WIND SPEED MODELLING AND SHORT-TERM PREDICTION USING WAVELETS

Piers R.J. Campbell and Ken Adamson
Faculty of Engineering
University of Ulster
Shore Road, Belfast
Northern Ireland
Email: pr.campbell@ulster.ac.uk

KEYWORDS
Wavelets, Wind Energy, Prediction

ABSTRACT

The mathematical method of wavelets is explained and used to forecast wind conditions using short-term data collected at a site and referred to long term data from meteorological stations. We model the response time-series in terms of a multi-scale wavelet decomposition of the explanatory time-series. Preliminary results of this method, using hourly 10 minute averaged data from six locations in the British Isles, allow comparison with a linear regression method in terms of prediction errors over 21days.

INTRODUCTION

This paper examines the applicability of the mathematical method wavelets to the forecasting of wind speed for wind parks in the UK and Ireland. Wind energy is the fasting growing alternative fuel in the world today. Coupled with solar, hydro and biomass fuel the UK government hopes that 15% of the UK's energy will be generated by renewable energy by 2015.

A number of technologies have been applied to the forecasting or prediction of wind regime; these have included statistical approaches, Artificial Intelligence approaches and parallel computing. Of these a hybrid approach has to date produced the most satisfactory results, at least in terms of international application. Implementations of this technology such as WaSP have performed well when producing forecasts for specific geographical areas but have underachieved when applied to the UK and Irish markets. The reasons for the poor performance of these technologies may have to do with the different ornographic conditions in the UK/RoI region. In addition the systems implemented have been designed specifically for the weather patterns of the Scandinavian countries and may require redevelopment to achieve satisfactory results in the UK.

The wavelet approach does not require the significant computing power of the NWP technique but does require some historical data from the target location and also a more substantial quantity of historical data from a near by location. This is normally achieved as follows; a small amount of wind speed data is collected at a site (this is normal procedure when assessing a locations suitability as a wind park installation) and that a nearby reference location, for example, a meteorological station, provides long term data. The Wavelet approach then models the data at the target site, known as the *response time-series* (y_t), in terms of the data at the reference location, known as the *explanatory times series* (x_t).

The forecasting or predicting of future wind speeds at a target site using data from a reference location is generally known in the wind energy industry as "measure-correlate-predict", MCP.

Linear regression is a popular industry method for constructing the statistical model which will predict future values of a response time-series, (Derrick 1992). Other approaches such as Artificial Neural Networks also rely on historical data in order to produce forecasts. As an alternative, we propose a wavelet method which takes a multi-resolution approach. We model the response time-series in terms of a multi-scale wavelet decomposition of the explanatory time-series. We provide some preliminary results of this method using hourly, 10-minute averaged data from six locations in the British Isles and show how our model compares with a linear regression method in terms of prediction errors over 21 days. Typically the industry will require accurate predictions across a range of time horizons; short term (between 2 and 48 hours) for energy trading (due for implementation during 2005) and for longer term forecasts (2 to 10 days) for maintenance scheduling. For a more detailed discussion of the wavelet methodology refer to (Nason *et al* 2001).

Proceedings 19th European Conference on Modelling and Simulation
Yuri Merkuryev, Richard Zobel, Eugène Kerckhoffs © ECMS, 2005
ISBN 1-84233-112-4 (Set) / ISBN 1-84233-113-2 (CD)

A BRIEF INTRODUCTION TO WAVELETS

The following section is intended to give only a brief introduction the concept of wavelets, for a more detailed description we would suggest Burrus, Gopinath and Guo (Burrus *et al* 1998). Multi-resolution analysis provides the framework for examining functions at different scales. In a multi-resolution analysis a father wavelet, Φ(x), is a function constructed to approximate general functions to a certain scale by using shifted copies of itself. The mother wavelets, $\nu(x)$ derived from the father wavelets, represent the difference between father wavelet approximations at two different scales. A mother wavelet is a localised oscillating function from which a family of wavelets, $\nu_{j,k}(x)$, can be constructed by dilation and translation, i.e.

$$\nu_{j,k}(x) = 2^{j/2}\nu(2^j x - k) \qquad (1)$$

for integers *j, k*.

The dilation parameter *j* controls the scale (or size) of the wavelet and the translation parameter *k* controls the location of the wavelet. For suitable mother wavelets, $\nu(x)$, the set $\{\nu_{j,k}(x)\}_{j,k}$ provides a basis that can be used to approximate functions, i.e.

$$f(x) = \sum_j \sum_k d_{j,k}\nu_{j,k}(x) \qquad (2)$$

where $d_{j,k}$ are the wavelet coefficients.

Wavelets have enjoyed particular success in representing complex various types of complex signals and as a result have been implemented in domains such as image compression. Wavelets are particularly useful for representing signals with discontinuities due to their excellent localisation ability.

For some time-series wavelet packets may be of more use as they provide a wider choice of decompositions of the frequency domain. A wavelet packet is a particular linear combination of wavelets that retains many of the orthogonality, smoothness and localisation properties of wavelets (Wickerhauser, 1994).

Using wavelets (or wavelet packets) in our model allows us to attach a physical interpretation to the model. For example, a wavelet packet is given in Figure 1 which could be used to represent daily variation in wind speeds over the past 2.75 days. The wavelet packet provides valuable information about which components in the explanatory time series drive the response time series, i.e. which types of oscillatory behaviour in x_t influence y_t.

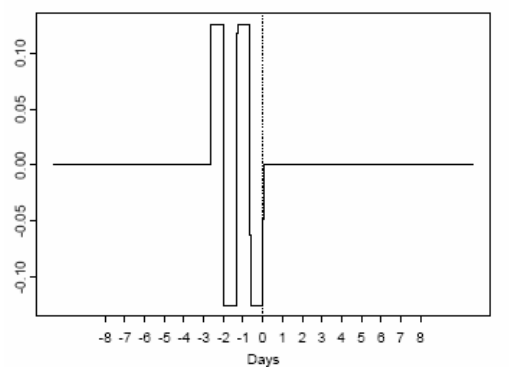

Figure 1: A Haar Wavelet Packet

A WAVELET MCP MODEL

To represent the explanatory time-series in terms of wavelet packets we compute a time ordered non-decimated wavelet packet transform. This calculates wavelet packet coefficients for the *j* scales and *k* locations in the analysis. These coefficients are stacked together to form a multivariate time-series matrix. Each variable in the multivariate time-series matrix corresponds to a particular wavelet packet and quantifies how similar the time-series is to the wavelet packet at each time point.

Principal Components Analysis (PCA) is performed on the multivariate time-series matrix because of some high correlations between the wavelet packets. The 3 resulting linear combinations of wavelet packets, the principal components (pc), are uncorrelated and are such that a few usually will explain most of the variation in the explanatory time-series.

We model the response time-series in terms of the principal components (pc). We assume that the residuals of the fitted model follow the normal distribution and this is reasonable in each example we have seen. We use the routines available in S-Plus for fitting a standard multivariate linear regression.

TEST LOCATIONS

In order to prove the validity of this approach three test sites have been selected on which to test the model. Each of the sites is an existing

wind park installation which will record the data required for result validation. The test data (response time-series) is a subset of the historical data recorded for that site and the reference data (explanatory time-series) is recorded by a meteorological mast or station at a variety of locations throughout the United Kingdom.

Site A; Met Mast 1

For our first example we use data from a wind park installation in northwestern England as the response time series and a meteorological mast in northern Wales as the explanatory time-series. The locations are approximately 85 miles (136 km) apart and the correlation between the time-series was 0.702. A graphical representation of the time-series data is given in Figure 2.

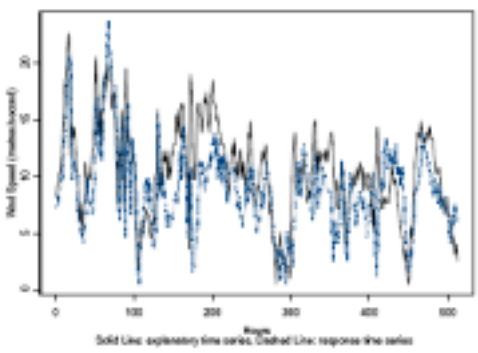

Figure 2: Time Series Plot of Site A

We perform the modelling procedure described in Section 3 to fit the model;

Site A ~ pc 1 + pc 5 + direction factor (3)

The 1st and 5th principle components (PC1 and PC5) were chosen for inclusion in the model.

We have chosen to include the first principal component (pc1) and the fifth principal component (pc5) in the model. The principal components represent linear combinations of non-decimated wavelet packet coefficients – so (3) expresses a statistical relationship between the response time series to the coefficients through the principal components. The direction factor is added to permit the statistical model to take account of the variability in wind direction.

The factor itself simply records the direction "bin" (the process of binning is the division of 360° into 12 distinct divisions of wind direction). To help us choose which components to include in the model we used a stepwise analysis to determine which components are most useful for the prediction of wind speeds at the target site. We adjust the estimate of the wind speed at the target site by a given direction factor. This is simply a constant that is determined by the direction of the wind at the reference location.

The first principal component is a mixture of father wavelets and wavelet packets. Father wavelets average wind speeds over a given time determined by the dilation level and translation index of the father wavelet. Wavelet packets capture a wide variety of high and low frequency oscillations in the data, the exact frequency determined by the dilation and

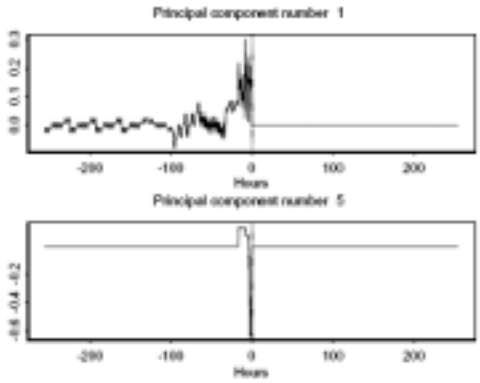

Figure 3: Principal Components from Site A Mast 1 model

translation index of the packet. The second principal component is made up of mother wavelets. Mother wavelets capture up to one oscillation over a given time determined by the dilation and translation index of the mother wavelet. The direction factor can have significant consequences in terms of swirl and wake effects for installations.

In Figure 3 we have plotted a graphical representation of the principal components. As an example, we suggest an approximate physical interpretation for principal component 5. In the first two hours there is a negative relationship between the two sites. As the sites are 136 km apart, unless the winds were extremely strong, we would not initially expect to see the same behaviour at the two sites. After two hours there is a gradual build up of the positive relationship between the two sites as wind speeds from the reference location

reach the target site. This continues up until 16 hours, by which time the weather system from Valley will have passed over St Bees Head.

Site B; Met Mast 2

For our second example we use data from a Wind Park in southern Wales as the response time-series and data from meteorological station (Met Mast 2) as the explanatory time-series. Met Mast 2 is approximately 87 miles (136 Kms) to the North of Site B and the correlation between the time-series is 0.762. A time-series plot is given in Figure 4. We explored various model relationships and the following model was found to fit well;

Site B ~ pc 10 + pc 41 + direction factor (4)

The forty-first component is similar to the first component picked out in the Site A example which represents a mixture of father wavelets and wavelet packets. The tenth principal component is a mixture of wavelet packets. These components are picking up more high frequency characteristics in the data than the components in the previous model.

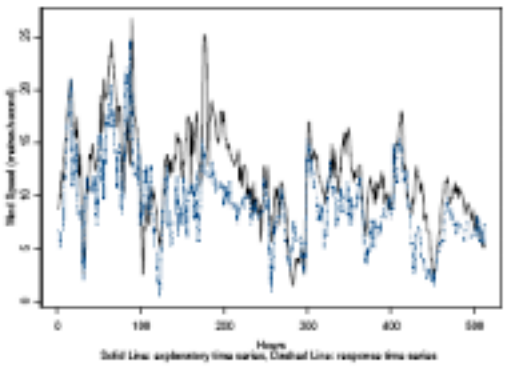

Figure 4: Time Series Plot of Site B

The direction factor is significant when the wind direction is from the Northwest, which we can again relate to site alignment, or from the Southwest, the direction of prevailing winds.

Site C & Met Mast 3

For the third test example we take Wind Park C on the Northwest coast of England as the response time-series and Met Mast 3 on the Southwest coast of the Isle of Man as the explanatory time-series. The two meteorological stations are approximately 45 miles (64 km) apart. The correlation between the two time-series is 0:697 and a time-series plot is shown in Figure 5. For this data we fit the model;

Site C ~ pc 1 + pc 4 + direction factor (5)

The principal components are similar to those illustrated in the first example. They pick out similar wavelets and wavelet packets that represent the major characteristics of the wind speed data but with some variation that may be specific to local conditions. The direction factor is significant for winds from a Southern to a Northwest direction encompassing both prevailing wind and site alignment direction.

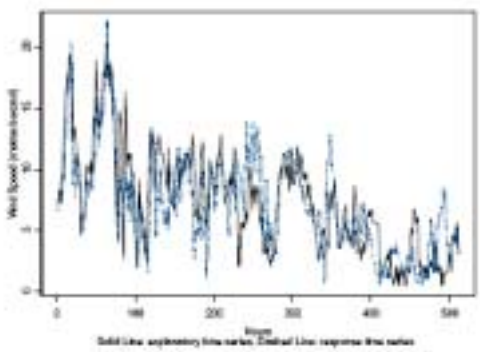

Figure 5: Time Series Plot of Site C

PREDICTING FUTURE WIND SPEEDS

We measure the accuracy of our method by looking at the mean residual sum of squares (MRSS) error of the predictions generated by our model. The models were fitted using a training dataset of 512 observations and a consecutive data run on 512 observations was used for the prediction. To evaluate our model we compare our predictions with those generated by the simple linear regression model.

It can be seen from Table 1 that the wavelets perform significantly better over the early prediction period. The results are closer for the longer prediction period as only a small quantity of data is used to formulate the models. We believe these preliminary results suggest our wavelet method is worth investigating as an alternative method for the short-term prediction of wind speeds.

CONCLUSIONS AND FUTURE WORK

We have proposed an alternative method for the prediction of wind speeds at a target site using wind speeds from a reference location. The preliminary results on relatively short time-series have been encouraging. Wavelet methods have been shown to provide more reliable estimates than a prediction method using simple linear regression over 21 days.

Our wavelet models also have the added bonus of often being physically interpretable. While it may be sensible to continue the study over a longer period of time it is important to remember the initial scope of the investigation; to assess wavelets ability to forecast wind speeds over a short term horizon, to satisfy trading requirements and over a longer term horizon for maintenance scheduling.

Table 1: Mean Residual Sum of Squares (MRSS) of the Predictions

Location	Forecast Horizon	Wavelets	LR
Site A & Mast 1	5 days	6.20	8.08
	10 days	5.40	6.89
	21 days	5.46	6.48
Site B & Mast 2	5 days	4.80	16.69
	10 days	4.36	12.74
	21 days	6.95	9.09
Site C & Mast 3	5 days	3.29	7.00
	10 days	3.59	6.57
	21 days	4.39	5.51

Wavelets show that they offer a significant improvement over linear regression and similar statistical approaches to wind spectra forecasting. It is intended to conduct further tests to assess their performance when compared to other leading techniques such as ARIMA, Moving Averages and Neural Networks, which have produced impressive results in recent studies (Campbell and Adamson 2004).

The authors would like to thank B9 Energy Ltd, Northern Ireland for their support of this project.

REFERENCES

Derrick, A. 1992. "The development of the measure-correlate-predict strategy 9 for site assessment". *Proc. 14th British wind energy association conference*, 259-265.

Nason, G.P., Sapatinas, T. , Sawczenko, A. 2001. "Wavelet packet modelling of infant sleep state using heart rate data". *Sankhya*, Series B, Vol. 63, 199-217

Burrus, C. S., R. A. Gopinath, & Guo, H. 1998. *Introduction to wavelets and wavelet transforms: A primer*. Upper Saddle River, NJ: Prentice Hall.

Wickerhauser, M. V. 1994. *Adapted wavelet analysis from theory to software*. Wellesley, Massachusetts.

Nason, G.P. & Sapatinas, T. (2002). "Wavelet packet transfer function modelling of nonstationary time series". *Statistics and Computing*, Vol. 12, 45-56.

Campbell PRJ, Adamson K, 2004. "A Comparison of Forecasting Methodologies for Wind Speed Time Series", *In Proc. Fourth International Conference on Modelling, Simulation and Optimization,* Kauai, USA, 211-218.