Solar Energy 166 (2018) 529–541

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Operational photovoltaics power forecasting using seasonal time series ensemble

Dazhi Yang[a,*], Zibo Dong[b]

[a] Singapore Institute of Manufacturing Technology, Agency for Science, Technology and Research (A*STAR), Singapore
[b] Solar Energy Research Institute of Singapore, National University of Singapore, Singapore

ARTICLE INFO

Keywords:
Ensemble forecast
Forecast combination
Time series
NWP

ABSTRACT

The hypothesis, namely, ensemble forecasts improve forecast accuracy, is herein investigated. More specifically, this paper considers an application of day-ahead operational PV power output forecasting using time series ensembles. Since numerical weather prediction (NWP) is strongly favored for day-ahead solar forecasting, the motivation of using seasonal time series forecasting models is made clear at the beginning of the paper. A total of 142 models from six families are considered: the SARIMA family of models (36 models), ETS family of models (30 models), MLP (1 model), STL decomposition (2 models), TBATS family of models (72 models) and the theta model (1 model), see main text for descriptions. These models first undergo a within-group competition judged using the Akaike information criterion (AIC). The forecasts made by the six winning models, one from each family, are then combined using eight different methods: (1) simple averaging, (2) variance-based combination, (3) combination through ordinary least squares regression, (4) least absolute deviation regression, (5) constrained least squares regression, (6) complete subset regressions, (7) AIC-weighted subset regressions, and (8) lasso regression. Methods (3) to (8) cover most aspects of regression-based forecast combinations.

The heavy empirical evidence from the case study suggests that the ensemble forecasting using seasonal time series models only provides marginal improvements over the best component model, in the day-ahead forecasting exercise. To that end, the pitfalls of using only time series ensembles are subsequently identified and discussed. A simple remedy, namely, adding an (uncorrected) NWP model to the ensemble, is proposed. The refined ensemble results show significant improvements over the best component model, due to the expanded information set provided by the NWP model. When this additional information set is exploited, e.g., adding a model-output-statistics-corrected NWP model, further improvements on ensemble-forecast accuracy are observed. Although a theoretical discussion on the benefits of ensemble forecasting is lacking in this paper, based on the empirical study, it appears that using forecast combination is less risky than using the best component model.

1. Introduction

Combining forecasts is perhaps one of the few things that forecasters can agree on. Since at least 1969, the seminal paper by Bates and Granger (1969), forecast combination has been well utilized in many scientific domains including solar engineering. In a recent review paper, forecast combination was identified to be one of the most important directions of future solar forecasting research (Yang et al., 2018). The naming convention for forecast combination is dispersed in the solar forecasting literature. Several terms have been used to describe it, e.g., reforecast (Chu et al., 2015), ensemble forecasting (Sperati et al., 2016), multi-modeling (Sanfilippo et al., 2016), reconciliation (Yang et al., 2017b) and ensemble learning (Jiang et al., 2017); they are collectively referred to as *ensemble methods* by Ren et al. (2015). Despite the subtle differences among these methods, the two terms, forecast combination and ensemble, are used interchangeably hereafter.

In the review by Ren et al. (2015), ensemble methods for solar and wind forecasting are classified into two categories, each containing two sub-categories, as shown in Fig. 1. *Competitive ensemble* uses multiple models and parameters to perform forecasts. The final forecasts are computed by (weighted) averaging of individual forecasts. On the other hand, *cooperative ensemble* divides a forecasting task into several steps, where different methods are used in different steps to complete the final forecasts. For a more detailed discussion on the classification and example works in each class, the reader is referred to Ren et al. (2015).

* Corresponding author.
E-mail addresses: yangdazhi.nus@gmail.com, yangdz@simtech.a-star.edu.sg (D. Yang).

https://doi.org/10.1016/j.solener.2018.02.011
Received 2 November 2017; Received in revised form 11 January 2018; Accepted 2 February 2018
0038-092X/

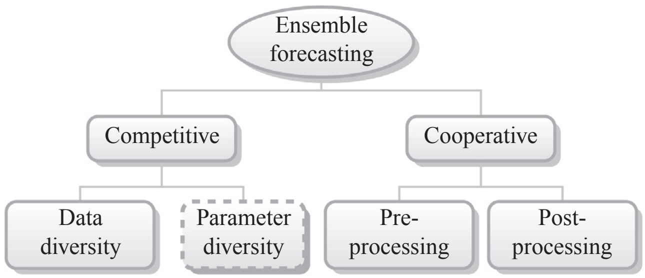

Fig. 1. Classification of ensemble methods for wind and solar power forecasting according to Ren et al. (2015). The box with dashed frame indicates the class for the method herein presented.

The present paper is concerned with the competitive ensemble methods. More specifically, an application of using *parameter diversity* to improve day-ahead operational forecasting is investigated. In contrast to *data diversity*, where N different datasets, $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N$, are used to generate the final forecast,[1] parameter diversity aims at exploiting various parameter settings using the same dataset $\boldsymbol{x}$. The forecasts can then be generated, according to Ren et al. (2015), using:

$$\hat{y}_{t+h} = \frac{1}{N}\sum_{i=1}^{N} f_i(\boldsymbol{x},\theta_i), \tag{1}$$

where θ_i is the parameter for model f_i, $i = 1,\ldots,N$. It should be noted that the combination in Eq. (1) uses equal weighting. This is most likely due to the fact that most solar forecasting studies using parameter diversity consider numerical weather prediction (NWP) models,[2] for which the initial conditions need to be perturbed. It is thus reasonable not to have any preference on a particular setting. However, if time series or machine learning ensemble is considered, a more appropriate combination equation would be

$$\hat{y}_{t+h} = \frac{1}{N}\sum_{i=1}^{N} w_i f_i(\boldsymbol{x},\theta_i), \tag{2}$$

or simply

$$\hat{y}_{t+h} = \sum_{i=1}^{N} w_i f_i(\boldsymbol{x},\theta_i), \tag{3}$$

if the weights, w_i, are scaled. Most of the frequently used ways to generate w_i—simple averaging, variance-based combination, ordinary least squares, constrained least squares, etc.—are explored in this paper. The majority of the forecast combination methods used in this paper are based on regressions.

Having discussed the forecast combination approach, the application, namely, day-ahead operational forecasting using *seasonal time series* ensemble, is motivated next. It is now well known that NWP models are preferred in day-ahead solar forecasting. Despite that seasonal time series models have also been explored in day-ahead forecasting exercises (e.g., Aryaputera et al., 2015), the lack of physical modeling of the atmosphere often limits the accuracy of time series models. Therefore, if the power output forecasts of a single photovoltaic (PV) power plant are of interest, very few arguments can be made on not using NWP. However, if the power output forecasts of many PV power plants are of interest, NWP forecasts may not be sufficient. To reason this statement, we digress and discuss the so-called *geographical hierarchy* of solar forecasting.

PV power output can be modeled using a geographical hierarchy, where strings aggregate to inverters, inverters aggregate to plants, plants aggregate at various distribution and transmission nodes, and finally, all PV power aggregates to the overall regional PV generation. For illustration, Fig. 2 shows a simple geographical hierarchy containing a total of m transmission zones. A total of $n = n_1 + n_2 + \cdots n_m$ PV plants are tied to those transmission zones. Whereas NWP models can be used to generate forecasts at the PV-plant level (or $\mathscr{L}_{\text{PV}}$), the forecasts at the transmission-zone level and the regional level ($\mathscr{L}_{\text{Trans}}$ and $\mathscr{L}_{\text{Reg}}$, respectively) can be generated by adding up the $\mathscr{L}_{\text{PV}}$ forecasts according to the structure of the hierarchy. Such aggregation is known as the *bottom-up reconciliation*. In a recent publication by Yang et al. (2017a), it was demonstrated that the bottom-up reconciliation in fact performs the worst as compared to other reconciliation methods, such as the *MinT reconciliation*,[3] which not only use $\mathscr{L}_{\text{PV}}$ forecasts, but also the forecasts made using data at $\mathscr{L}_{\text{Trans}}$ and $\mathscr{L}_{\text{Reg}}$. To produce $\mathscr{L}_{\text{Trans}}$ and $\mathscr{L}_{\text{Reg}}$ forecasts independently without utilizing the PV-plant level information, a time series method, namely, exponential smoothing (ETS), was used by Yang et al. (2017a). Since ETS is a seasonal time series method, the motivation of considering seasonal time series ensemble is made clear, i.e., an ensemble is hypothesized to outperform ETS.

Following the motivation above, the same data—solar power data for integration studies (SPDIS)—as used by Yang et al. (2017a,b) is considered in this paper. Since it is only of interest to apply the time series ensemble on the geographically-aggregated time series, the time series at $\mathscr{L}_{\text{Reg}}$ and $\mathscr{L}_{\text{Trans}}$ are considered. The $\mathscr{L}_{\text{Reg}}$ time series is the sum of the normalized PV power output from 318 PV systems in California (Yang et al., 2017a); it contains one year (2006) of hourly values. The data is visualized in Fig. 3 using a 3-dimensional plot. Similarly, a total of five $\mathscr{L}_{\text{Trans}}$ time series were used in (Yang et al., 2017a), which are the results of aggregating various PV systems to hypothetical transmission nodes. Therefore, a total of six time series ($1 \times \mathscr{L}_{\text{Reg}}$ and $5 \times \mathscr{L}_{\text{Trans}}$) are used in the empirical part of the paper.

The remaining part of the paper is organized as follows: Section 2 introduces the seasonal time series models, i.e., the component models. A total of six component models[4] are considered in this paper. Section 3 discusses the various ways of combining the forecasts made by the component models. Motivated by the unsatisfactory results from ensembles that only consider time series models, potential remedies are proposed and discussed in Section 5. Conclusions follow at the end.

2. Component models for time series ensemble

The time series ensemble in this paper includes six families of models: (1) the SARIMA family of models[5] (Box and Jenkins, 1994); (2) the ETS family of models[6] (Hyndman et al., 2008); (3) a version of

[1] The combination can be performed in one of two ways:

$$(a): \hat{y}_{t+h} = \frac{1}{N}\sum_{i=1}^{N} w_i f_i(\boldsymbol{x}_i), \text{ or}$$

$$(b): \hat{y}_{t+h} = f(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N),$$

where f denotes a forecasting model, w_i is the weight for model i. It is thus clear that in method (a), the final forecasts are produced by applying N different models on N datasets, respectively, and combining them. Method (b) trains one model that is best for all N datasets.

[2] In fact, the entire section on parameter diversity in Ren et al. (2015) is revolved around NWP models.

[3] MinT stands for minimum trace. Some statistical approach was used by Yang et al. (2017a) to optimize the reconciliation weights.

[4] As stated in the abstract, the total number of models is a lot more than six, due to the various parameter settings. For examples, there are 30 models in the ETS family alone, Akaike information criterion is used to select the best model. The winning model from the ETS family of models is considered as one component model here.

[5] A seasonal autoregressive integrated moving average model is defined by six process order parameters, p, d, q, P, D and Q. We set $d = 0, D = 1$, and the max values of $p = q = 2, P = Q = 1$, so that a total of $3 \times 3 \times 2 \times 2 = 36$ models are available for within-group AIC-based selection.

[6] The exponential smoothing models are described by three components, namely, seasonality, trend and error components. The seasonal component can be additive, multiplicative or none; the trend component can be additive, additive damped, multiplicative, multiplicative damped and none; and the error component can be either additive or multiplicative. In total, there are $3 \times 5 \times 2 = 30$ models.