ISSN 2004-2965 Energy Proceedings, Vol. 21, 2021

International Conference on Applied Energy 2021
Nov. 29 - Dec. 5, 2021, Thailand/Virtual
Paper ID: 055

Increasing the skill of short-term wind speed ensemble forecasts combining forecasts and observations via EMOS

Gabriele Casciaro[1*], Andrea Lira-Loarca[1], Andrea Mazzino[1,2]

1 Department of Civil, Chemical and Environmental Engineering. University of Genoa, Via Montallegro 1, Genoa, 16145, Genoa, Italy

2 Istituto Nazionale di Fisica Nucleare, Sezione di Genova, Via Dodecaneso 33, Genoa, 16146, Genoa, Italy

* Corresponding Author

ABSTRACT

All numerical weather prediction models used for the wind industry need to produce their forecasts starting from the main synoptic hours 00, 06, 12, and 18 UTC, once analysis become available. The six-hour latency time between two consecutive model runs calls for strategies to fill the gap by providing new accurate predictions having, at least, hourly frequency. This is done to accommodate the request of frequent, accurate and fresh information from traders and system regulators to continuously adapt their work strategies. Here, we propose a strategy where quasi-real time observed wind speed and weather model predictions are combined by means of a novel Ensemble Model Output Statistics (EMOS) strategy. The success of our strategy is measured by comparisons against observed wind speed from SYNOP stations over Italy in the years 2018 and 2019.

Keywords: wind forecasting, probabilistic forecasting, dynamic forecast calibration, ensemble model output statistics, wind forecast based on real-time conditions, Numerical Weather Prediction models

1. INTRODUCTION

Global cumulative installations of onshore and offshore wind are expected to exceed 1 TW before 2025 (Global Wind Energy Council, 2021). This means that the contribution of wind power in power systems is becoming increasingly important. The downside is that detailed schedule plans and reserve capacity must be properly set by power system regulators (Impram et al., 2020) facing the intrinsic problem of the highly intermittent nature of wind, making this very hard to predict. The accuracy of wind forecasts thus becomes an issue of paramount importance for the wind industry.

In a recent work by Casciaro et al., 2021, a novel accurate Ensemble Model Output Statistics (EMOS) strategy for calibrating wind speed/power forecasts from an Ensemble Prediction System (EPS) has been proposed and its superiority when compared against more parsimonious strategies in the 0-48 h look-ahead forecast horizon clearly emerged. However, because all global weather models start their run from analysis corresponding to the main synoptic hours 00, 06, 12, and 18 UTC, weather predictions (of any forecast horizons) necessarily remain frozen for six hours. This limitation is in sharp contrast with the needs of power system regulators, as well as of traders for marketing wind energy, who need to adapt their strategies hour after hour. It is thus highly important to propose accurate strategies which give fresh information on the wind speed in a given location continuously evolving between two consecutive main synoptic hours. Proposing a strategy with such characteristics is the main aim of the present paper. In plain words, we propose a novel EMOS strategy where the mean entering in the EMOS predictive probability density function now also depends on observed wind speed data. For our strategy to be used operatively, observed wind speed on the site of interest must be available in quasi-real time, other than as a record of past observations for calibration purposes.

2. WIND DATA

2.1 Observed data: SYNOP stations

From 2018 to 2019, SYNOP meteorological measurements were collected at 43 locations around

Selection and peer-review under responsibility of the scientific committee of the 13th Int. Conf. on Applied Energy (ICAE2021).

Italy. According to ICAO regulations, the SYNOP anemometers record the wind speed as an average over 10 minutes (ICAO, 2007). Selected stations have an hourly time interval of the observations.

2.2 Forecast data: the ECMWF Ensemble Prediction System (EPS)

The ECMWF Ensemble Prediction System (EPS) has 51 members: 50 forecasts which add a small perturbation to the best-known initial condition and one control forecast with no perturbations (Buizza, 1995; Leutbecher and Palmer, 2008). The EPS used in this study has a resolution of about 18 km (Persson, 2001) and a spectral triangular truncation, a cubic-octahedral grid Tco639, and 91 layers with a top of atmosphere pressure of 0.01 hPa (Buizza, 2018).

3. METHODS

Gneiting et al. (2005) presented the Ensemble Model Output Statistic (EMOS), an easy-to-implement statistical post-processing tool that permits the calibration of an ensemble forecast as the EPS. Let us analyze the method, from its standard form to more sophisticated settings.

3.1 The standard EMOS

Let us denote by $X_1, \dots, X_k$ the *K* ensemble member forecasts of a univariate continuous, positive-defined variable *Y*, here the wind speed at a given location and look-ahead time. The EMOS method uses a parametric distribution of the following general form:

$$Y|X_1, \dots, X_k \sim f(Y|X_1, \dots, X_k)$$

where the left-hand-side means that the distribution is conditional on the ensemble members.

For this study the gamma distribution is used, and its Probability Density Function (PDF) is denoted as

$$\mathcal{G}(\mu, \sigma^2)$$

where the mean, μ, and variance, σ^2, are defined as,

$$\mu = a + b_1 X_1 + \dots + b_K X_K$$
$$\sigma^2 = c + dS^2$$

with, $a, b_1, \cdots, b_K, c, d$ representing the non-negative EMOS coefficients, and S^2 the ensemble spread defined as the EPS members variance.

Gneiting et al. (2005) suggested an approach based on the minimization of the Continuous Ranked Probability Score (CRPS) (Hersbach, 2000) to find the EMOS coefficients. The latter is defined as follows:

$$crps(F, Y) = \int_{-\infty}^{\infty} [F(t) - H(t - Y)]^2 dt$$

where F is the cumulated probability of $\mathcal{G}$, Y is the observation, and H is the Heaviside function, which returns 0 if $t < Y$, and 1 otherwise. So, considering the forecast vector $\boldsymbol{X} = (X_1, \cdots, X_K)$, in a training set of pairs of forecasts and observations, the quantity to be minimized is:

$$CRPS = \frac{1}{N} \sum_{i=1}^{N} crps(X_i, Y_i)$$

where i represents the i-th pair observation-forecast and N is the total number of pairs in the training set.

3.2 The EMOS$_{+4r}$ strategy

The standard EMOS approach (Gneiting et al., 2005, Thorarinsdottir and Gneiting, 2010) is characterized by three main aspects:

1. the length of the training period is usually of 40 days before the day of the forecast;
2. the considered forecasts are those corresponding to the grid point closest to the observations;
3. the predictive distribution is conditional on the sole ensemble observables that need to be forecasted.

Casciaro et al., 2021 proposed an evolution of the EMOS strategy divided in three steps by relaxing these points using a larger training set.

The first step is to add conditionings to the parametric distributions in addition to those on variables representing the weather quantity of interest.

So, the EMOS method in its standard form is modified as follows:

$$Y \mid X_1, \dots, X_K; Z_1, \dots, Z_M \sim \mathcal{G}(\, Y \mid X_1, \dots, X_K; Z_1, \dots, Z_M \,)$$

where $Z_1, \cdots, Z_M$ are M variables used for conditioning.

The second step is to consider the calibrated forecasts on the 4 grid points around the considered station by merging them with a conditional EMOS as in the first step, obtaining the mean value as:

$$\mu_i = a(Z) + b_{ij}(Z) X_i^j, \qquad i = 1, \cdots, K \qquad j = 1, \cdots, 4$$

where j, spans over the 4 model grid points around the station.

The final step is to make an EMOS using as training set a rolling window of 40 days as proposed by Gneiting et al., 2005.

Such combined strategy is called EMOS$_{+4r}$ by Casciaro et al. (2021).

3.3 The use of the observed data

Our idea here for improving short-time forecasts is to combine model forecasts with quasi-real time observations.

Two quantities are used for this purpose. The first is the observed wind speed at a given time (e.g., 06 UTC), while the second is the error of the prediction with respect to the observation at the same reference time. To take advantage of this strategy without losing the benefits of EMOS$_{+4r}$, we simply combined both of them in terms of a new EMOS we will detail below.

Let us consider P_h the observation and E_h the error of the forecast with respect to the observation at the reference hour h (e.g., 06 UTC). The standard EMOS is modified as:

$$Y \mid X_1, \dots, X_K, P_h, E_h; Z_1, \dots, Z_M \\ \sim \mathcal{G}(Y \mid X_1, \dots, X_K, P_h, E_h; Z_1, \dots, Z_M)$$

with the mean, μ, of the predictive gamma Probability Density Function (PDF) given by:

$$\mu = a + b_1 X_1 + \dots + b_K X_K + b_p P_h + b_e E_h$$

where b_p and b_e are the non-negative EMOS coefficients associated to the P_h and E_h variables. The variance, σ^2, is assumed as in the standard EMOS.

All coefficients will be determined by CRPS minimization. We dub this approach EMOS$_{+4ro}$ where "o" stays for observation.

4. STATISTICAL INDICES

The Skill Score (SS) metric (Wilks, 2011) will be used to evaluate the goodness of this new approach. The skill score compares the calibrated forecast to a reference forecast to quantify how better it is. Lower bounds differ depending on the score used to compute the skill (the MAE normalized with the mean of the observations, i.e. NMAE, and the correlation coefficient, i.e. Pearson coefficient; for more details see Wilks, 2011) as well as the reference forecast used. Upper boundaries, on the other hand, are always 1 and means a perfect performance. In plain words, the skill score is defined as

$$SS = \frac{A - A_{ref}}{A_{opt} - A_{ref}}$$

with A being the error index value of the calibrated forecast, A_{ref} is the error index value of the reference forecast, and A_{opt} is the optimal index value.

5. RESULTS

Let us now analyze the improvement of our forecast using observed data starting from the persistence as a reference to arrive at the best calibration strategy presented by Casciaro et al. (2021) as a reference.

Let us assume to have an observation available at h=06 UTC and need to predict the hours from 07 UTC.

Fig. 1 shows the NMAE and the correlation coefficient skill scores of the forecast calibrated with the EMOS$_{+4ro}$ using as reference the persistence. Results are shown as the mean of 43 SYNOP stations over Italy. Except for the first hour in which the NMAE of the persistence turns out

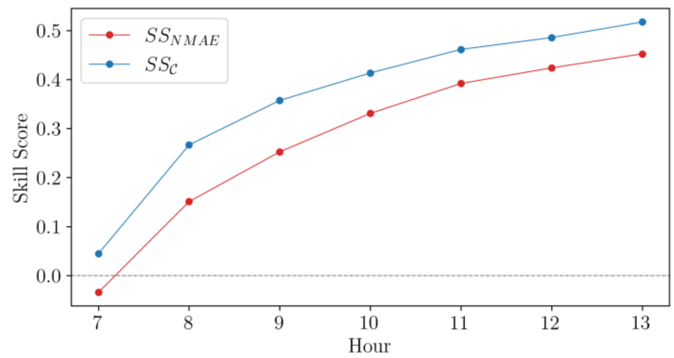

Fig. 1. Skill score of NMAE and correlation coefficient of the forecast calibrated with the EMOS$_{+4ro}$, using the persistence as reference forecast.

to be slightly better, the use of the EMOS calibration turns out to be significantly better for both the NMAE and the correlation coefficient.

The NMAE and the correlation coefficient skill score of the EMOS-calibrated forecast conditioning on the sole day hours and using the observed data (a strategy here denoted by EMOS$_0$) using as a reference the same

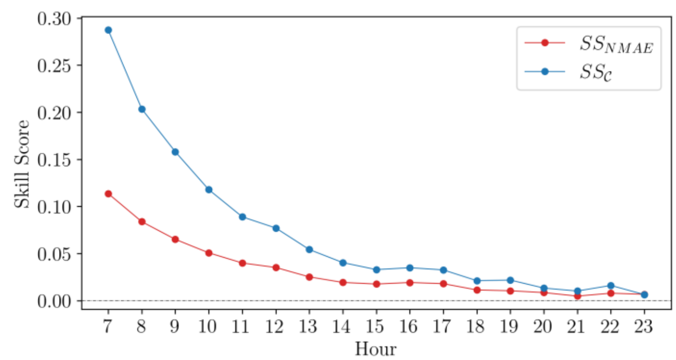

Fig. 2. Skill score of NMAE and correlation coefficient of the forecast calibrated with the EMOS$_0$ conditionated on day hours and using the observation data, using the EMOS$_0$ conditionated on day hours as reference forecast.

EMOS$_0$, but with no use of observed data, is shown in Fig. 2. From the figure, using the observed data leads to an overall improvement of the forecast. As expected, the improvement is greater closest to the time of the observation.

Ascertained the fact that compared to a standard EMOS the observed data leads to an improvement of the prediction, we now want to assess if, and how much, this improvement appears when observed data are used in combination with the best calibration strategy EMOS$_{+4r}$. Dubbing the resulting strategy as EMOS$_{+4ro}$, Fig. 3 shows the skill score of NMAE and correlation coefficient of the prediction calibrated with the EMOS$_{+4ro}$ compared with the prediction calibrated with the EMOS$_{+4r}$. The result is similar to that reported in Fig. 2 with a relevant improvement in the first hours of the forecast. Unlike the previous case, however, the decay appears to be faster,

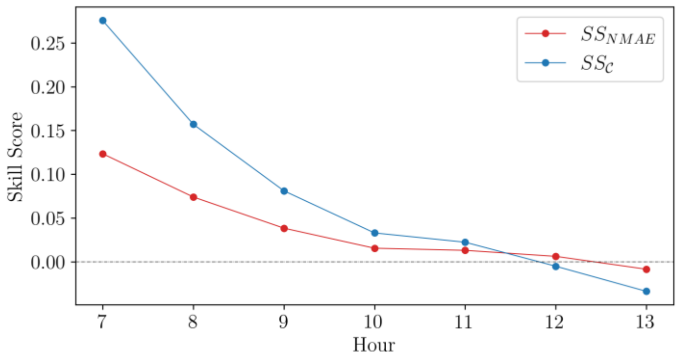

Fig. 3. Skill score of NMAE and correlation coefficient of the $EMOS_{+4ro}$ calibration using $EMOS_{+4r}$ calibration as reference.

leading to an improvement only in the first 6 hours after the observation.

Finally, to assess how this improvement changes by varying the observation hour, h. Fig. 4 shows the mean

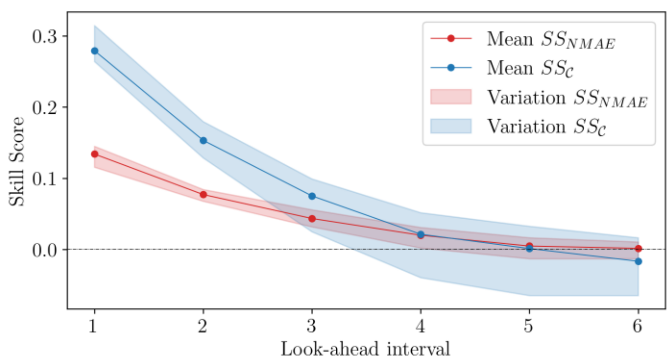

Fig. 4. Skill score of NMAE and correlation coefficient of the forecast calibrated with the $EMOS_{+4r}$ using the observation data, with the one calibrated with $EMOS_{+4r}$ as reference forecast. Continuous lines: averages of the skill scores obtained for observed data available at 00, 06, 12, and 18 UTC; shaded areas respresent the variation of the skill score.

and variation of the NMAE and correlation coefficient skill score of four cases in which h = 00, 06, 12, and 18 UTC. From the figure it is confirmed that the use of observed data leads to a considerable improvement in the first few hours and then decay to zero after about 5/6 hours, regardless of the selected hour of observation, h.

6. CONCLUSIONS

Quasi-real time wind speed observations and model forecasts have been combined in a new EMOS strategy to increase the forecast skill in the 6-hour horizon measured from the hour at which the observation is available. The added value of our strategy clearly emerged from our study. The improved forecast for the wind speed can be trivially translated in added value for the wind power forecast with many useful applications for the whole wind industry.

REFERENCE

Buizza, R., 1995. Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. Quarterly Journal of the Royal Meteorological Society 121, 1705–1738.

Buizza, R., 2018. Ensemble forecasting and the need for calibration, in: Statistical Postprocessing of Ensemble Forecasts. Elsevier, pp. 15–48.

Casciaro, G., Ferrari, F., Mazzino, A., 2021. Comparing novel strategies of ensemble model output statistics (EMOS) for calibrating wind speed/power forecasts. arXiv:2108.12174.

Global Wind Energy Council, 2021. GWEC, Global Wind Report 2021. URL: https://gwec.net/wp-content/uploads/2021/03/GWEC-Global-Wind-Report-2021.pdf.

Gneiting, T., Raftery, A.E., Westveld III, A.H., Goldman, T., 2005. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. Monthly Weather Review 133, 1098–1118.

Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting 15, 559–570.

ICAO, 2007. Meteorological Service for International Air Navigation: Annex 3 to the Convention on International Civil Aviation, 16th edn. International Civil Aviation Organization: Montreal, Canada.

Impram, S., Nese, S.V., Oral, B., 2020. Challenges of renewable energy penetration on power system flexibility: A survey. Energy Strategy Reviews 31, 100539.

Leutbecher, M., Palmer, T.N., 2008. Ensemble forecasting. Journal of Computational Physics 227, 3515–3539.

Persson, A., 2001. User guide to ECMWF forecast products. ECMWF Meteorological Bulletin M3 2.

Thorarinsdottir, T.L., Gneiting, T., 2010. Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. Journal of the Royal Statistical Society: Series A (Statistics in Society) 173, 371–388.

Wilks, D.S., 2011. Statistical methods in the atmospheric sciences. Volume 100. Academic Press.