18th World IMACS/MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Modelling residual wind farm variability using HMMs

Ward, K.[1], **Korolkiewicz, M. and Boland, J.**

[1] *School of Mathematics and Statistics, University of South Australia, Australia*
Email: Kathryn.Ward@postgrads.unisa.edu.au

Abstract: The South Australian government has made a strong commitment to replacing the use of fossil fuelled energy with renewable energy. There are a number of wind farms in South Australia and approximately fifteen percent of the electricity currently used in South Australia is generated by wind farms. The topography of the sites chosen for the wind farms can make the energy produced from the farm challenging to predict.

There is now a large body of work on forecasting wind output generated by correlations with wind speeds at nearby meteorological sites. The wind speed is used as a predictor of wind farm output via power models that incorporate turbine energy losses. We take a different approach by using output data from the wind farms themselves and in this paper we consider just the residual variability in the data.

A hidden Markov model is used when we have a set of observations but know little of the process that controls them. We assume that there is a number of states that the observations can belong to. We also assume that there is a process which controls the switching of the states. This process is the 'hidden' part of the Markov model.

In this paper we present a hidden Markov model (HMM) for the residual variability in wind farm output. The HMM gives us an estimate of the state transition probabilities and the variability level. For an extensive discussion of hidden Markov models for when a Markov chain is observed as Gaussian noise please see Elliott (1993). This version of the HMM theory is dynamic and updates the estimates of the parameters as new observations arrive.

Using the HMM algorithm, we are able to learn about the states that our wind farm system takes and also the process controlling them. We learn that there is two states in our wind farm systems and that we can predict the effect that variability will have on our predictions. We find that there is always a high probability of staying in the current state for the wind farm data. This demonstrates the persistence that is apparent in the wind farm residuals. Most of the time we will remain in the state that we are currently in, so that if we know that we are in state 1 we have greater confidence in our forecasts.

We present here a series of results for the use of hidden Markov models with residual wind farm variability. We show how to capture the dynamics of the residual variability and how to determine the correct number of states for a system. We find that for each of the wind farms, having just two states of variability is the most appropriate model. This indicates two variability levels 'high' and 'low' capture the dynamics of the data well.

Keywords: *Wind Farm, Hidden Markov Model, Variability, Transition Probabilities*

1. INTRODUCTION

The South Australian government has made a strong commitment to replacing the use of fossil fuelled energy with renewable energy. There are a number of wind farms in South Australia ranging from the south-east of the state to the Eyre Peninsula in the state's mid-west. Approximately fifteen percent of the electricity currently used in South Australia is generated by wind farms.

Each wind farm has a number of turbines that are arrayed in a manner that best suits the location. The topography of the sites chosen for the wind farms can make the power produced from the farm challenging to predict. Big changes in wind flowing through the farms (i.e gentle breezes or gusting) can cause dramatic differences in the power output observed from the farms which makes forecasting quite difficult.

A large body of work has previously been performed on forecasting wind power (van Lieshout, 2004). Many models have been considered and the required forecast time horizon generally indicates the best approach to take (Giebel et al, 2003). Generating prediction intervals is the next evolution in confident wind forecasting. Both static and dynamic models have been examined previously (Dobschinski et al, 2008) with adaptive models, neural network approaches and linear regression on numerical weather predictors most common. Combining these models into an ensemble is so far the most effective way of modelling variability (Kariniotakis & Pinson, 2004). We investigate a different technique which comes from financial mathematics and is known as Hidden Markov Modelling.

Time series approaches have been shown to forecast wind power well in short time horizons (less than three hours) (Giebel et al, 2003) and we have used these models (nominally ARMA models) to predict the wind farm output (Ward & Boland 2007). In this paper, however, we are interested only in the residual variability, after the persistent effects have been removed.

Consider the residual variability for one of the wind farms. We have removed seasonality, cyclic and ARMA effects (Ward & Boland, 2007) as shown in Figure 1. We see that the level of the residuals is very low for a period of time but is then followed by a period of very high variability. We do not have random residual levels over the whole time series, but time dependent levels, so we need to model the variability separately.

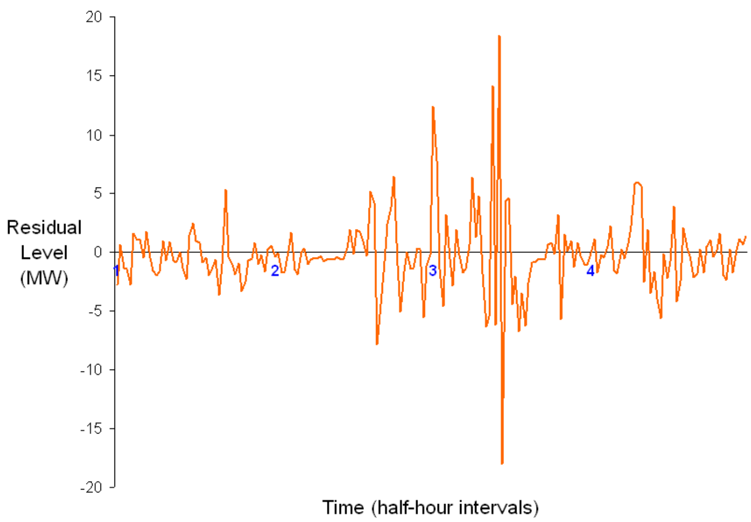


Figure 1. Residual variability of wind farm 1 over four days.

In this paper we present a hidden Markov model (HMM) for the variability in wind farm output. The HMM gives us an estimate of the state transition probabilities and the variability level. We begin modelling with two states 'high' and 'low' and introduce more states later in the paper.

2. THE HIDDEN MARKOV MODEL

A hidden Markov model (HMM) is used when we have a set of observations but know little of the process that controls them. We assume that there are a number of states which the observations can be derived from. We also assume that there is a process which controls switching between the states. This process is the 'hidden' part of the Markov model (see Figure 2) .

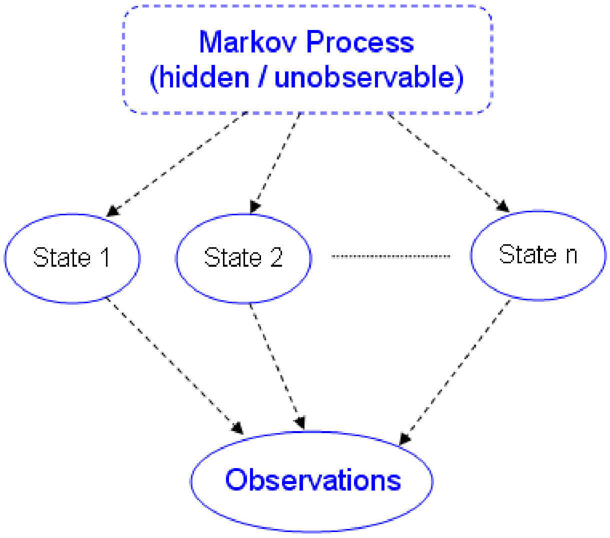


Figure 2. Depiction of an HMM: the Markov process and number of states are hidden; we observe a noisy series.

In this paper we present a hidden Markov model (HMM) for the residual variability in wind farm output. The HMM gives us an estimate of the state transition probabilities and the variability level. For an extensive discussion of hidden Markov models for when a Markov chain is observed as Gaussian noise please see Elliott (1993). This version of the HMM theory is dynamic and updates the estimates of the parameters as new observations arrive.

The basis of our HMM assumes that the residual wind variability is in one of two states - high or low - at any one time k. The observations we use are the residuals from the removal of the persistent parts of the wind farm output (Ward & Boland, 2007) and are denoted by y_k, $k = 1, \ldots, 30520$ (for two years of half-hour output observations). X_k, $k = 0, \ldots, 30520$ is the state that the wind farm residuals are in at time k. This setup assumes that the states have the form of the basis vectors $(0, \ldots, 0, 1, 0, \ldots 0)$ which makes calculations simpler (Elliott, 1993). That is $X_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or $X_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We have the state updating dynamics $X_{k+1} = AX_k + V_{k+1}$, where $E[V_{k+1}|\mathcal{F}_k] = 0 \in \mathbb{R}^2$, where $\mathcal{F}_k$ is the history of the system up to time k, as well as the property that $E\left[X_{k+1}|X_k\right] = AX_k$. Assume that

$$y_{k+1} = \sigma(X_k)w_{k+1} \tag{1}$$

is the observation dynamic, where $\sigma(X_k) = \langle \sigma, X_k \rangle$, $\sigma = \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix}$ represents the variability level of the two states and $\{w_k\}$ is an independent and identically distributed sequence of standard Normal random variables. As the model updates, the variability level and the transition matrix are updated. Thus, no knowledge of the original ‘high’ and ‘low’ variability level is needed, as this is estimated in the process.

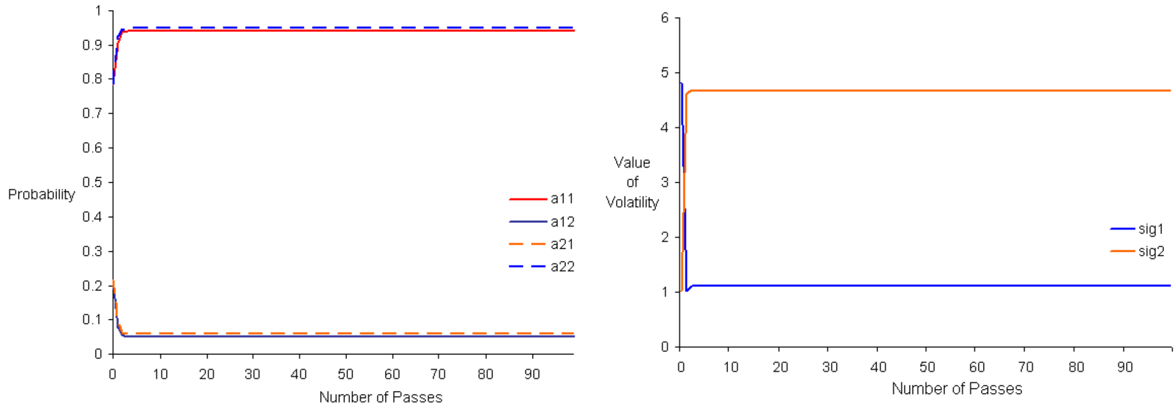
To begin our analysis we assume that the distribution of the wind farm output residuals is in one of two independent states. The measure change that takes place in the HMM algorithm gives observations is independent and identically distributed Normal variables. Thus we can assume that each state follows a Normal distribution with mean 0 and standard deviation σ as we have removed the level component and we are only interested in the variability component.

3. RESULTS

We present here the results obtained from running the HMM algorithm through various sets of residual data. We show the effect that over fitting or under fitting the number of states has on a two-state system. We also show an interesting phenomena observed in some of the wind farm data.

3.1. HMM for wind farm 1

We consider the wind farm data from the South Australian site whose residual variability is shown in Figure 1. To build our HMM model, we begin with the lowest possible number of states,2, and then test to see if that is inadequate by considering more states (see Section 3.2. We initially assume two states because we can see two distinct variability levels in Figure 3. We begin our HMM algorithm with initial states of $A_0 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ and $\sigma_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. These initial guesses are arbitrary and the algorithm soon forgets them when new observations are introduced. If we had some prior knowledge about the two states, we could introduce this to the initial guesses. We run the algorithm through 100 loops of the data which is again arbitrary - we just need to run through the data until convergence is achieved.



(a) Transition probabilities (b) Variability levels

Figure 3. HMM algorithm results for wind farm 1: transition probabilities and variability levels

Figure 3 shows the convergence of the transition probabilities and the variability levels for wind farm 1. We can see that the transition probabilities converge to $A_0 = \begin{bmatrix} 0.94 & 0.05 \\ 0.06 & 0.95 \end{bmatrix}$ and the variability levels converge to $\sigma_0 = \begin{bmatrix} 1.311 \\ 4.66 \end{bmatrix}$. If we initialise with other values, we will eventually arrive at these convergent results. Thus we have a very high probability of remaining in the state we are currently in.

We run the HMM algorithm over the first half of the observed data to obtain estimates of the transition probabilities and the variability levels. We then use these parameter estimates to forecast the state that the system is in at each time step. Figure 4 shows the residual values of this system with the forecasted states overlaid. We can see that when the variability is high in magnitude, we are in state 2, but when it is low in magnitude, we are in state 1. A possible explanation for the rapid fluctuations between states on some days in Figure 4 could be that this reflects the effects of the interaction of the wind with the local terrain, producing gusts.

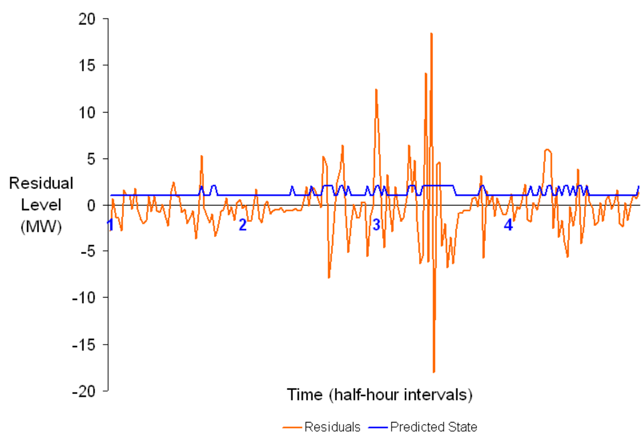


Figure 4. Half hour residual values over four days with estimated variability state (either 1 or 2) overlaid.

3.2. More than two states

We consider a system that has been built with two states. We want to know the effect of fitting more than the appropriate number of states for the model. That is, we want to check if there are actually more than two discernible states embedded in the output (or in the system).

We construct a system with known transition probabilities and variability levels. We have two states with the state transition probabilities such that $A = \begin{bmatrix} .95 & .40 \\ .05 & .60 \end{bmatrix}$ and variability levels $\sigma = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$. Using these values, we build a time series of 1000 'observations' by simulating a set of state changes and then generating a value from $N(0, \sigma_i)$ for state i.

We run the HMM algorithm with arbitrary initial values $A_0 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ and $\sigma_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. The algorithm updates the transition probabilities and variability levels very quickly, so that the initial guesses soon become irrelevant.

We obtain the solutions for the two state model as $\hat{A}_{\text{2-state}} = \begin{bmatrix} .94 & .38 \\ .06 & .62 \end{bmatrix}$ and $\hat{\sigma}_{\text{2-state}} = \begin{bmatrix} 1.02 \\ 4.81 \end{bmatrix}$. We can see that these are very good estimates of the true values and they give us a base line to compare our following results to.

Now we see what happens when we try to describe this system using three states which would coincide with variability levels of 'low', 'high' and 'very high'. We run the HMM algorithm with initial values of $\{a_{ij}\} = 1/3,\ i = 1, 2, 3, j = 1, 2, 3$ and $\sigma = [1\ 2\ 10]^T$.

We obtain the solutions to the three state problems as $\hat{A}_{\text{3-state}} = \begin{bmatrix} 0.94 & 0.38 & 0 \\ 0.06 & 0.62 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $\hat{\sigma}_{\text{3-state}} = [1.02\ \ 4.81\ \ 1.38]$. We can see from the bottom row of $\hat{A}_{\text{3-state}}$ that the probability of being in state 3, given we were in any other state is zero. This means that our time series will never be in state 3. We also notice that the transition probabilities for the first two states and the variability level estimates are identical to our two state model. Thus, this third, superfluous state has dropped out and does not affect the parameter estimates of the other states.

When we consider four states a similar outcome appears. We again start the algorithm with the very uniform $\{a_{ij}\} = 1/4,\ i = 1, \ldots, 4, j = 1, \ldots, 4$ and the variability levels of $\sigma = [1\ 2\ 5\ 10]^T$. We run the HMM algorithm and at convergence we obtain $\hat{A}_{\text{4-state}} = \begin{bmatrix} 0.94 & 0 & 0 & 0.38 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ .06 & 1 & 1 & 0.61 \end{bmatrix}$ and $\hat{\sigma}_{\text{4-state}} = [1.02\ \ 1.37\ \ 1.38\ \ 4.81]$.

Once again we see that the additional two states have become absorbed. We see that in the bottom row of $\hat{A}_{\text{4-state}}$ the probability of being in state 4, given that we are in state 2 or 3, is 1. However rows 2 and 3 show that the probability of being in state 2 or 3 is zero. Thus we never use state 2 or 3. This problem once again reduces to the two state problem and the superfluous states are essentially ignored by the HMM algorithm.

3.3. Less than two states

We have shown the results of trying to fit too many states to a system that actually has just two states. We want to know want the results of the HMM algorithm would be if we didn't have multiple states at all (remember that we are using real life data and it is sometimes difficult to tell how many states will be needed).

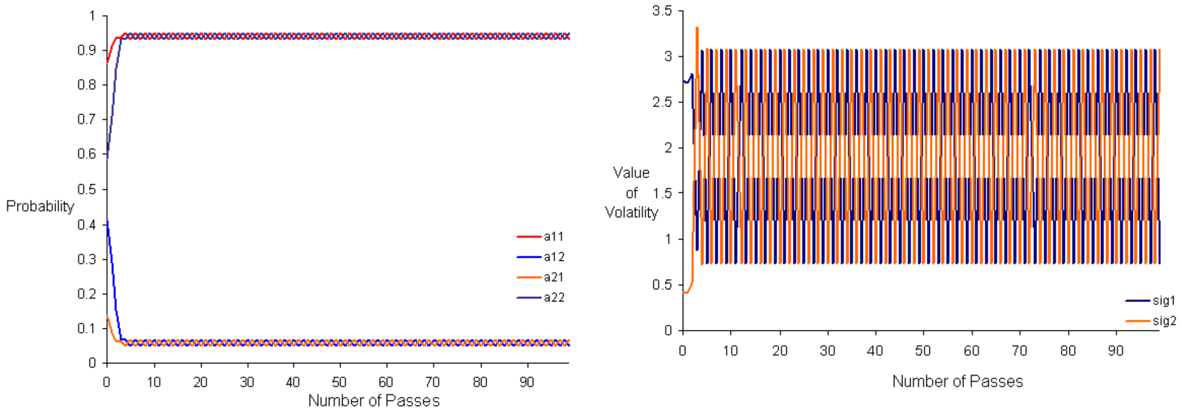
We build a time series that has just one variability level over the length of the series. We then run a 2-state HMM using the initial conditions seen in Section 3.1. After 100 loops through the algorithm, we obtain the following 'convergent' results:

$$A = \begin{bmatrix} .59 & .49 \\ .41 & .51 \end{bmatrix} \qquad \sigma = \begin{bmatrix} 1.039 \\ 1.039 \end{bmatrix}.$$

We see that the transition matrix has not deviated a long way from the initial conditions, and that the two variability levels are the same. The chance of switching states is practically 50-50. The algorithm has detected that there is just one state so it has given the variability level, and essentially the transition matrix is meaningless.

3.4. Oscillating states

Consider Figure 5 which shows the transition probabilities and variability levels for wind farm 2. The HMM algorithm is initialised in the same manner as in Section 3.1.



(a) Transition probabilities

(b) Variability levels

Figure 5. HMM algorithm results for wind farm 2: transition probabilities and variability levels

We see here that the state probabilities and the variability levels are oscillating. Each of the two states has a probability profile (a column of the A matrix) and an associated variability level. The association between the transition probabilities and the variability level does not change with each oscillation. Instead, it is the label that changes. In this case, at loop 10 state 1 has the transition probabilities $\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix}$ and $\sigma = 3.08$ and state 2 has transition probabilities $\begin{bmatrix} 0.93 \\ 0.07 \end{bmatrix}$ and $\sigma = 0.74$. When we cycle through to loop 11 we obtain state 1 with transition probabilities $\begin{bmatrix} 0.93 \\ 0.07 \end{bmatrix}$ and $\sigma = 0.74$ and state 2 has transition probabilities $\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix}$ and $\sigma = 3.08$. The nature of the states remain the same but the label changes.

4. ANALYSIS OF RESULTS

Knowledge of the dynamics of the states gives us valuable insight into the wind farm variability. We are able to forecast the state that our system is in. This allows us to determine the level of confidence we have in a wind farm variability forecast. If we are in the ‘low’ state then we know that our prediction should be reliable as the variability is low. If we are in the ‘high’ state though, we will know that the system is in a very volatile stage which makes accurate prediction difficult.

Through the HMM algorithm we can also predict the level of the variability. Thus, even though we know we have two states, we also know how the variability will have a relative effect on our wind farm output predictions. If we have two states that have variability 1 and 10 then we know that we will have trouble with accuracy when the system is in state 2. We use the predicted variability level to build a confidence interval for the predicted values of the deterministic components. If the relative variability is low (like in state 1) then the confidence interval is narrow and we have greater confidence in our prediction. If the variability is high, however, our confidence interval will be very wide and we can’t be confident in the prediction.

We can see from the level of the variability the differences between the two states. We see that the ‘high’ state is at least three times as volatile as the ‘low’ state which is consistent with our other farms (results not published in this paper). This shows us that the ‘low’ state is considerably more reliable to make forecasts in than the ‘high’ state.

We find when we run the HMM algorithm through the data for our wind farms, that there is always a high probability of staying in the current state. This demonstrates the persistence that is apparent in the wind farm variability. Most of the time we will remain in the state that we are currently in, so that if we know that we are in state 1 we can remain confident in our forecasts.

Getting the correct number of states is important in order to forecast with accuracy. With the HMM algorithm we can obtain the most appropriate number of states. If we should only have one state, the algorithm will conclude that the variability levels are equal and thus we know we only need one state. If we try to fit too many states, we will obtain a resultant transition matrix where the probability of getting into the extra states is zero (within reason - some noise is to be expected in real world data). Thus we should always be able to obtain the correct number of states, their transition probabilities and the corresponding variability levels.

We find, through the HMM algorithm, that for each wind farm, the most appropriate number of states is two. This could be caused by the control process used at the individual turbines. Two different control mechanisms are used at the turbine depending on the wind speed. When the wind speed is low, the control process aims to obtain as much electricity as possible. When the wind speed is high, the control process aims to regulate (steady) the electricity production as much as possible. This two tiered system may account for the results obtained.

We observe the phenomena of oscillating states in a number of our wind farms. Changing initial conditions does not effect the outcome of the algorithm (it soon forgets the initial conditions). We do not see any of the characteristics of too many or too few states in the results. The farms that oscillate are not generally the biggest or smallest farms, do not generally have the highest variability and are not by any statistical measure different from the farms that do not oscillate. This is an interesting finding that warrants further investigation.

5. CONCLUSIONS

We presented here a series of results for the use of hidden Markov models with wind farm variability. We showed how to capture the dynamics of the variability and how to determine the correct number of states for a system.

We find that for each of the wind farms, two states is the most appropriate model. This indicates two variability levels ‘high’ and ‘low’ capture the dynamics of the data well. Each new time step has a high probability of remaining in the current state, with few jumps noticeable.

If designed correctly, the variability of a network of wind farms should be less than the variability of one wind farm. Across South Australia, the wind farms are placed so that their production does not rise and fall at the same time thus making the output more steady. Future work will involve investigating the variability over this network.

Another avenue for future work is to investigate the grid links between South Australia and Victoria and New South Wales. At the moment there is only limited transmission between South Australia and the rest of the grid. If the transmission links were improved, more wind electricity could be produced as the variability would be spread over a greater region.

We have presented a new way of forecasting wind farm variability. This methodology is state-dependent and will be dynamic over the observations of the time series. Hidden Markov models can be used to estimate state changes and variability levels and thus give us more confidence in our forecasts.

REFERENCES

Dobschinski, J. et al. (2008) *Estimation of Wind Power Prediction Intervals using Stochastic Methods and Artificial Intelligence Model Ensembles*. Instutut fur Solare Energieversorgungstechnik, Kassel, Germany.

Elliott, R.J., (1993) New Finite-Dimensional Filters and Smoothers for Noisily Observed Markov Chains, *IEEE Transactions on Information Theory*, 39(1), 265-271.

G. Giebel, R. Brownsword, and G. Kariniotakis., (2003) State of the art on short-term wind power prediction, *ANEMOS Report D1.1*, available online: http://anemos.cma.fr.

Karianiotakis, G.N. & P. Pinson (2004). *Uncertainty of Short-term Wind Power Forecasts - a methodology for online assessments*. 8th International Conference on Probabilistic Methods to Power Systems, Iowa State University, Ames, Iowa.

van Lieshout, P., (2004) Predicting Wind Output: Time dependent energy calculations, *Refocus*, 5(3), 40-43.

Ward, K. and Boland, J., (2007) Modelling the Volatility in Wind Farm Output, *MODSIM 2007 International Congress on Modelling and Simulation*, Christchurch, N.Z., Modelling and Simulation Society of Australia and New Zealand, 1858-163.