Advanced statistical methods for short-term wind power forecasting
Research proposal draft

Alex Barnett

July 2001

1 Background

Over the last decade wind power has become a cost-effective alternative to non-renewable energy sources, with the potential to reduce pollution on the worldwide scale that will be required in the near future (Brower 2000). However, wind is an intermittent resource, and if it is to achieve significant penetration into the market it must not reduce the utilities' ability to meet electricity demand reliably, nor the grid's ability to handle the fluctuations. Accurate forecasts of wind turbine power generation on short-term timescales of 1–48 hours into the future are therefore of great value both to utilities and to the environment (Milligan 1995).

Forecasting methods have generally fallen into two categories: 1) extrapolation of a single variable (the wind speed, or power output at a turbine) using linear or nonlinear time-series analysis (Alexiadis 1999), or 2) forecasting wind velocity using regional ("mesoscale") numerical weather models initialised from national weather data (TrueWind 2001, Beyer 1999). The latter is far superior at providing forecasts beyond a couple of hours, and is becoming standard in the industry. However, due to computational limitations, mesoscale models can only provide wind velocity predictions averaged over grid cells of several kilometers in size. In practise, due to the turbine location, height, and the local topography on length scales down to a few tens of meters, there are systematic errors between the model forecasts at the local grid cell and the observed wind velocity at the turbine. Correction of these errors has been handled by simple linear rescaling (Beyer 1999), or by a more successful linear regression using hand-picked variables (TrueWind). My interest is in improving and automating this fitting stage using more flexible neural network methods from the machine learning (AI) community.

Neural networks have successfully been applied to "downscaling", a similar problem of learning local environmental variables from larger-scale gridded variables which arises in climate modelling (Sailor 1999). In preliminary research I have applied Bayesian neural networks using a promising Monte-Carlo training scheme (Neal 1995) to data from the wind-energy industry, with some success (Report for TrueWind 2000). However this work has served mainly to highlight the questions which deserve investigation before a practical and robust automatic scheme can be put into use.

2 Directions

Research would focus on the general issue of prediction of the future of a time-series A given a potentially correlated time-series B whose future is known. Classical linear techniques are known (Box, Jenkins and Reinsel 1994), but more powerful nonlinear techniques have received little attention (MacKay 1995). In the wind-energy application, A is the desired wind speed (or power output) at the turbine, and B is the numerical weather model variables. Training the neural network for a prediction involves learning the (nonlinear) map from inputs B to output A. (In fact this training is performed using many past versions of B corresponding to weather model runs used for previous prediction attempts).

Testing would be performed on both machine-generated time-series which embody successively more real-world (*i.e.* non-idealised) features, and on actual data from the wind power industry. Robustly handling unexpected and missing data is a necessity. As well as these empirical approaches, I will be developing the theoretical underpinnings to the forecasting problem, within a Bayesian framework wherever possible.

Some of the particular questions I plan to address are:

1. Reliable improvement over the linear regression performance. Currently the neural networks can do better but can on occasion do much worse. Creating a training scheme which cannot do worse than linear regression is a priority.

2. Selection of training data from a time-series when there are rare events whose occurrence it is important to learn (for instance high wind-speeds responsible for the peak output power). This involves a balance between using most recent training examples and sufficient examples of rare events from the more distant past. It is related to the general problem of choosing training data which samples the input space well. This would also address seasonal (annual) changes, whose effect on local weather patterns is significant.

3. How should the *crossover* from extrapolation of series A (method 1 above) to prediction using series B (method 2) be best handled? This crossover is expected to happen beyond a certain characteristic timescale, related to time-correlations in series A and the strength of A-B correlations. In the wind application, the timescale seems to be a few hours, however until now no work has been done to analyse how a system could optimally combine both these approaches.

4. Development of other *noise models*. In the real world, the prediction error is nonconstant (there is more uncertainty in high than low wind speeds), and this has a detrimental effect on the neural network's ability to learn. Robust ways to handle this need to be developed. Also in standard neural network models, only the output is assumed to be noisy. In the wind application, it is the inputs which contain the dominant noise, and a noise model could Furthermore the noise has correlations in time whose explicit inclusion would improve performance.