Norwegian
Meteorological
Institute

No. 6/2018
ISSN 2387-4201

METreport

Meteorology

On the use of NWP forecasts in wind power forecasts for the next few hours

John Bjørnar Bremnes

Title On the use of NWP forecasts in wind power forecasts for the next few hours	Date 31 May 2018
Section Meteorology	**Report no.** No. 6/2018
Author(s) John Bjørnar Bremnes	**Classification** ● Free ○ Restricted
Client(s) Norwegian Research Council	**Client's reference** Project no. 256399 Nowcasting for wind energy production -- an integrated modelling approach[1]

Abstract

Wind power forecasts for the next few hours are ideally based on forecasts from numerical weather prediction (NWP) models and recent measurements of weather conditions and power production. The main objective of the study is to assess the potential benefit of using a deterministic NWP model (MET Norway's MEPS control member) with one hour forecast generation time compared to an NWP ensemble (MEPS 10 members) with 2.5 hour generation time. Based on various use of the predictive information sources hourly quantile wind power forecasts are made in three steps. First, tree-based gradient boosting (xgboost) is applied to predict expected wind power at each wind turbine. Second, predictions are aggregated to wind farm level. Third, constrained quantile regression splines are used to make quantile forecasts at wind farm level. Nine months of data for the Norwegian wind farms Bessakerfjellet and Hitra were organized to evaluate several forecast models. In terms of mean absolute errors the results were neutral, but in situations where moderate to large changes in wind speed were forecast, the scores were in favour of the deterministic NWP model with one hour generation time.

Keywords

wind power forecast, NWP models, statistical methods, nowcasting

Disiplinary signature

Responsible signature

[1] The report summarizes the work and outcome of Task H3.1 *Statistical post-processing* in work package H3.

Table of contents

Introduction

During the last two decades there has been a tremendous growth in installed wind power production capacity worldwide. With this, new challenges have also arisen. Due to the non-dispatchable nature of wind power, the demand for information about future power production has become stronger amongst energy producers and grid operators. High quality wind power forecasts for the next hours and days are therefore essential in power production planning and trading.

Wind power forecasts are based on two sources of predictive information: weather forecasts from numerical weather prediction (NWP) models and recent measurements of weather conditions and power production. Their respective degree of importance are highly dependent on the forecast horizons. The longer the horizon the more impact NWP forecasts have and vice versa for recent measurements data. The equilibrium is often roughly found to be between two to four hours, but that can vary considerably with location, whether upstream measurements are available, and not least the application of the wind power forecasts.

The main focus in this study is forecasts for the next few hours which are invaluable for optimal trading in for example Nord Pool's intraday market. Making wind power forecasts for these forecast horizons is a challenging task due to that the two sources of predictive information must be combined. The aim here is to investigate the role of the NWP forecasts for these horizons. Normally, the process of making an operational NWP forecast takes more than two hours to complete, and it is of interest to know how much the quality of wind power forecasts can be improved by making NWP forecasts available only one hour after initialization time. In this study, this is emulated by using MET Norway's operational NWP ensemble model system MEPS and assuming that forecasts are ready after one hour instead of about 2.5 hours as normal.

Data

Forecast and measurement data were collected and organized for the two wind farms Bessakerfjellet (57.5 MW, TrønderEnergi) and Hitra (55.2 MW, Statkraft) for the period 10 November 2016 to 31 August 2017.

The forecasts were extracted from MET Norway's operational weather forecast model MEPS with 2.5 km spatial horizontal resolution and 10 scenarios (ensemble members). The forecasts were all initiated at 00, 06, 12, and 18 UTC and available for operational use about 2.5 hours later. Only hourly forecasts of wind speed and direction at 10 meter height with lead times up to +40 hours ahead were used in this study. These were bi-linearly interpolated to the locations of the turbines.

Measurements of wind speed and direction at the nacelles and power production for each turbine were also made available. A very brief quality control was carried out by plotting measured wind speed against wind power production. Data far from the idealized power curve was somewhat subjectively removed before further use. About 4.8 % and 2.0 % of the measurement data at Bessakerfjellet and Hitra, respectively, were excluded in this process. The measurements at Hitra were available with a 10-minute time resolution, but was aggregated to hourly resolution after the quality control. For Bessakerfjellet the provided data already had hourly temporal resolution and no further preprocessing was necessary.

Methods

In order to make good wind power forecasts for the next hours information from MEPS and recent measurement data at the time the forecast is being created need to be combined using statistical methods. Here, only one statistical approach is applied. However, by varying the input we are able to test various use of the predictive information sources and to quantify their degree of importance.

All statistical models need to be fit or trained on data. In this study, the statistical models are fitted on the first day of each month using data from the previous three months as training data. For example, all predictions for April are based on models trained on data for March, February, and January. The time-adaptive training regime makes it possible to account for seasonal variations and changes in the NWP models.

Statistical approach

Wind power forecasts for the wind farms are generated separately for each lead time in three steps:

1. For each turbine, predictions of the expected power production are computed by the gradient tree boosting algorithm *xgboost* using square error loss function, see Chen and Guestrin (2016) for details[2]. The number of iterations/rounds are determined by a two-fold cross-validation (odd and even days) on the first training data set, while the remaining tuning parameter are set to default values. The same fitted model is applied to all ensemble members resulting in an ensemble of wind power forecasts for expected wind power production.
2. The wind power predictions are aggregated to wind farm level, and ensemble statistics such as the mean is computed.
3. Constrained quantile regression splines with the ensemble mean wind power prediction as input is used to make predictions of the 5, 25, 50, 75, and 95 % quantiles. The spline functions are based on linear B-splines with five equally spaced interior knots and constrained to be within the range of possible power production values for the wind farm. In addition, they are forced to be non-decreasing by re-ordering of the spline coefficients. The problem of crossing quantiles is also handled by re-ordering of the spline coefficients in ascending order with increasing quantile levels.

Forecast models

Various forecast models are defined by which data is used as input to the statistical forecast method. The three basic models are

- OBS
    - latest measured wind power production
    - latest measured wind speed

[2] The R-package *xgboost* version 0.6-4 for R version 3.4.4 is applied.

        - latest measured hourly change in wind speed
        - hour of the day
- MEPS
    - wind speed forecast at the end of the hour
    - hourly mean wind speed forecast
    - change in wind speed forecast during the 1-hour period
    - hourly mean wind speed forecast averaged over all turbines in the wind farm
    - wind direction at the end of the hour
    - hour of the day
- ALL
    - all of the above

The latter two are further refined by making assumptions on how long it takes to generate the MEPS forecasts and whether an ensemble or only the control member is available. The complete list of forecast models is given in Table 1.

Table 1. Statistical forecast models. The input variables are listed above.

Acronym	Description of input data
OBS	only based on observations
MEPS_3	MEPS with 3 hours generation time (as current operational forecast)
ALL_3	observations and MEPS with 3 hours generation time
MEPS_1	MEPS with 1 hour generation time
ALL_1	observations and MEPS with 1 hour generation time
MEPS_1C	MEPS control member with 1 hour MEPS generation time
ALL_1C	observations and MEPS control member with 1 hour generation time

The forecast model ALL_3 represents the best use of MET’s current operational weather forecasts, while ALL_1C is most similar to the proposed NWP nowcasting system that currently are being developed. Table 2 shows which MEPS forecasts are used for issuing wind power forecasts every hour. The upper table assumes three hours generation time, while the lower only one hour.

Table 2. Forecast reference hours. In the upper table three hour MEPS generation time is assumed, while in the lower only one hour.

Reference hour	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
MEPS reference hour	18	18	18	00	00	00	00	00	00	06	06	06	06	06	06	12	12	12	12	12	12	18	18	18
Minimum MEPS lead time in hours	6	7	8	3	4	5	6	7	8	3	4	5	6	7	8	3	4	5	6	7	8	3	4	5

Reference hour	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
MEPS reference hour	18	00	00	00	00	00	00	06	06	06	06	06	06	12	12	12	12	12	12	18	18	18	18	18
Minimum MEPS lead time in hours	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5

Results

The quality of the various forecast models is quantified by the mean absolute error (MAE) of the 50 % quantile forecast, which in theory is optimal with respect to absolute error. The MAE is presented in percentage of total power production capacity and its precise definition here is

$$MAE = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{c_t} \left| q_t - p_t \right| 100$$

where T is the number of forecasts, c_t is the power production capacity at time t, q_t is the 50 % quantile forecast at time t, and p_t the measured power production at time t.

Bessakerfjellet

The overall results for Bessakerfjellet are summarized in Table 3 where forecasts for all 24 reference hours have been included. The main findings were:

- The differences between ALL_3 and ALL_1C are practically insignificant. For perhaps the most important lead time for intraday trading, +2h, ALL_1C is only 0.3% better.
- The MAEs for models using NWP forecasts vary from about 4% to 11%
- Using recent measurement data has mostly impact on forecasts up to 3 hours ahead. Beyond +12 hours their impact is slightly negative which indicate that there is no additional information in them when NWP forecasts also are included
- The quality of wind power forecasts based only on NWP forecasts does not vary much with lead time (from about 7% to 10%). The relative large MAE for the first hours may indicate that there is scope for tuning NWP models better to these lead times. On the other hand, it could also imply that the weather conditions within a wind farm are too complex to be modeled accurately with NWP models on kilometer-scale horizontal resolutions.

Table 3. Mean absolute error (%) at Bessakerfjellet for selected lead times. All reference hours.

	+1h	+2h	+3h	+4h	+5h	+6h	+12h	+24h	+36h
OBS	4.61	7.12	9.01	10.54	11.87	13.01	17.55	21.28	23.04
MEPS_3	7.53	7.68	7.80	7.92	8.06	8.23	8.85	9.50	10.79
ALL_3	4.31	6.16	6.98	7.47	7.74	8.06	8.94	9.86	11.16
MEPS_1	7.39	7.45	7.54	7.71	7.82	7.92	8.66	9.43	10.63
ALL_1	4.29	6.04	6.84	7.33	7.60	7.83	8.72	9.77	10.91
MEPS_1C	7.71	7.83	7.93	7.97	8.17	8.32	8.84	10.04	11.27
ALL_1C	4.34	6.14	6.91	7.36	7.76	8.02	8.82	10.04	11.47

Since the MEPS forecasts were only updated four times a day, only the reference hours 01, 07, 13, and 19 UTC actually represented one hour forecast generation time. The results for these reference hours are given in Table 4. The scores are quite similar to those in Table 3, but it can be noticed that they now are in slightly more favour of ALL_1C compared to ALL_3; for example, for lead time +2h the MAE was about 2.9 % better.

Table 4. Mean absolute error (%) at Bessakerfjellet for selected lead times. Reference hours 01, 07, 13, and 19 UTC.

	+1h	+2h	+3h	+4h	+5h	+6h	+12h	+24h	+36h
OBS	4.53	7.43	8.81	10.61	11.54	12.53	17.02	20.93	22.88

MEPS_3	7.56	8.12	8.25	7.88	7.95	8.23	8.80	9.57	10.73
ALL_3	4.35	6.66	7.06	7.26	7.52	8.10	8.97	9.85	11.05
MEPS_1	7.34	7.79	7.46	7.35	7.48	7.59	8.30	9.43	10.06
ALL_1	4.24	6.43	6.80	7.04	7.33	7.67	8.33	9.66	10.39
MEPS_1C	7.51	8.21	7.95	7.65	7.89	8.04	8.50	10.36	10.69
ALL_1C	4.29	6.47	6.81	7.03	7.52	7.75	8.48	10.14	11.10

In the previous two tables scores are averaged over all forecasts. The averaging may conceal possible differences between the forecast models in extremal errors which are of special importance. In Table 5 selected quantiles of the errors (forecast minus measurement) for lead time +2 hours are given. The comparison of ALL_3 and ALL_1C shows no obvious systematic differences.

Table 5. Selected quantiles of forecast error distribution at Bessakerfjellet for lead time +2h.

	0.5%	1%	25%	50%	75%	99%	99.5%
OBS	-48	-39.9	-3.7	0.4	3.5	33.7	41.6
MEPS_3	-48.4	-40.1	-6.4	-0.1	3.2	27.8	33.4
ALL_3	-41.0	-34.1	-4.0	0.2	2.5	26.9	33.7
MEPS_1	-47.6	-39.2	-6.5	-0.2	3.1	26.0	32.8
ALL_1	-40.6	-31.8	-3.9	0.1	2.3	27.2	33.3
MEPS_1C	-54.3	-44.1	-6.0	-0.2	3.1	30.0	35.8
ALL_1C	-43.1	-34.2	-4.1	0.1	2.4	27.1	34.5

Making good wind power forecasts in persistent weather conditions poses no major challenges. However, to forecast changes in wind speed is demanding as both the timing and magnitude may have severe impact on decisions to be made. In order to assess the skill of the forecast models in these weather situations, the MAE scores were stratified according to hourly change in wind speed forecast of the control member (average for the wind farm), see Table 6. Only results for lead time +2h are provided. In cases where wind speed was forecast to decrease more than 1 m/s, ALL_1C was clearly better than ALL_3, while for wind speeds increasing by more than 1 m/s the scores were similar.

Table 6. Mean absolute error (%) at Bessakerfjellet for lead time +2h conditioned on hourly change in wind speed forecast (wind farm average for control member). Reference hours 01, 07, 13, and 19 UTC.

	< -1 m/s	(-1 ,-0.5] m/s	(-0.5 ,0] m/s	(0, 0.5] m/s	(0.5, 1] m/s	> 1 m/s
OBS	12.67	10.29	6.74	5.59	9.29	21.47
MEPS_3	15.56	8.79	7.86	6.76	9.80	15.98
ALL_3	12.24	8.85	6.32	4.96	8.09	17.95
MEPS_1	11.61	9.81	6.07	7.08	10.34	18.14
ALL_1	9.30	7.88	4.80	5.58	9.20	19.15
MEPS_1C	13.34	11.34	6.61	7.25	10.33	17.48
ALL_1C	8.75	8.68	5.03	5.52	8.95	17.19

Hitra

The results for Hitra are reported in Tables 7-10. Most of the results are similar to those for Bessakerfjellet. However, from Tables 7 and 8 it can be noticed that the forecast model ALL_3 was better than ALL_1C. On the other hand, when more than one m/s change in wind speed was forecast, ALL_1C was clearly better than ALL_3 (Table 10). It should be added that these important cases only amounts to about 4 % of the total number of forecasts. Their impact on the overall scores are therefore modest, as is seen.

Table 7. Mean absolute error (%) at Hitra for selected lead times. All reference hours

	+1h	+2h	+3h	+4h	+5h	+6h	+12h	+24h	+36h
OBS	4.70	7.77	9.93	11.72	13.30	14.39	19.32	23.66	23.95
MEPS_3	10.34	10.33	10.37	10.42	10.50	10.57	11.05	11.82	12.74
ALL_3	4.49	6.89	8.02	8.81	9.33	9.64	10.67	11.56	12.97
MEPS_1	10.14	10.23	10.31	10.34	10.38	10.44	10.85	11.65	12.59
ALL_1	4.46	6.78	7.94	8.69	9.14	9.50	10.50	11.50	12.81
MEPS_1C	10.49	10.59	10.77	10.92	10.95	11.02	11.22	12.24	13.52
ALL_1C	4.64	7.32	8.50	9.13	9.46	9.96	10.84	12.03	13.28

Table 8. Mean absolute error (%) at Hitra for selected lead times. Reference hours 01, 07, 13, and 19 UTC.

	+1h	+2h	+3h	+4h	+5h	+6h	+12h	+24h	+36h
OBS	4.61	7.94	10.17	12.13	13.07	13.90	18.79	23.29	23.40
MEPS_3	10.32	10.45	10.57	10.77	10.66	10.36	10.78	11.45	12.28
ALL_3	4.54	7.03	8.16	8.95	9.25	9.42	10.61	11.22	12.37
MEPS_1	9.99	10.27	10.41	10.52	10.04	10.15	10.31	11.06	12.17
ALL_1	4.44	6.82	7.89	8.62	8.67	9.11	9.99	11.00	12.12
MEPS_1C	10.33	10.57	10.53	11.19	10.43	10.81	10.65	11.70	13.02
ALL_1C	4.69	7.36	8.39	9.13	8.80	9.55	10.21	11.48	12.65

Table 9. Selected quantiles of forecast error distribution at Hitra for lead time +2h.

	0.5%	1%	25%	50%	75%	99%	99.5%
OBS	-53.5	-41	-4.6	0.6	3.8	38.7	45.5
MEPS_3	-54.5	-45.7	-6.6	0.1	3.9	71.2	89.2
ALL_3	-42.1	-32.5	-4.3	0.4	3.6	33.0	40.1
MEPS_1	-50.6	-44.5	-6.6	0.1	4.0	71.3	90.0
ALL_1	-40.6	-32.0	-3.9	0.4	3.6	33.2	41.2
MEPS_1C	-53.8	-47.6	-5.5	0.5	5.1	75.9	88.6
ALL_1C	-37.7	-31.2	-4.1	0.6	4.0	42.0	50.2

Table 10. Mean absolute error (%) at Hitra for lead time +2h conditioned on hourly change in wind speed forecast (wind farm average of control member). Reference hours 01, 07, 13, and 19 UTC.

	< -1 m/s	(-1 ,-0.5] m/s	(-0.5 ,0] m/s	(0, 0.5] m/s	(0.5, 1] m/s	> 1 m/s
OBS	19.82	8.56	6.48	7.04	9.46	28.18
MEPS_3	10.92	9.17	10.43	10.16	10.28	20.16

ALL_3	13.88	7.57	6.32	6.21	8.55	18.53
MEPS_1	12.22	11.20	10.31	9.04	12.09	14.93
ALL_1	10.98	6.56	6.26	6.00	7.95	19.81
MEPS_1C	13.60	11.91	9.93	9.18	15.83	13.46
ALL_1C	10.10	7.60	7.21	6.26	10.03	12.02

Discussion and summary

While the overall results were quite neutral with respect to reducing the generation time of NWP forecasts, the scores in situations with moderate to strong changes in forecast wind speeds were more promising and requires further attention. In particular, it would be interesting to test the statistical wind power forecasting models with input from a real NWP nowcasting system with hourly updates that is tuned towards optimal forecasts for lead times up to a few hours ahead. It should be added that an operational nowcasting system will not directly be able to assimilate as much observational data as the MEPS, but since MEPS forecasts are used as first guess most, if not all, observational data has impact on the analyses. Further, longer generation times, as in the MEPS, is not necessarily a disadvantage. Forecasts for lead times up to the generation time are possible to evaluate, that is, the quality of the scenarios are known for the first lead times which may be used to weigh them or even discard some ensemble members.

In this study, the wind power forecasts were only evaluated in terms of mean absolute error. However, the main objective of the wind power forecasts was energy trading and the actual usefulness of the forecasts would ideally be measured in terms of euros or Norwegian kroner. The two metrics may not necessarily lead to the same conclusions. For energy producers rapid changes in energy production are a major concern, and If these are not well forecasted high balancing costs will occur.

Regarding the use of NWP forecast data, there are several issues that can be refined. First, only wind forecasts at 10 meter heights were available, while forecasts at the height of the nacelles would probably be beneficial and improve the wind power forecasts. Second, it would be interesting to explore the potential of using NWP forecasts on a grid around the wind farms; here, only forecasts at the respective turbine and wind farm averages were used.

Concerning the statistical modeling, recent measurement data and NWP forecasts were combined in only one model. However, quite often, as in this case, longer time series of measurement data were available and these are not easy to take advantage in one statistical model. To fully utilize all data it would likely be better to make two separate statistical models, one based on measurement data and another based on NWP forecasts as input, and combine the output from these in a third statistical model. For operational forecasting this may also be a more stable and robust forecast system as it is less vulnerable to missing measurement data.

Finally, it should be recognised that the quality of wind farm data has an impact on the quality of wind power forecasts. For the statistical modeling it is in particular important to know the maximum possible power production at any point in time or at least some information about the state of the turbines. Further, it may also be beneficial to have wind farm data available at a higher temporal resolution than hourly, especially for forecasts for the first few hours ahead. Cleaning the data prior to training the statistical models is also essential.

Acknowledgements

The wind farm data for Bessakerfjellet and Hitra were kindly provided by TrønderEnergi and Statkraft, respectively. The work is part of the project "*Nowcasting for wind energy production -- an integrated modelling approach*" funded by the Norwegian Research Council (project number 256399) and industry partners Kjeller Vindteknikk, Statkraft, TrønderEnergi, Vestas Wind Systems, and WindSim.

References

Chen T. and C. Guestrin (2016). XGBoost: A Scalable Tree Boosting System. arXiv:1603.02754v3.