Solar Energy 158 (2017) 332–346

Contents lists available at ScienceDirect

Solar Energy

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

Reconciling solar forecasts: Temporal hierarchy

Dazhi Yang[a,*], Hao Quan[b], Vahid R. Disfani[c], Carlos D. Rodríguez-Gallegos[d]

[a] Singapore Institute of Manufacturing Technology, Agency for Science, Technology and Research (A*STAR), Singapore
[b] Experimental Power Grid Centre, Institute of Chemical and Engineering Sciences, A* STAR, Singapore
[c] Department of Electrical Engineering, University of Tennessee at Chattanooga, United States
[d] Solar Energy Research Institute of Singapore, National University of Singapore, Singapore

ARTICLE INFO

Keywords:
Forecast post-processing
Temporal reconciliation
Numerical weather prediction
Operational forecasting

ABSTRACT

Previously in “Reconciling solar forecasts: Geographical hierarchy” [Solar Energy 146 (2017) 276–286], we have demonstrated that by reconciling the forecasts obtained from a geographical hierarchy (formed by levels such as transmission zones, distribution nodes, PV plants, subsystems and inverters), more accurate forecasts can be obtained as compared to forecasts produced using only the information from a single level. Furthermore, these reconciled forecasts are aggregate consistent, facilitating different power system operations and power plant management. In this paper, we continue our discussion on hierarchical reconciliation and consider another type of hierarchy, namely, temporal hierarchy.

A temporal hierarchy can be constructed by aggregating a high-frequency time series (e.g., hourly) into different low-frequency time series (e.g., 6-hourly and daily). While the measurements in a temporal hierarchy sum up consistently (e.g., the first two points in an hourly time series sum up to be the first point in the 2-hourly time series), the forecasts in general do not sum up, due to the different models used to generate forecasts. This inconsistency poses an apparent problem when selecting forecasts. Using reconciliation could mitigate the problem by reconciling different forecasts obtained at different timescales, and thus generating aggregate consistent forecasts. In addition, the reconciled forecasts reduce modeling uncertainties using combinations – similar to the ensemble forecasting used in numerical weather prediction (NWP) – and thus improve forecast accuracy.

The empirical part of the paper consists an NWP-based day-ahead forecast reconciliation example. Hourly data from 318 sites in California, over a period of 1 year, is used. Over 50 million forecasts are produced in this paper, in an operational forecasting context. Temporal reconciliation is shown to significantly outperform NWP forecasts produced by 3TIER, with a forecast skill (computed based on 3TIER forecasts) up to 0.3. Since reconciliation does not require any additional information, such improvement is thus encouraging.

1. Introduction

It is now well-known that the variability of solar irradiance, or other forms of it, depends on timescale. On this point, forecasting methods based on multiscale decomposition of time series, such as the multi aggregation prediction algorithm (Kourentzes et al., 2014), the theta model (Assimakopoulos and Nikolopoulos, 2000), wavelet decomposition (Zhang et al., 2017), empirical mode decomposition and many variants, have shown improved results in solar forecasting, as compared to the results of conventional time series modeling (e.g., Monjoly et al., 2017; Deo et al., 2016; Sharma et al., 2016; AlHakeem et al., 2015; Yang et al., 2015a, 2012; Mohammadi et al., 2015). The merit of these methods goes to their ability to model variability at different timescales separately, so that some nonlinear and nonstationary properties of the irradiance time series can be better handled. In general, these methods share the procedure outlined by Wang and Wu (2016): (1) raw time series is decomposed into a set of sub-series via decomposition algorithms; (2) forecasts are made using each of these sub-series; and (3) forecasts of individual sub-series are recombined to form the final forecasts.

A particular drawback of the above class of forecasting methods is that the recombination is often limited to simple addition of forecasts.

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

http://dx.doi.org/10.1016/j.solener.2017.09.055
Received 9 June 2017; Received in revised form 22 September 2017; Accepted 25 September 2017

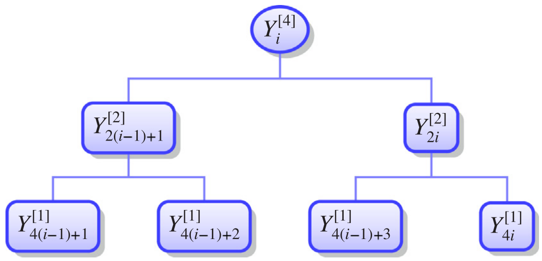

Fig. 1. An illustration of the indexing of time series in a hierarchy, with $m = 4$.

In other words, there is no weighting during the recombination, and thus the benefits of a better forecast sub-series cannot be amplified through weighting. Another drawback of these methods is that they only generate forecasts at the timescale of the original series. When the original time series is aggregated into new series with longer timescales, the forecasts produced by these series are not aggregated consistent,[1] largely owing to modeling errors.

In this paper, a temporal reconciliation framework (first proposed by Athanasopoulos et al., 2017) is applied to solar forecasting, which addresses both of the above drawbacks. The framework models a time series as hierarchical structures and produces consistent forecasts across all possible timescales of the given time series. Higher levels in the hierarchy contain more aggregated series, while lower levels contain more disaggregated series. By construction, the values at lower levels sum up to the values at higher levels. This construction is very similar to the ones reported in an earlier work (Yang et al., 2017b), where geographically distributed PV systems are modeled using a geographical hierarchy (formed by levels such as transmission zones, distribution nodes, PV plants, subsystems and inverters). To that end, many techniques introduced earlier can be transferred to the present work. For readers who are unfamiliar with the motivation, concepts and benefits of reconciling solar forecasts, we strongly encourage reading the first part of this work, namely, "Reconciling solar forecasts: Geographical hierarchy" [Solar Energy 146 (2017) 276–286].

The remaining part of the paper is organized as follows: Section 2 formulates the reconciliation techniques for temporal hierarchies. A day-ahead operational forecasting case study using data from 318 simulated PV systems in California is presented in Section 3. Reconciliation can be thought of as a post-processing technique. In Section 4, it is compared to two other commonly used post-processing techniques, namely, model output statistics and Kalman filter. Concluding remarks follow at the end. As usual, we provide the data and code necessary to reproduce some results shown in this paper in Appendix D.

2. The temporal hierarchical reconciliation framework

For a given seasonal – let us assume that for now – time series $\{Y_t\}$ where $t = 1,...,T$, containing the most disaggregated observations with sampling frequency m per cycle, T is a multiple of m, it can be aggregated into a new time series when all non-overlapping and consecutive k observations are combined, where k is a factor of m. Mathematically, the aggregated series is

$$Y_j^{[k]} = \sum_{t=1+(j-1)k}^{jk} Y_t, \tag{1}$$

for $j = 1,...,T/k$. The time series $\{Y_j^{[k]}\}$ in Eq. (1) has a seasonal period of $M_k = m/k$. In situations where T is not a multiple of m, the first few observations in $\{Y_t\}$ are discarded, so that the remaining observations form $\lfloor T/m \rfloor$ complete cycles. Eq. (1) becomes

$$Y_j^{[k]} = \sum_{t=t^*+(j-1)k}^{t^*-1+jk} Y_t, \tag{2}$$

where $j = 1,...,\lfloor T/m \rfloor$ and $t^* = T-\lfloor T/m \rfloor m + 1$. As for different k, the indexing for various aggregated time series has different ranges. To unify the indexing, we set $j = i$ for the most aggregated series, i.e., $\{Y_i^{[m]}\}$, and $i = 1,...,\lfloor T/m \rfloor$. In this way, each observation in time series of other levels of aggregation can be written as $Y_{M_k(i-1)+z}^{[k]}$, for $z = 1,...,M_k$. For every step increment in i, the index for other time series advances M_k steps, or one period. This is more illustrated by the $m = 4$ example shown in Fig. 1. For example, when $i = 2$, the time series element $Y_2^{[4]}$ corresponds to elements $\{Y_3^{[2]},Y_4^{[2]}\}$ and $\{Y_5^{[1]},Y_6^{[1]},Y_7^{[1]},Y_8^{[1]}\}$ in their respective time series.

If we denote all q factors of m, in descending order, as $\{k_q,...,k_1\}$, the temporal hierarchy can be defined with a k_q-aggregated top level, where $k_q = m$, and a k_1-aggregated bottom level, where $k_1 = 1$. In the above example with $m = 4$, $k \in \{4,2,1\}$. If we write those observations in the k-aggregated series indexed by i as a vector

$$\boldsymbol{Y}_i^{[k]} = (Y_{M_k(i-1)+1}^{[k]},Y_{M_k(i-1)+2}^{[k]},...,Y_{M_k i}^{[k]})^\top, \tag{3}$$

and stack all such vectors corresponding to the q factors[2]

$$\boldsymbol{Y}_i = (Y_i^{[k_q]},\boldsymbol{Y}_i^{[k_{q-1}]^\top},...,\boldsymbol{Y}_i^{[k_2]^\top},\boldsymbol{Y}_i^{[k_1]^\top})^\top, \tag{4}$$

we have

$$\boldsymbol{Y}_i = \boldsymbol{S}\boldsymbol{Y}_i^{[1]}, \tag{5}$$

where $\boldsymbol{S}$ denotes the summing matrix. In the $m = 4$ case above, $\boldsymbol{S}$ is given by

$$\boldsymbol{S} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{6}$$

The expressions in Eqs. (4) and (5) can be immediately related to Eqs. (1) and (3) of Yang et al. (2017b), which were used to describe the aggregation in a geographical hierarchy.

There is, however, a slight complication in dealing with temporal hierarchy. Since the aggregation of the time series is based on the factorization of m, the constructed hierarchy is often not *unique*.[3] For example, consider an hourly time series with diurnal periodicity ($m = 24$), we have $k \in \{24,12,8,6,4,3,2,1\}$. Two hierarchies can be defined in this case by disaggregating the observation $Y_i^{[24]}$ into either $\{Y_{2(i-1)+1}^{[12]},Y_{2i}^{[12]}\}$ or $\{Y_{3(i-1)+1}^{[8]},Y_{3(i-1)+2}^{[8]},Y_{3i}^{[8]}\}$. The corollaries below describe the uniqueness of a temporal hierarchy. Corollary 2 was stated by Athanasopoulos et al. (2017) without a proof; we provide a proof here.

Corollary 1. *For any given positive integer $m > 1$, there is a single unique temporal hierarchy only if $m = p^\alpha$, where α is a positive integer and p is a prime number.*

Proof. The unique factorization theorem states that any positive integer $n > 1$ can be written in the form $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$, where the powers α_1

[1] Let us consider an example here. An hourly time series can be aggregated into a 2-hourly series by adding the adjacent data points together. Suppose the total power in the next 2 h is of interest, we can produce this forecast either using the 2-hourly series directly (with a 1-step-ahead forecast) or using the sum of two 1 h forecasts (a 1-step-ahead forecast and a 2-step-ahead forecast) produced by the hourly series. Aggregate inconsistency refers to the situation where the sum of those two 1 h forecasts does not equal to the one 2 h forecast.

[2] Take note that only $Y_i^{[k_q]}$ is not a vector, as the top level contains a single series.

[3] Uniqueness of a hierarchy is only used to describe *full hierarchies*. For example, the hierarchy in Fig. 1 is a full hierarchy. However, if we take the middle level out, a "new" hierarchy is obtained, but it is not a full hierarchy. Since any non-prime number of members in a level would suggest the existence of a higher level, a full hierarchy must contain an $\mathscr{L}_1$ that contains prime number of members. $\mathscr{L}_1$ is the level immediately below the top level, i.e., $\mathscr{L}_0$.