Poster Abstract: Combining multiple forecast for improved day ahead prediction of wind power generation

Milan Jain
IIIT Delhi, India
milanj@iiitd.ac.in

Amarjeet Singh
IIIT Delhi,India
amarjeet@iiitd.ac.in

ABSTRACT

Wind, a major alternative source of energy, provides dynamic output due to frequent weather changes, which introduces one of the biggest challenges in integrating it with the existing power system. Commercial wind power forecasters vary in their prediction accuracies both across the wind farms and for different time periods within a farm. Therefore, wind power generators (WPGs) employ multiple such forecasters and heuristically choose day-ahead-prediction from one of them (baseline model). In this work, we combine multiple forecasters to generate a superforecast for the day-ahead-prediction which is, expected to be better than individual forecasters in terms of penalty – the cost a WPG has to pay for inaccurate predictions. Performance evaluation using 6 months of SCADA and forecaster data, from a WPG, of a wind farm located in India, shows that superforecast reduced the penalty by 7% and 13% when compared with the least penalised forecaster for each month and the baseline model.

1. INTRODUCTION

Wind power production capacity reached 336GW (approx. 4% of the total electricity demand) by the end of June 2014 worldwide[1]. A single wind farm consists of multiple wind turbines that are expected to follow a standard power curve, as shown in Figure 1. However, each turbine deviates from the ideal power curve due to multiple reasons, including random variations in the wind, that are complex to model. Such deviations at the turbine level result in significant variation in power produced at the wind farm level complicating their integration with the existing power system.

Across the world, there exist regulations that require wind power generators to submit prediction for each 15-minute window (09:00-09:15, 09:15-09:30, so on) for the next day (day-ahead-prediction). These predictions are then used for appropriate supply side management by the electricity distribution centers. However, due to weather based deviations, even commercial wind power forecasters are not able to accurately perform day-ahead-prediction. For some countries, there exist incentives and penalties for correctly and incorrectly predicting the day-ahead-generation. To optimise on such incentives and penalties, often wind power generators employ multiple wind power forecasters to provide the day-ahead-predictions and then choose the best forecaster heuristically. In this paper, we present some initial insights from a real world wind farm dataset along with a basic model to generate superforecast, combining multiple commercial forecasters. Proposed model is evaluated using real data from a wind farm in south-east India that includes `SCADA` data from the farm and forecaster data from 4 commercial forecasters employed by the generator. We evaluate our approach in terms of the penalty that a generator would pay in comparison to the least penalised forecaster for each month and the baseline model (currently adopted by the generator to heuristically select one of the forecasters). Proposed model shows an improvement of approx. 7% and 13% (averaged over six months) in terms of penalty when compared to the monthly least penalised (MLP) forecaster and the baseline model, respectively.

[1]http://goo.gl/xSj50E

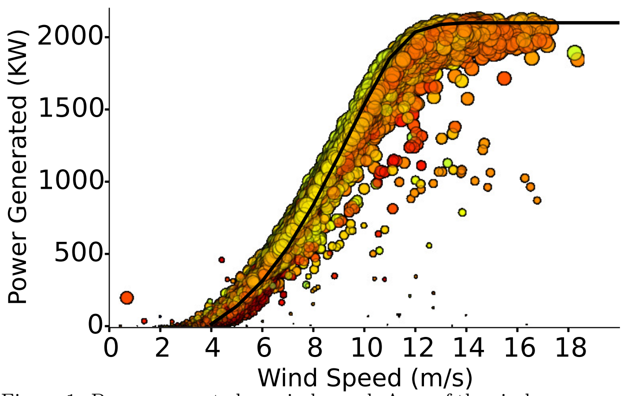


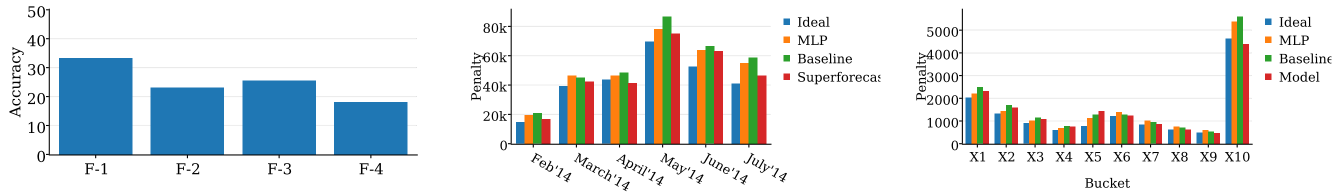
Figure 1: Power generated vs wind speed; Area of the circle represents average time per 10 minutes for which each turbine was operational; color indicates environmental temperature with darker shade representing higher temperature.

2. STATE OF THE ART

The current state of the art in wind power forecasting includes deterministic and probabilistic forecasting methods that can provide short, medium and long-term predictions. Deterministic forecasting methods generate spot predictions i.e. a single value of future wind output. In accordance with a survey by Zhang et al. [3], many of these prediction approaches have inherent and irreducible uncertainty. On the other hand, probabilistic forecasts usually take the form of a probability density function (PDF) and can be categorized into parametric (wherein the shape of PDF is assumed) and non-parametric (wherein PDF is estimated at a finite number of points and its full description is obtained by interpolation). Various wind power forecasters work on these existing techniques and generate predictions for wind power generators. Randomness in wind behavior leads to variations in their prediction accuracy, thus motivating research in hybrid approaches[1] and adaptive combination[2] of forecasts. Treating each of the forecasts as a time series, this work presents a simple regression based approach to combine predictions from multiple forecasters and generate a superforecast.

3. MODELING AND EVALUATION

`SCADA` and forecaster data used for evaluation was spread over 6 months from Feb'14 to July'14 and was collected from a wind farm located in the state of Andhra Pradesh



(a) Percentage of time a forecaster is most accurate across all 15 minute blocks

(b) Monthly comparison of superforecast with Ideal, MLP and baseline models.

(c) Bucket-wise comparison of superforecast with Ideal, MLP and baseline models averaged over six-months.

Figure 2: Results showing reduction in penalty in comparison to MLP and baseline models.

Case	A	P	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	200	100	120	0	0	0	0	0	20	20	20	20	0
2	200	300	200	20	20	20	20	0	0	0	0	0	0

Table 1: Distribution of predicted power in buckets for two different scenarios

(AP), India, with power generation capacity of 75.6 MW and 36 operational wind turbines. `SCADA` data is collected from each turbine in the wind farm at a frequency of 1 (average) reading every 10 minutes. For this wind farm, wind power generator employed four wind power forecasters represented by F-1, F-2, F-3 and F-4 (names anonymized). Each forecaster provides day-ahead-prediction for each 15 minute window. Figure 2a presents the percentage of times each of these forecaster provided the best forecast across all the 15 minute windows over the evaluation period, demonstrating the absence of single accurate forecaster and thus motivating the need to generate a superforecast.

$$\begin{gathered}
X_0 = \min(P + 0.1A, A), \text{if } P \in [0.9A, 1.1A] \\
X_1 = \max(\min(P - 1.1A, 0.1A), 0.0), \text{if } P \in [1.1A, 1.2A] \\
X_2 = \max(\min(P - 1.2A, 0.1A), 0.0), \text{if } P \in [1.2A, 1.3A] \\
X_3 = \max(\min(P - 1.3A, 0.1A), 0.0), \text{if } P \in [1.3A, 1.4A] \\
X_4 = \max(\min(P - 1.4A, 0.1A), 0.0), \text{if } P \in [1.4A, 1.5A] \\
X_5 = \max(P - 1.5A, 0.0), \text{if } P \geq 1.5A \\
X_6 = \max(\min(0.9A - P, 0.1A), 0.0), \text{if } P \in [0.8A, 0.9A] \\
X_7 = \max(\min(0.8A - P, 0.1A), 0.0), \text{if } P \in [0.7A, 0.8A] \\
X_8 = \max(\min(0.7A - P, 0.1A), 0.0), \text{if } P \in [0.6A, 0.7A] \\
X_9 = \max(\min(0.6A - P, 0.1A), 0.0), \text{if } P \in [0.5A, 0.6A] \\
X_{10} = \max(0.5A - P, 0.0), \text{if } P \leq 0.5A
\end{gathered} \tag{1}$$

Now we discuss the metric, *penalty*, used for the evaluation of our model. We divide predicted power into the 10 buckets, each of size 10% of the actual power generated. Table-1 shows two scenarios; Case-1: When predicted power (P) is less than actual power generated (A) and Case-2: When predicted power is greater than the actual power generated. Buckets start filling up in order $X_1 \rightarrow X_5$ and in order $X_6 \rightarrow X_{10}$ for Case-1 and Case-2 respectively. Values in each of these buckets are calculated using Equation 1. X_0 represents the exempted part of the predicted power, i.e. $[0.9A, 1.1A]$, corresponding to zero penalty ($p_0 = 0$). Total penalty for any prediction is calculated as weighted sum of power in each bucket (as shown in Equation 2) with weights (p_i) increasing as the predicted power deviates further away from the actual power. We simply assigned $p_i = i \ \forall_{i \in [0,5]}$. The baseline model, currently used by the generator for selecting the best forecaster, selected the forecaster with least penalty from the previous week and used its day-ahead-prediction for the current week. Our simple regression-based model (Equation 3) learns weights for each forecaster from the last 15 days of SCADA and forecaster data for day-ahead-prediction. We also generate ideal scenario in which we select least penalised forecaster for a day and use prediction from the same forecaster for the same day. This gives us an idea of minimum penalty that we can achieve using our model (when using the same forecaster for the whole day).

$$penalty = p_0^* X_0 + \sum_{i=1}^{5} p_i^* (X_i + X_{i+5}) \tag{2}$$

$$y = \sum_{i=1}^{4} \mathbf{F_i} \times w_i + \epsilon \tag{3}$$

Figure 2b presents monthly distribution of penalty for superforecast in comparison with least penalised forecaster for that month(MLP), baseline scenario and the ideal scenario. Our model reduced the average penalty by 7% and 13% when compared to most accurate forecaster for the month and baseline model, respectively. Figure 2c shows the average of monthly variation in buckets for superforecast, baseline and least penalised forecaster for the month.

4. CONCLUSIONS

Wind is one of the major renewable sources of energy and require accurate day-ahead-prediction for integration into the existing power systems. Wind power generators use multiple forecasters to generate day ahead predictions for their wind farms as their accuracy varies across the year, day and wind speed. In this work, we propose a basic model to generate superforecast using predictions from multiple forecasters to reduce penalty over wind power generators in comparison to their baseline model. We analysed six-month data from a wind farm located in the southern east part of India and results show a reduction of approx. 7% and 13% in terms of penalty when compared with least penalised forecaster for the month and the current heuristic for selecting the best forecaster respectively. We are working on developing improved models to generate superforecast with better accuracy than simple regression-based approach.

5. REFERENCES

[1] A. Haque, P. Mandal, H. Nehrir, A. Bhuiya, and R. Baker. A hybrid intelligent framework for wind power forecasting engine. In *Electrical Power and Energy Conference (EPEC), 2014 IEEE*, pages 184–189, Nov 2014.

[2] I. Sánchez. Adaptive combination of forecasts with application to wind energy. *International Journal of Forecasting*, 24(4):679–693, 2008.

[3] Y. Zhang, J. Wang, and X. Wang. Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*, 32(0):255 – 270, 2014.