
- Steppeler J, Doms G, Schattler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G. 2003. Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteorology and Atmospheric Physics* **82**(1–4): 75–96, DOI: 10.1007/s00703-001-0592-9.
- Sweeney C, Lynch P. 2011. Adaptive post-processing of short-term wind forecasts for energy applications. *Wind Energy* **14**(3): 317–325, DOI: 10.1002/we.420.
- Torres JL, García A, De Blas M, De Francisco A. 2005. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy* **79**(1): 65–77.
- Untch A, Miller M, Hortal M, Buizza R, Janssen P. 2006. Towards a global meso-scale model: the high-resolution system T799L91 and T399L62 EPS. *ECMWF Newsletter* **108**: 6–13.
- Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S*, 4th edn. Springer: New York, NY.