Probabilistic Wind Farm Group Forecasting using Bayesian Model Averaging

Pål Preede Revheim, Hans Georg Beyer
Dept. of Engineering Sciences
University of Agder, Grimstad, Norway

Outline

- Introduction
- Wind power forecasting
- Data
- Bayesian model averaging (BMA)
  - Choice of component PDFs
  - Estimation of component PDF parameters
  - BMA weights
  - BMA predictive PDF
- Results
- Conclusions
- Remaining challenges and opportunities

Introduction

- For operational purposes forecasts of the lumped output of a group of wind farms spread over a larger geographic area will often be of interest
  - Spatial smoothing effects
  - Power system management

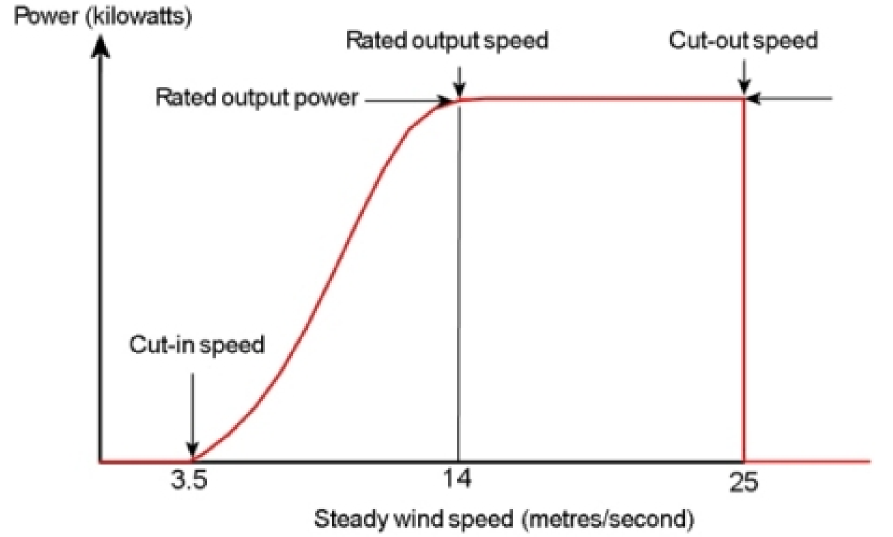
- Increasing interest in probabilistic wind power forecasts
  - Power system management
  - Electricity trading

- Bayesian model averaging (BMA) is a statistical post-processing method for producing probabilistic forecasts from ensemble information
  - Aim at making probabilistic forecasts for the lumped output of the ensemble members

- Normally BMA uses the outcome of various models as ensemble. Here we apply a BMA scheme treating the contributions of single sites to form the ensemble of the area-wide lumped production

Wind power forecasting

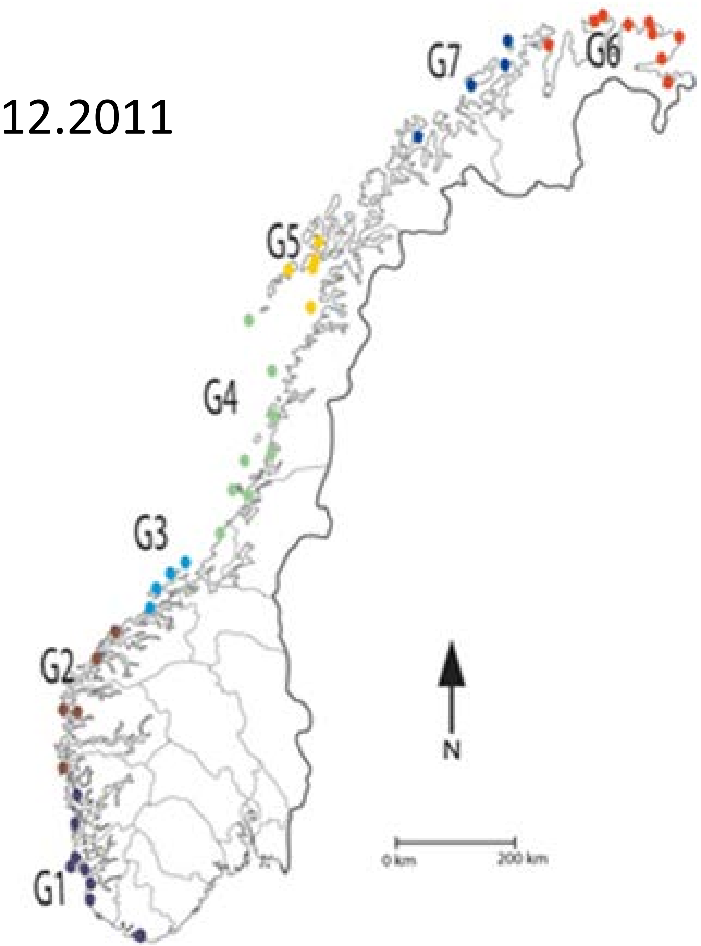
- The transformation from wind to power is highly non-linear
- Typically represented by a power transformation curve
  - Lower cut-in speed (~ 3.5 m/s) and upper cut-out speed (~25 m/s)
  - Flat lower and upper part, steep middle part
  - The lower and upper cut-out speeds causes the wind power production to be zero much more frequently than what is the case with the wind speed



Typical wind turbine power output with steady wind speed.

Data

- Data from time period 01.01.2009 to 17.12.2011
- 43 weather stations run by Met.no
  - 10 m height
  - Hourly updates
- Hirlam 4*4 km forecasts
  - Hourly forecasts
  - Updated every 24 hours
- Height transformation
- Transformation to power output
- Data grouped into 7 groups
- All forecasts in examples in this presentation is for a 24 hour forecast horizon



Location of the 43 data points and the 7 groups created through a clustering procedure

Bayesian Model Averaging (BMA)

- BMA can be used to produce PDFs of future group wind power from the group members forecasts

- Each group member forecast is associated with a component PDF, $g_k(y|f_k)$

- The BMA predictive PDF of the future wind power is a combination of the component PDFs

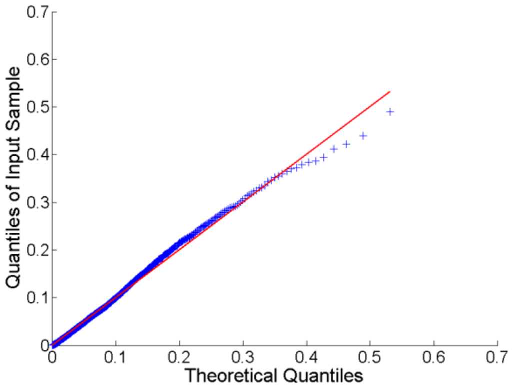
$$p(y|f_1, \ldots, f_K) = \sum_{k=1}^{K} w_k g_k(y|f_k),$$

where w_k is the weight put on the forecast for group member *k* based on its performance in a training period (here 30 days)

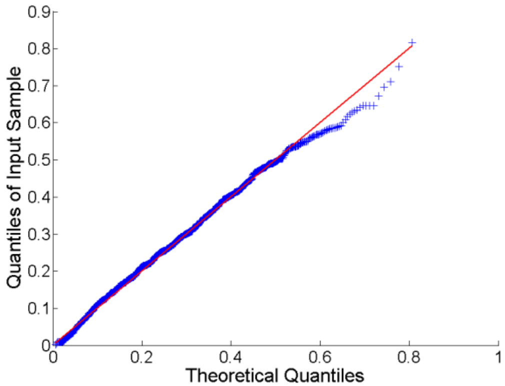
BMA – Choice of component PDFs

- For normalized wind power beta-distribution is a good choice of component PDF
  - Limited to [0,1]
  - Shows a good fit with the observed data (quantile-quantile plotts for wind power between 0%-10%, 40%-50% and 90%-100% shown below)
  - Parameters can easily be estimated form data

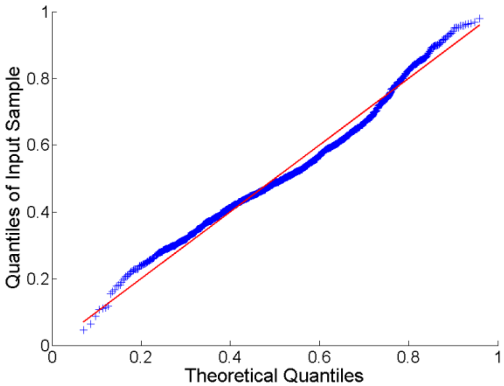
- Beta PDF: $g(y;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}$, where $0 \le y \le 1$ and $\alpha, \beta > 0$



0% - 10% norm. wind power



40% - 50% norm. wind power



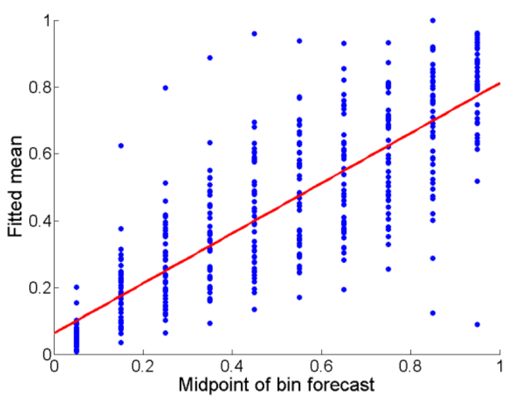
90% - 100% norm. wind power

UNIVERSITY OF AGDER

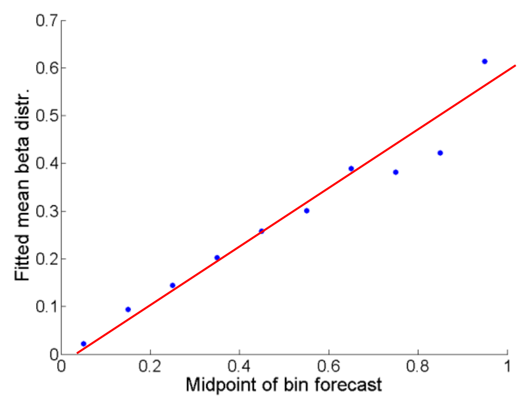
BMA – Estimation of component PDF parameters

- The parameters of the beta-distribution are estimated from the data
  - The means by linear regression for each site on the data from the training period
  - For the standar deviations regression with a quadratic term from the data from all site seemed promising, but the standard deviations were later found to be too noisy
  - For the standard deviations the regression was replaced by a per-site lookup table

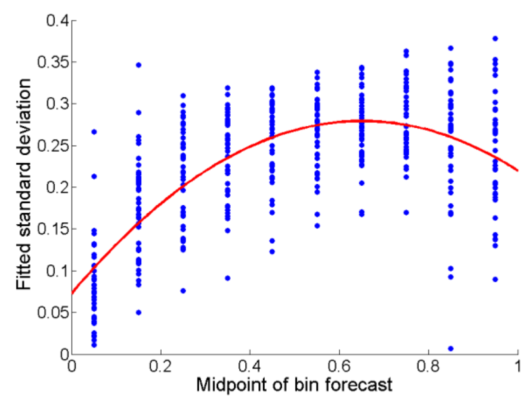
- The beta distribution parameters are calculated from the means and st. dev. by:
  - $\alpha = \left(\frac{1-\mu}{\sigma^2} - \frac{1}{\mu}\right) * \mu^2$ and $\beta = \alpha * \left(\frac{1}{\mu} - 1\right)$



Fitted mean values all sites



Fitted mean values single site



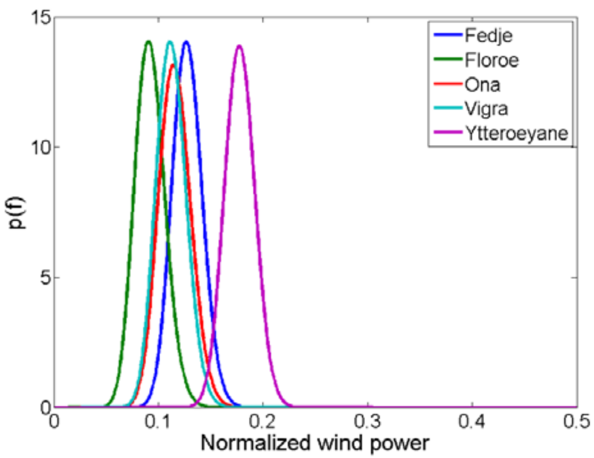
Fitted st. dev. All sites

BMA – BMA weights

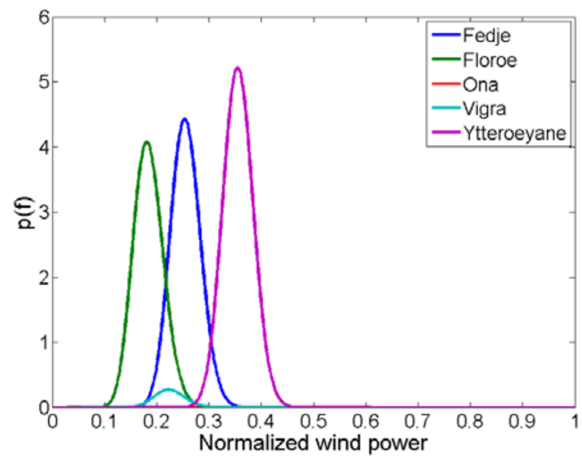
- The single-site component PDFs with parameters estimated from the training period are weighted by their relative contribution to forecast skill in the training period
- Weights are found by maximum likelihood, minimizing the log-likelihood function

$$\ell(w_1, \dots, w_k) = \sum_t \log\left(\sum_{k=1}^{K} w_k g_k(y|f_k)\right),$$

where t is the number of observations in the training period, k is the number of single sites and g the component beta PDF.



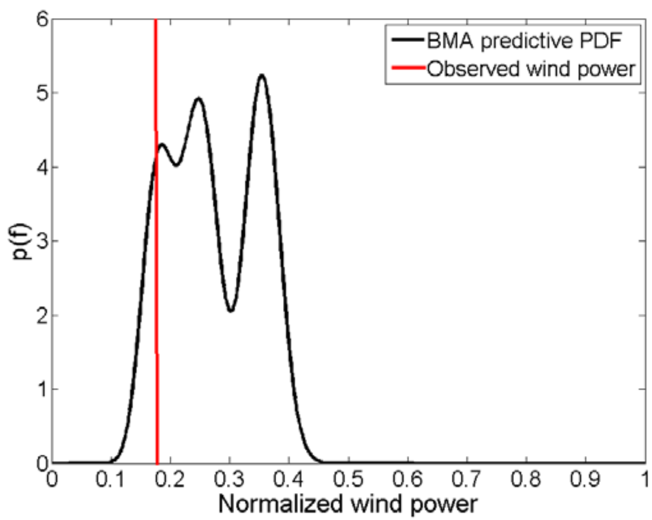
Unweighted component PDFs



Weighted component PDFs

BMA – BMA predictive PDF

- The BMA predictive PDF for the future wind power is found by summation of the weighted component PDFs, $p(y|f_1,\ldots,f_K) = \sum_{k=1}^{K} w_k g_k(y|f_k)$
- As a result of the summation it can be rather oddly shaped
- The BMA PDF can be used for e.g. making confidence intervals of the future wind power, make estimates of quantiles etc.

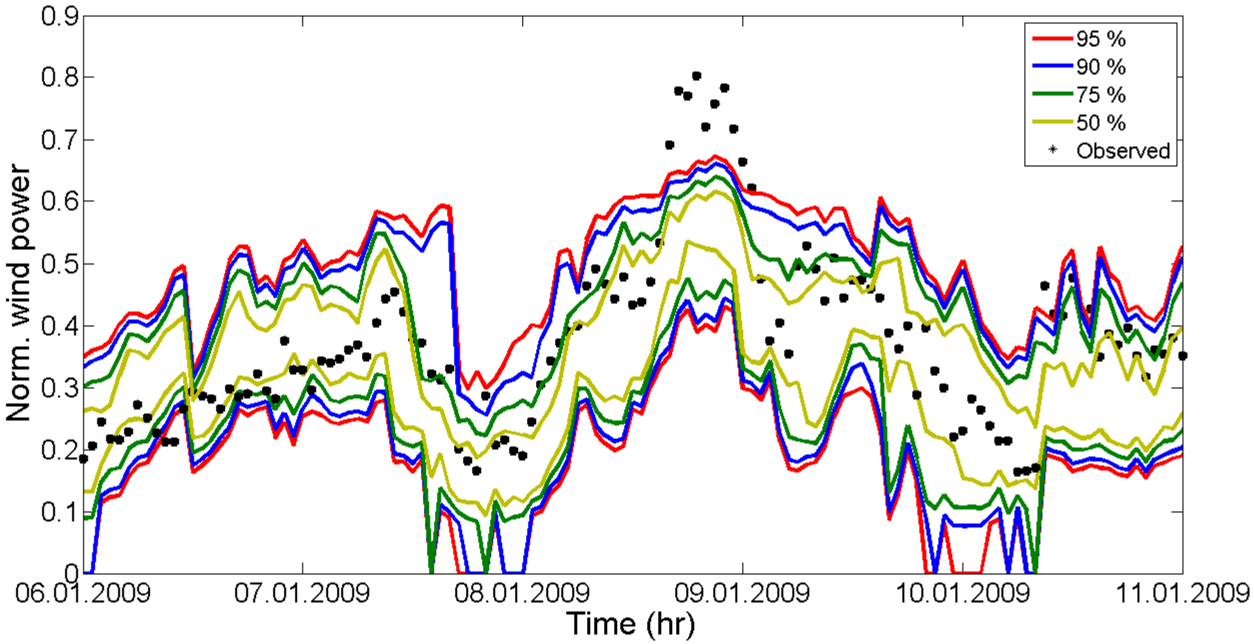


BMA predictive PDF for the wind dower of Group 2 for one sample hour

Results

- Model evaluated by
  - It's ability to produce accurate confidence intervals for different levels
  - The RMSE of the P50 estimate of wind power production

- Sample confidence interval for five days in January 2009 for Group 4 shows good coverage except for at the peak 08-09 January.



Results

- The table show the shares of observations covered by different level confidence intervals
  - «Low» – group with lowest coverage
  - «Mean» – mean coverage for the 7 groups
  - «High» – group with highest coverage)

Conf. level	24 hour forecasts		
	Low	Mean	High
95 %	84 %	86 %	90 %
90 %	80 %	83 %	87 %
75 %	70 %	73 %	75 %
50 %	51 %	56 %	61 %

- For 6 out of 7 groups the BMA P50 estimate for wind power has a lower RMSE than the combined NWP-forecasts (but only by an average of 1 percentage point)

Conclusions

- Based on the beta distribution a BMA PDF produces fairly accurate confidence intervals for the future lumped wind power production of groups of wind farms
- The mean of the beta distribution can be estimated from training data by linear regression.
- Different approaches for estimating the standard deviation of the beta distribution were tested. To create site-specific lookup-tables found to be the best functioning.
- The beta distribution is not capable of modeling the high number of zero-production observations found in the data
  - Here this is solved through setting a lower threshold value for when the lower confidence limits
- The BMA PDFs for the larger groups (number of sites) is found to be more accurate than for the smaller groups
  - Mainly caused by lower shares of zero-production

Remaining challenges and opportunities

UNIVERSITY OF AGDER

- The problems with zero-values could be solved more efficiently
  - One option would be to have piecewise component PDFs
  - Another to make forecasts for larger groups, thus limiting the share of zero-productions
- The lookup-table approach for the beta standard deviations is not ideal
  - Makes forecasts less time-adaptive
- For meteorological applications BMA is most commonly applied to NWP ensemble forecasts. Would be interesting to test this kind of data for wind power forecasting.
- The possibilities for using the BMA weights w_k to identify sample sites giving a good representation of a larger group will be investigated further.

Thank you for the attention!