INCREASED FORECASTING PERFORMANCE AND ROBUSTNESS IN FORECAST COMBINATIONS USING CORRECTED INDIVIDUAL FORECASTS

Mathias Käso, University of Technology Cottbus-Senftenberg, 0049 355 69 4040, mathias.kaeso@b-tu.de
Oliver Grothe, Karlsruhe Institute of Technology, grothe@kit.edu

Overview

In numerous countries, one can observe an increasing generation of intermittent renewable energy sources (RES) such as wind and photovoltaics. As a result, electricity systems face growing challenges with respect to system operations and balancing of supply and demand. It can be expected that this development will continue and most likely accelerate.

A demand covering electricity generation and a reliable operation of the electrical system is considerably more difficult, if the system is influenced by the uncertain energy production of RES. Therefore, balancing group operators as well as transmission grid operators need accurate forecasts for the electricity generation of these fluctuating sources. This is important to ensure low system balancing costs and guarantee a high level of security of supply.

This paper will focus on forecasting the variable wind energy production one day ahead in a quarter-hourly resolution which is characterized by strong varying statistical properties and a high volatility over time. Therefore, it is often challenging to predict the generated amount of wind energy with individual forecast models. For a better use of the strengths and to compensate the weaknesses of different individual model, combinations of individual forecasting models are frequently applied to forecasting problems. This results in an increased prediction quality.

In this context we analyse the effect of improving the individual forecasts before they are combined to a new meta forecast. We start from an empirical investigation using the combined error minimisation model (CEMM) as well as real wind energy forecast data from Germany. We find a further improvement of the prediction quality compared to the simple combination of the uncorrected forecasts. Motivated by this observation, we extend the investigation by a simulation experiment and some theoretical considerations. This approach confirms the observed improvement from the empirical study. In addition, the occurrence of extreme events in the mean squared error distribution over all CEMM-forecasts is reduced. Moreover we find that the CEMM shows more resistance against the mostly negative effects resulting from highly correlated individual forecasts.

Methods

We use the CEMM which is a two-step forecasting model. In the first step, we correct individual forecast by a (dynamic) non-linear error minimisation model based on polynomial approach. This leads to improved individual forecasts in the sense of the mean squared error. In the second step these forecasts are included in a linear combination with in general time varying combination coefficients.

We apply the CEMM to five different wind forecasts in a dynamic out-of-sample procedure using different correction and combination approaches. For a deeper analysis of the CEMM, we use a regression based CEMM in a Monte Carlo simulation with synthetic data which can represent typical properties of wind energy forecasts. To this end, we generate random samples of correlated forecast errors. The frequency of the errors is a function of the predicted energy feed-in and follows the typical Weibull-distribution. Each error is normal distributed around the predicted wind energy value with mean and standard deviation as a function of this energy. The degree of correlation between the errors, the mean and variance of the individual errors as well as of the entire resulting error distribution can be adjusted separately. Based on this synthetic data set, we generate out-of-sample forecasts according to the two steps of the CEMM leading to further new insights in its functionality.

Results

Our investigation shows three main results:

1. Empirically, we find an increased forecast performance of the CEMM compared to the individual forecasts and to the ordinary combination of uncorrected forecasts. These findings are valid for different correction and combination approaches. We can confirm the empirical observed improvement of the forecast quality in a Monte Carlo experiment with synthetic data and a regression based CEMM. In the theoretical analysis of these results, we find

evidence that the improvement effect by using corrected individual forecasts in a regression based combination should be a general property and therefore does not depend on the special correction method.

2. In the Monte Carlo simulation, we show furthermore that the out-of-sample forecasts of CEMM benefits from an increasing order of the correction polynomial. This is the case, if the sample size is not too small (≥ 100) which is practically no difficulty in energy forecasts. In this context, we find a decreasing mean value of the distribution of the mean squared errors over all CEMM-forecasts with increasing order of the correction polynomial. Simultaneously, we observe a reduction of variance of the mean squared errors of the CEMM-forecasts. Therefore, in addition to a better forecast on average, we have in the CEMM a lower probability of extreme mean squared errors occurring.

3. The given individual forecasts in a forecast combination can be highly correlated, because they are frequently based on similar information and model approaches. Usually the mean squared error of the combined forecast is then depending on the correlation between the individual forecasts and shows an increasing behaviour for a wide range of positive correlation coefficients. While this is a general problem in forecast combinations, this effect is significant smaller in the CEMM compared to the ordinary combination of uncorrected forecasts. This result is also consistent with our theoretical considerations. Therefore, the CEMM is more robust against negative influences which can be arising from highly positive correlated individual forecasts.

Conclusions

We present forecasting models combining RES-infeed forecasts from different commercial forecasting institutions. We show empirically that forecasting models combining individual wind energy forecasts are better than each individual forecast. A further increase of the wind energy prediction accuracy can be obtained by using corrected individual forecasts in different linear combination methods. We emphasise these empirical findings by a Monte Carlo experiment. In addition, the simulation and the analytical investigation shows that the CEMM is more robust in dealing with highly correlated forecast data and can reduce the probability of extreme mean squared errors.

Our results can help to integrate RES into energy systems worldwide. In addition, because of the general structure of the methodology, the results can potentially improve the forecasting accuracy of further RES-technologies such as photovoltaics or solar heat.