Article

Forecasting Hierarchical Time Series in Power Generation

Tiago Silveira Gontijo [1],[*] and Marcelo Azevedo Costa [1],[2]

[1] Graduate Program in Industrial Engineering, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte 31270-901, MG, Brazil; azevedo@est.ufmg.br

[2] Department of Industrial Engineering, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte 31270-901, MG, Brazil

[*] Correspondence: tsgontijo@hotmail.com

check for
updates

Abstract: Academic attention is being paid to the study of hierarchical time series. Especially in the electrical sector, there are several applications in which information can be organized into a hierarchical structure. The present study analyzed hourly power generation in Brazil (2018–2020), grouped according to each of the electrical subsystems and their respective sources of generating energy. The objective was to calculate the accuracy of the main measures of aggregating and disaggregating the forecasts of the Autoregressive Integrated Moving Average (ARIMA) and Error, Trend, Seasonal (ETS) models. Specifically, the following hierarchical approaches were analyzed: (i) bottom-up (BU), (ii) top-down (TD), and (iii) optimal reconciliation. The optimal reconciliation models showed the best mean performance, considering the primary predictive windows. It was also found that energy forecasts in the South subsystem presented greater inaccuracy compared to the others, which signals the need for individualized models for this subsystem.

Keywords: power generation; electrical subsystems; time series

1. Introduction

The advent of Industry 4.0 revolutionized factories worldwide, since it allowed the connectivity between measuring machines and the automation of companies, distributing the capacity to collect massive volumes of data [1]. In high-level data analysis, forecasting models allow the extraction of behavior patterns, as well as the prediction of future values for the collected data set [2].

In the above-mentioned scenario, the construction of predictive models is gaining prominence in the literature [3–5], since economic agents deal with uncertainty in multiple spheres and aim to achieve the best results using available resources [6]. Developing acceptably accurate models presents a meaningful challenge, as prediction is a technique that deals with risk and there will always be a fundamental error associated with it. The best model is the one that most adequately represents the phenomenon of interest.

In relation to the object of our study, power generation, there are several forecasting applications: (i) classical time series models like the autoregressive moving average, autoregressive integrated moving average, and generalized autoregressive conditional heteroscedastic among others [7,8]; (ii) pre-processing techniques like spectrum analysis, wavelets, and Fourier analysis [9]; and, (iii) machine learning approaches such as neural networks, fuzzy systems, and support vector machine [10]. Alternatively, hybrid models aim to combine machine learning representations with different methods. These methods include focused time-delay neural networks [11], wavelet neuro-fuzzy systems [12], finite-impulse response neural networks [13], local feedback dynamic fuzzy

neural networks [14], type recurrent fuzzy networks [15], and neuro-fuzzy inference systems [16] among others.

Additionally, an alternative class known as hierarchical forecasting [17–19] deals with organized time series that can be aggregated at different levels into groups based on geography, sources of energy, or other, specific features. Despite this being a recent topic, there is already research that has addressed the use of hierarchical forecasting models in the energy sector. Examples of hierarchical forecasting include electrical grids [20], solar power generation [21], energy transport [22], short-term load forecasting [23], long-term load forecasting [24], energy consumption [25], and air pollution [26] among others.

The papers identified above have calibrated the forecasts using only the bottom-up, top-down or Ordinary least squares (OLS) assumptions [19]. Thus, the following research question is formulated: how is it possible to make hierarchical predictions using advanced linear regression models with regularization? In this way, it is expected to obtain more reliable forecasts by rewriting the hierarchical problem in terms of finding a set of unbiased, minimum variance measures of projected values across the whole array of data. It is possible to minimize the sum of variances of the reconciled estimate errors under the property of unbiasedness, using the procedure called MinT (minimum trace) reconciliation [27].

The present paper presents a case study using a power generation data set from Brazil (2018–2020) organized by electrical subsystems and different generating sources. Specifically, the main approaches used to aggregate and disaggregate predictions made for grouped time series are examined, namely: (i) bottom-up, (ii) top-down and (iii) optimal reconciliation models (OLS, WLS and MinT). The ARIMA and ETS predictive models were used to test the performance of these reconciliation methods, since these are the default models available in the R-package HTS. Further descriptions can be found in the materials and methods section.

The remainder of the present paper is organized as follows. Section 2 defines the study methodology, describing the data set, hierarchical procedures, and forecasting models employed. Section 3 presents the results and discussions of the techniques, in addition to the limitations of this paper. Finally, Section 4 presents the conclusions and guidelines for future work.

2. Materials and Methods

The secondary data used in this study correspond to the amounts of power generated by each of the Brazilian electrical subsystems (North, Northeast, Southeast/Midwest, and South). We separated these data according to the source of energy (wind, hydroelectric, thermal, solar, and nuclear). Data were obtained from the National Electric System Operator [28], due to their reliability. The observations of hourly power generation (GWh) were made during the period from January 2018 to January 2020, making a total of 17,521 h.

Based on Hyndman et al. [19], we present a schematic representation of the Brazilian energy generation system, comprising a three-level hierarchical structure (Figure 1). Level 0 represents the total energy generated in Brazil (completely aggregated series). Level 1 denotes each of Brazil's electrical subsystems (first level of disaggregation). The last level, Level 2, represents each of the energy generating sources (Level k). According to this framework, it is possible to identify the most disaggregated time series (in this case $k = 2$).

Table 1 shows the amounts of power generation in Brazil (GWh), according to generating sources and electrical subsystems. There is a predominance of hydroelectric generation (73%), making the Brazilian electrical matrix one of the cleanest in the world. At the same time, the Southeast/Midwest subsystem accounts for more than half (56%) of all energy generated in the country.
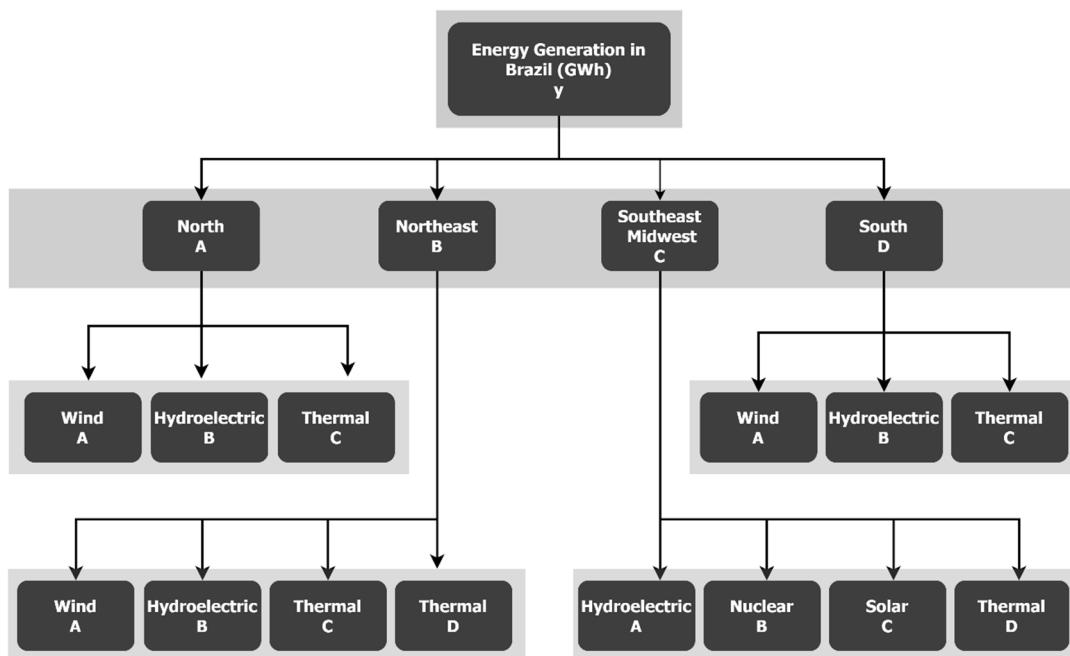
Figure 1. Hierarchical aggregation structure for the energy generation in Brazil.

Table 1. Amounts of power generation in Brazil (GWh).

Subsystem/Source		Wind	Hydro	Thermal	Solar	Nuclear	Total (GWh—Subsystem)	%
North	(A)	2688	125,182	31,489	0	0	159,359	**14.3%**
Northeast	(B)	85,377	37,705	36,699	4626	0	164,407	**14.7%**
Southeast/Midwest	(C)	0	518,714	73,555	2437	31,805	626,511	**56.1%**
South	(D)	11,326	135,914	19,472	0	0	166,712	**14.9%**
Total (GWh—Source)		99,391	817,516	161,215	7063	31,805	1,116,989	**100%**
%		**8.9%**	**73.2%**	**14.4%**	**0.6%**	**2.8%**	100%	-

Routines were implemented using the R® programming language [29]. The R-package HTS was used to calculate the bottom-up, top-down, optimal combination reconciliation and trace minimization reconciliation. HTS is available at: https://cran.r-project.org/web/packages/hts/index.html. Although HTS includes functions for creating, plotting and forecasting hierarchical time series, it has some limitations. Those limitations include the fact that it has only three built-in forecasting options: ARIMA, ETS, and random walks [19]. This paper will use the ARIMA and the ETS models since they have automatic adjustment and allow consideration of factors such as the trend and seasonality of the data set. The computer used to execute the algorithms had CPU Intel Core i5-7200 2.70 GHz, RAM of 16 GB, and operating system Windows 10 x64. In the next subsection, we present the hierarchical reconciliation models used in the present paper, as well as the forecasting models.

2.1. The Bottom-Up (BU) Approach

The BU procedure requires first providing forecasts for every series at the bottom-level, and then summing these to generate forecasts for all the levels of the hierarchical structure [30]. In its simplicity, this approach neglects the relations between time series and works, mainly unsuccessfully, on highly disaggregated data. These data tend to have a low signal-to-noise ratio [27]. According to the hierarchy (Figure 1), we first make h-step-ahead forecasts for all the bottom-level time series ($n = 14$):

$$\hat{y}_{AA,t}, \ \hat{y}_{AB,t}, \ \hat{y}_{AC,t}, \ \hat{y}_{BA,t}, \ \hat{y}_{BB,t}, \ \hat{y}_{BC,t}, \ \hat{y}_{BD,t}, \ \hat{y}_{CA,t}, \ \hat{y}_{CB,t}, \ \hat{y}_{CC,t}, \ \hat{y}_{CD,t}, \ \hat{y}_{DA,t}, \ \hat{y}_{DB,t}, \ \hat{y}_{DC,t}. \quad (1)$$

Summing these, we obtain h-step-ahead forecasts for the rest of the series:

$$
\begin{aligned}
\widetilde{y}_t &= \hat{y}_{AA,t} + \hat{y}_{AB,t} + \hat{y}_{AC,t} + \hat{y}_{BA,t} + \hat{y}_{BB,t} + \hat{y}_{BC,t} + \hat{y}_{BD,t} + \hat{y}_{CA,t} + \hat{y}_{CB,t} \\
&\quad + \hat{y}_{CC,t} + \hat{y}_{CD,t} + \hat{y}_{DA,t} + \hat{y}_{DB,t} + \hat{y}_{DC,t}. \\
\widetilde{y}_{A,t} &= \hat{y}_{AA,t} + \hat{y}_{AB,t} + \hat{y}_{AC,t}. \\
\widetilde{y}_{B,t} &= \hat{y}_{BA,t} + \hat{y}_{BB,t} + \hat{y}_{BC,t} + \hat{y}_{BD,t}. \\
\widetilde{y}_{C,t} &= \hat{y}_{CA,t} + \hat{y}_{CB,t} + \hat{y}_{CC,t} + \hat{y}_{CD,t}. \\
\widetilde{y}_{D,t} &= \hat{y}_{DA,t} + \hat{y}_{DB,t} + \hat{y}_{DC,t}.
\end{aligned}
\tag{2}
$$

According to [19], it is possible to arrange the equations expressed in (2) into an algebra notation. Below is a complete notation for this problem:

$$
\begin{bmatrix}
\widetilde{y}_t \\
\widetilde{y}_{A,t} \\
\widetilde{y}_{B,t} \\
\widetilde{y}_{C,t} \\
\widetilde{y}_{D,t} \\
\widetilde{y}_{AA,t} \\
\widetilde{y}_{AB,t} \\
\widetilde{y}_{AC,t} \\
\widetilde{y}_{BA,t} \\
\widetilde{y}_{BB,t} \\
\widetilde{y}_{BC,t} \\
\widetilde{y}_{BD,t} \\
\widetilde{y}_{CA,t} \\
\widetilde{y}_{CB,t} \\
\widetilde{y}_{CC,t} \\
\widetilde{y}_{CD,t} \\
\widetilde{y}_{DA,t} \\
\widetilde{y}_{DB,t} \\
\widetilde{y}_{DC,t}
\end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\hat{y}_{AA,t} \\
\hat{y}_{AB,t} \\
\hat{y}_{AC,t} \\
\hat{y}_{BA,t} \\
\hat{y}_{BB,t} \\
\hat{y}_{BC,t} \\
\hat{y}_{BD,t} \\
\hat{y}_{CA,t} \\
\hat{y}_{CB,t} \\
\hat{y}_{CC,t} \\
\hat{y}_{CD,t} \\
\hat{y}_{DA,t} \\
\hat{y}_{DB,t} \\
\hat{y}_{DC,t}
\end{bmatrix}
\tag{3}
$$

Alternatively, the notation presented in (3) can be reformulated in a compact way by applying the summing matrix. Thus, the bottom-up approach can be represented as:

$$
\widetilde{y}_t = S\hat{b}_t,
\tag{4}
$$

where $\widetilde{y}_t$ is an *n*-dimensional vector of *h*-step-ahead forecasts for the total energy, *S* is the summing matrix, and $\hat{b}_t$ is an *m*-dimensional vector of *h*-step-ahead forecasts for each of the sources of energy at bottom-level. An advantage of this procedure is that we are forecasting at the bottom-level of a hierarchy. Consequently, no information is missed due to aggregation [17].

2.2. The Top-Down (TD) Approach

Top-down methods operate with strictly hierarchical aggregation structures, not with grouped structures. They involve first making forecasts for the Total level y_t, and next disaggregating these down the hierarchy [17]. Let $p_1, \ldots, p_m$ be a set of disaggregation proportions that deliver the forecasts of the Total series, which are to be distributed in order to obtain forecasts for all series at the bottom-level

of the structure. To illustrate, concerning our hierarchy by applying proportions to Figure 1, we get $p_1, \ldots, p_{14}$:

$$\widetilde{y}_{AA,t} = p_1\hat{y}_t, \ \widetilde{y}_{AB,t} = p_2\hat{y}_t, \ \widetilde{y}_{AC,t} = p_3\hat{y}_t.$$
$$\widetilde{y}_{BA,t} = p_4\hat{y}_t, \ \widetilde{y}_{BB,t} = p_5\hat{y}_t, \ \widetilde{y}_{BC,t} = p_6\hat{y}_t, \ \widetilde{y}_{BD,t} = p_7\hat{y}_t.$$
$$\widetilde{y}_{CA,t} = p_8\hat{y}_t, \ \widetilde{y}_{CB,t} = p_9\hat{y}_t, \ \widetilde{y}_{CC,t} = p_{10}\hat{y}_t, \ \widetilde{y}_{CD,t} = p_{11}\hat{y}_t.$$
$$\widetilde{y}_{DA,t} = p_{12}\hat{y}_t, \ \widetilde{y}_{DB,t} = p_{13}\hat{y}_t, \ \widetilde{y}_{DC,t} = p_{14}\hat{y}_t.$$

(5)

This can be rewritten using matrix notation. If we stack the set of proportions in an m-dimensional vector $p = (p_1, \ldots, p_m)'$, we have the bottom-level h-step-ahead predictions. Overall, for a given set of proportions, top-down approaches can be written as:

$$\widetilde{b}_t = p_j\hat{y}_t.$$
$$\widetilde{y}_t = Sp_j\hat{y}_t.$$

(6)

The main TD models stipulate disaggregation proportions according to the historical proportions of the data. Among the main models of this approach, we highlight the following three: (i) top-down Gross–Sohl method A (TDGSA), (ii) top-down Gross–Sohl method F (TDGSF), and (iii) Top-down forecast proportions (TDFP) (Table 2). Additional details and demonstrations of Table 2 can be obtained from [18,31].

Table 2. TD disaggregation proportions according to the historical proportions of the data.

TD Gross-Sohl Method A TDGSA	TD Gross-Sohl Method F TDGSF	TD Forecast Proportions TDFP
$p_j = \frac{1}{T}\sum\limits_{t=1}^{T}\frac{y_{j,t}}{y_t}$	$p_j = \sum\limits_{t=1}^{T}\frac{y_{j,t}}{T} / \sum\limits_{t=1}^{T}\frac{y_t}{T}$	$p_j = \prod\limits_{l=0}^{K-1}\frac{\hat{y}_{j,t}^{(l)}}{\hat{S}_{j,t}^{(l+1)}}$
for $j = 1, \ldots, m$. Each proportion p_j reflects the average of the historical proportions of the bottom-level series $y_{j,t}$, t over the period $t = 1, \ldots, T$ relative to the total aggregate y_t.	for $j = 1, \ldots, m$. Each proportion p_j takes the average historical value of the bottom-level series $y_{j,t}$ related to the average value of the total aggregate y_t.	where $j = 1, \ldots, m$, $\hat{y}_{j,h}^{(l)}$ is the h-step-ahead forecast and $\hat{S}_{j,t}^{(l)}$ is the sum of the h-step-ahead forecasts below the node that is l levels above node j.

2.3. The Optimal Reconciliation Approaches

The optimal reconciliation approach proposed by [19] consists of an ordinary least squares problem based on the calculation of independent projections for all hierarchical levels, then applying a regression model to optimize the combination of these forecasts. According to [32], we can write the base prediction as:

$$\hat{y}_{t+h|t} = S\beta_{t+h|t} + \varepsilon_h,$$

(7)

where $\beta_{t+h|t}$ represents the unknown conditional mean of the most disaggregated series, and ε_h is the error with mean of zero and covariance matrix $\sum_h$. If $\sum_h$ were known, the estimator of $\beta_{t+h|t}$ would lead to the following weighted least squares, producing reconciled forecasts, as follows:

$$\widetilde{y}_{t+h|t} = S\hat{\beta}_{t+h|t} = S\left(S'\sum\limits_h^{-1}S\right)^{-1}S'\sum\limits_h^{-1}\hat{y}_{t+h|t} = SP\hat{y}_{t+h|t},$$

(8)

where $P = \left(S'\sum_h^{-1}S\right)^{-1}S'\sum_h^{-1}S$. If the base forecasts $\hat{y}_{t+h|t}$ are unbiased, then the reconciled forecasts $\widetilde{y}_{t+h|t}$ will be unbiased, provided that $SPS = S$ [19]. This condition is valid for this reconciliation procedure for the bottom-up, although not for the top-down, methods. Consequently, the top-down approaches will never give unbiased reconciled forecasts, even if the base forecasts are unbiased.

Additionally, [27] proved that, in general, $\sum h$ is not known and not identifiable. The covariance matrix of the *h*-step-ahead reconciled forecast errors is given by the following expression:

$$Var_{(y_{t+h}-\widetilde{y}_{t+h|t})} = SPW_hP'S',$$ (9)

for any P such that $SPS = S$, then $W_h = Var_{(y_{t+h}-\hat{y}_{t+h|t})} = E(\hat{e}_{t+h|t}\hat{e}'_{t+h|t})$ is the covariance matrix of the corresponding *h*-step ahead base forecast errors. The purpose is to get the matrix P that minimizes the error variances of the reconciled forecasts which are on the diagonal of the covariance matrix $Var_{(y_{t+h}-\widetilde{y}_{t+h|t})}$. Finally, [27] demonstrated that the optimal reconciliation matrix P that minimizes the trace of $SPW_hP'S' =$, such that $SPS = S$, and the optimal reconciled forecasts, respectively, are given by:

$$P = (S'W_h^{-1}S)^{-1}S'W_h^{-1}$$
$$\widetilde{y}_{t+h|t} = S(S'W_h^{-1}S)^{-1}S'W_h^{-1}\hat{y}_{t+h|t},$$ (10)

which is introduced as the MinT (minimum trace) estimator. The next step consists of estimating W_h, a matrix of order *n*. Wickramasuriya, Athanasopoulos and Hyndman [27] proposed the following procedures (Table 3) to obtain the matrix:

Table 3. Hierarchical forecasting for electricity generation based on the ARIMA procedure.

Procedure	Description
OLS	$W_h = k_hI$, $\forall h$ where $k_h > 0$. This is the most simplifying premise, and collapses the MinT estimator to the OLS estimator, proposed by Hyndman et al. [19]. This is optimal when the base forecast errors are uncorrelated and equivariant.
WLSv	$W_h = k_h\text{diag}(\widehat{W}_1)$, $\forall h$ where $k_h > 0$ and: $\widehat{W} = \frac{1}{T}\sum_{t=1}^{T}\hat{e}_t(1)\hat{e}_t(1)'$, is the unbiased sample covariance estimator of the in-sample one-step-ahead base forecast errors. In this case, we can describe MinT as a WLS estimator applying variance scaling [27].
WLSs	$W_h = k_h\Lambda$, $\forall h$ where $k_h > 0$ and $\Lambda = \text{diag}(S1)$ with 1 being a unit column vector of dimension *n*. We assume that each of the bottom-level base forecast errors has a variance k_h and is uncorrelated between nodes. Consequently, every element of the diagonal Λ matrix receives the number of forecast error variances contributing to that aggregation level [27]. This estimator depends only on the grouping structure of the hierarchy.
MinT (Sample)	$W_h = k_w\widehat{W}_1$, $\forall h$ where $k_h > 0$, the unrestricted sample covariance estimator for $h = 1$ [27]. In the results section, we denote this as MinT (Sample).
MinT (Shrink)	$W_h = k_w\widehat{W^*_{1,D}}$; $\forall h$; $k_h > 0$; $W^*_{1,D} = \lambda_D\widehat{W}_{1,D} + (1 - \lambda_D)\widehat{W}_1$, is a shrinkage estimator with diagonal target, $\widehat{W}_{1,D}$, which is a diagonal matrix comprising the diagonal entries of $\widehat{W}_1$, and λ_D is the shrinkage intensity parameter. Thus, off-diagonal elements of $\widehat{W}_1$ are shrunk toward zero and diagonal elements (variances) remain unchanged [27]. Wickramasuriya, Athanasopoulos and Hyndman [27] suggested a scale and location invariant shrinkage estimator by parameterizing the shrinkage in terms of variances and correlations: $\hat{\lambda}_D = \frac{\sum_{i\neq j}\hat{var}_{(\hat{r}_{ij})}}{\sum_{i\neq j}\hat{r}^2_{ij}}$, where $\hat{r}_{ij}$ is the *ij*th element of $\hat{R}_1$, the 1-step-ahead sample correlation matrix to shrink it toward an identity matrix.

Source: adapted by authors from: [27].

2.4. ARIMA and ETS Formulation

ARIMA is one of the most-widely-used time series approaches for forecasting power generation [33]. Although studies have shown that ETS outperforms ARIMA [34], it is recommended to keep ARIMA

as a reference model during the forecasting process. Moreover, several statistical software packages, like R®, provide automatic model identification and parameter estimation skills for both ARIMA and ETS [17]. Professor Hyndman [19] developed the HTS package initially based on these predictive models. The present paper aims to test different approaches to optimal forecast reconciliation and, to do so, only the ARIMA and ETS models will be used. It is recommended that future studies extend these forecasting procedures using different predictive models, such as machine learning ones.

ARIMA was proposed by [33]. It is a linear forecasting method for dealing with stationary time series [34]. In the initial step, a time series is built stationary by differencing d times along with some nonlinear transformations, such as logging [34]. The consequential data are recognized as a linear function of past p data values and q errors (11), i.e., modeled as an autoregressive moving average (ARMA) model,

$$y_t = \varnothing_1 y_{t-1} + \varnothing_2 y_{t-2} + \ldots + \varnothing_p y_{t-p} + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + \ldots + \Theta_q \varepsilon_{q-1}, \tag{11}$$

where y_t denotes real value at time t, ε_t describes the error sequence: it is supposed to be white noise and Gaussian distributed $(0, \sigma^2)$. $\varnothing_i$ for $(i = 1, 2, \ldots, p)$ are autoregressive (AR) coefficients and Θ_j for $(j = 1, 2, \ldots, q)$ are moving average (MA) coefficients. p and q are integers referred to as model orders. The time series model is denoted as $ARIMA(p, d, q)$ [35,36].

According to [34], the group of exponential smoothing methods utilizes the principle of weighted averages of past information for making forecasts. Since its formulation in 1950, a variety of exponential smoothing methods have been developed. All exponential smoothing methods were initially classified by [37], which has been continued by [38–40]. ETS stands for error, trend, and seasonality elements. As pointed by [34], the usual representation for these patterns involves a state vector $x_t = (l_t, b_t, s_t, s_{t-1}, \ldots, s_{t-m+1})'$, and the state space equations [39] have the resulting structure:

$$\begin{aligned} y_t &= w(x_{t-1}) + r(x_{t-1})\varepsilon_t \\ x_t &= f(x_{t-1}) + g(x_{t-1})\varepsilon_t, \end{aligned} \tag{12}$$

where (ε_t) denotes a Gaussian white noise $(0, \sigma^2)$ and $\mu_t = w(x_t - 1)$. The model with additive error has $r_t(x_t - 1) = 1$, so $y_t = \mu_t + \varepsilon_t$. The model with multiplicative errors has $r_t(x_t - 1) = \mu_t = \mu_t$, so $y_t = \mu_t(1 + \varepsilon_t)$. Consequently, $\varepsilon_t = (y_t - \mu_t)/\mu_t$ is a relative error for the multiplicative model and any value of $r_t(x_t - 1)$ will lead to the identical point forecast for y_t [34,39].

2.5. Evaluating Forecast Accuracy

According to [20], there are several accuracy metrics, such as mean absolute percentage error (*MAPE*), mean absolute error (*MAE*), mean absolute scaled error (*MASE*), or root-mean-square error (*RMSE*), to evaluate the performance of point prediction methods, defined as follows:

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{y_t - \hat{y}_t}{y_t} \right|. \tag{13}$$

$$MAE = \frac{1}{T} \sum_{t=1}^{T} \left| y_t - \hat{y}_t \right|. \tag{14}$$

$$MASE = \frac{MAE}{MAE_{in-sample,naive}} \tag{15}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{y}_t)^2}, \tag{16}$$

where y_t is the amount of power generation at time t, $\hat{y}_t$ is the fitted value for power generation, and $MAE_{in-sample,naive}$ is the MAE generated by a naive forecast.

Specifically, in studies of hierarchical time series, the *MAPE* indicator appears the most frequently in the literature [41–43]. *MAPE* was also the selected metric for the present paper (Figures 2 and 3). Complementarily, *MAE*, *MASE*, and *RMSE* were estimated, and the results can be found in the Appendix A (Figures A3 and A4). The values of the *MAPE*, *MAE*, *MASE* and *RMSE* statistics were obtained using a weighted average, with proportions from Table 1.

		Predictive model: Autoregressive integrated moving average (ARIMA)									
MAPE		Forecast horizon (h)									
		1	2	3	4	5	6	7	8	9	Mean
		Hierarchical level 0 : Total - Brazil									
Method	BU	2.00	3.53	5.62	8.04	**10.17**	**11.45**	**11.17**	**10.76**	**10.48**	8.14
	TDGSA	2.07	3.76	6.10	8.79	11.25	12.87	12.95	12.72	12.63	9.24
	TDGSF	2.07	3.76	6.10	8.79	11.25	12.87	12.95	12.72	12.63	9.24
	TDFP	2.07	3.76	6.10	8.79	11.25	12.87	12.95	12.72	12.63	9.24
	OLS	1.98	3.65	5.94	8.59	10.99	12.54	12.55	12.23	12.06	8.95
	WLSv	1.91	3.51	5.70	8.24	10.51	11.93	11.79	11.32	10.99	8.43
	WLSs	1.88	3.46	5.63	8.15	10.39	11.77	11.59	11.09	10.76	8.30
	MintT (Sample)	**1.68**	**3.29**	**5.50**	**8.02**	10.24	11.57	11.31	10.85	10.60	**8.12**
	MinT (Shrink)	1.74	3.36	5.59	8.12	10.35	11.69	11.44	10.94	10.69	8.21
		Hierarchical level 1 : Electrical subsystems									
Method	BU	1.97	3.64	6.12	8.75	10.78	**11.93**	**11.90**	**11.70**	**11.88**	**8.74**
	TDGSA	31.97	31.74	30.37	28.93	28.12	27.49	26.71	26.04	25.36	28.53
	TDGSF	32.38	32.14	30.71	29.21	28.21	27.46	26.71	26.06	25.41	28.70
	TDFP	1.86	3.88	6.68	9.89	**9.89**	12.52	14.19	14.45	14.34	9.75
	OLS	1.90	3.55	6.30	9.20	11.70	13.36	13.64	13.56	13.66	9.65
	WLSv	1.77	3.35	5.84	8.62	10.84	12.38	12.56	12.40	12.41	8.91
	WLSs	1.81	3.41	5.92	8.74	11.00	12.57	12.79	12.68	12.75	9.07
	MintT (Sample)	**1.64**	**3.20**	**5.66**	**8.50**	10.76	12.23	12.40	12.21	12.20	8.76
	MinT (Shrink)	1.66	3.28	5.75	8.57	10.84	12.33	12.50	12.31	12.28	8.83
		Hierarchical level 2 : Energy sources									
Method	BU	2.66	5.05	**6.53**	**7.71**	**8.88**	**9.46**	**9.40**	**9.22**	**9.11**	**7.56**
	TDGSA	46.33	44.34	41.72	40.35	39.87	39.29	38.44	37.52	36.58	40.49
	TDGSF	47.66	45.70	42.87	41.24	40.42	39.64	38.80	37.90	36.96	41.24
	TDFP	2.83	5.51	7.53	9.45	9.45	11.46	12.79	13.20	13.33	9.50
	OLS	2.51	5.07	6.78	8.29	9.78	10.62	10.73	10.63	10.56	8.33
	WLSv	2.60	5.11	6.74	8.09	9.42	10.18	10.21	10.07	9.97	8.04
	WLSs	2.56	4.98	6.64	8.00	9.31	10.00	9.95	9.70	9.63	7.86
	MintT (Sample)	**2.48**	**4.96**	6.58	7.91	9.27	10.10	10.22	10.10	10.00	7.96
	MinT (Shrink)	2.52	5.04	6.68	8.02	9.38	10.20	10.30	10.19	10.10	8.05

Figure 2. Hierarchical forecasting for electricity generation based on the ARIMA procedure (MAPE). (Note: The performance was indicated into a color scale, where green means better values for calculated accuracy, and red means worse accuracy. The intermediate values are colored yellow.).

MAPE		Predictive model: Error, trend, seasonality (ETS)									
		Forecast horizon (h)									
		1	2	3	4	5	6	7	8	9	Mean
		Hierarchical level 0 : Total - Brazil									
Method	BU	2.60	4.89	7.73	10.90	13.79	15.78	16.20	16.28	16.48	11.63
	TDGSA	2.11	4.21	6.90	9.95	12.73	14.64	14.98	14.99	15.12	10.62
	TDGSF	2.11	4.21	6.90	9.95	12.73	14.64	14.98	14.99	15.12	10.62
	TDFP	2.11	4.21	6.90	9.95	12.73	14.64	14.98	14.99	15.12	10.62
	OLS	2.13	4.24	6.93	9.98	12.77	14.68	15.03	15.03	15.17	10.66
	WLSv	2.26	4.42	7.16	10.25	13.06	15.00	15.37	15.39	15.55	10.94
	WLSs	2.29	4.46	7.20	10.29	13.11	15.05	15.42	15.45	15.61	10.99
	MintT (Sample)	**2.06**	**4.14**	**6.80**	**9.84**	**12.61**	**14.51**	**14.84**	**14.84**	**14.96**	**10.51**
	MinT (Shrink)	2.07	4.15	6.82	9.86	12.63	14.53	14.87	14.87	14.99	10.53
		Hierarchical level 1 : Electrical subsystems									
Method	BU	2.59	4.94	8.09	11.43	14.21	16.13	16.71	16.93	17.36	12.04
	TDGSA	31.98	31.78	30.44	29.05	28.29	27.76	27.08	26.51	25.91	28.76
	TDGSF	32.38	32.17	30.78	29.32	28.37	27.72	27.08	26.53	25.96	28.92
	TDFP	2.17	4.37	7.37	10.58	**10.58**	**13.24**	15.05	15.54	15.68	10.51
	OLS	2.15	4.34	7.35	10.55	13.22	15.04	15.53	15.68	16.04	11.10
	WLSv	2.30	4.55	7.60	10.85	13.56	15.42	15.94	16.11	16.49	11.42
	WLSs	2.27	4.50	7.54	10.79	13.49	15.34	15.85	16.03	16.40	11.36
	MintT (Sample)	**1.89**	**3.98**	**6.92**	**10.09**	12.74	14.55	**15.03**	**15.20**	**15.56**	**10.66**
	MinT (Shrink)	1.94	4.04	7.00	10.17	12.83	14.64	15.13	15.29	15.64	10.74
		Hierarchical level 2 : Energy sources									
Method	BU	3.13	6.56	9.01	11.16	13.24	14.40	14.74	15.21	15.63	11.45
	TDGSA	46.34	44.42	41.84	40.48	40.00	39.45	38.69	37.87	37.00	40.68
	TDGSF	47.66	45.77	42.98	41.35	40.54	39.77	38.99	38.17	37.32	41.40
	TDFP	2.90	6.31	8.74	10.97	**10.97**	**13.05**	14.29	14.58	14.94	10.75
	OLS	3.14	6.60	9.07	11.30	13.38	14.62	14.92	15.31	15.74	11.56
	WLSv	2.76	6.08	8.42	10.56	12.53	13.72	13.97	14.18	14.56	10.75
	WLSs	3.13	6.57	9.02	11.21	13.29	14.49	14.79	15.21	15.63	11.48
	MintT (Sample)	2.67	5.88	8.10	**10.10**	11.99	13.05	**13.20**	**13.47**	**13.81**	**10.25**
	MinT (Shrink)	**2.63**	**5.89**	**8.17**	10.21	12.14	13.25	13.43	13.72	14.09	10.39

Figure 3. Hierarchical forecasting for electricity generation based on the ETS procedure. (Note: The performance was indicated into a color scale, where green means better values for calculated accuracy, and red means worse accuracy. The intermediate values are colored yellow.).

3. Results and Discussion

Figure 2, below, shows the predictive result obtained, using the ARIMA model, considering a predictive window of nine hours ($h = 1, \ldots, 9$). Note that the model was estimated, taking the main hierarchical adjustment approaches into account, for the following levels: (i) total power generation in Brazil (Level 0), (ii) total energy generation by electrical subsystem (Level 1), and (iii) total energy generation by the energy generating source (Level 2). For Level 1, four forecasts (one for each electrical subsystem) were estimated. For Level 2, 14 forecasts (one for each energy source) were estimated.

Therefore, we estimated 1539 predictive models satisfying the following proportions: (i) 81 models for Level 0, (ii) 324 models for Level 1, and (iii) 1134 models for Level 2. The MAPE calculation for Levels 1 and 2 was based on a weighted average of the predictive errors. The weighting factors used are shown in Table 1.

The performance of each predictive model, divided by the forecast horizon, is illustrated by a color scale. The green colors indicate the most accurate forecasts, while the red colors symbolize less accurate forecasts. The best forecasts, for each of the predictive horizons, are highlighted in bold. The last column of Table 1 presents the average performance for each forecast horizon (*h*) for each hierarchical approach.

As pointed by [27], the MinT procedure has a useful feature: it systematizes results into a unique analytical solution that incorporates information about the correlation structure of the entire dataset. Additionally, the minimum trace reconciliation, with or without regularization, presented the best results of all linear reconciliation methods, such as OLS and WLS, with variations. Moreover, the MinT (Sample) approach returns the most accurate, coherent forecasts for all levels considering just the first forecast horizons. However, as the predictive window grows, the BU method becomes more accurate. Furthermore, the performance of the BU model increases as the time series disaggregate.

As expected, the results obtained using the top-down technique did not present good predictive results, since it is intended to generate forecasts for level 0, with worse accuracy for the other levels. Both BU and TD present disadvantages: they do not take the correlation among the series at each level into account.

The other accuracy metrics presented in the Appendix A (*MAE*, *MAE*, and *RMSE*) reinforce the results found. In general, the performance of the optimal reconciliation models, by trace minimization, provides more uniform estimates and better predictive potential for the first hours of the predictive horizon (Figures A3 and A4).

In addition to the ARIMA predictive model, Figure 3 presents the same forecasting procedures. However, they are based on the ETS automatic adjustment model. The objective is to show the influence of different forecasting methods for each hierarchical reconciliation model. In general, the error percentage produced by the ETS model was slightly higher than that produced by the ARIMA model. Figure 3 also shows the influence of trace minimization procedures (MinT) on the improvement of predictive performance. In particular, the MinT models have good predictive performance, even with the increase of the forecast horizon hours.

The average performance of the trace minimization (MinT) models shows stability, considering all hierarchical levels. As shown in Figure 2, the ETS-based predictive model shares some similarities with the ARIMA model. The BU technique is better for the most disaggregated levels, whereas the TD technique stands out only at the more aggregated levels. Note that the trace minimization procedures show significant gains over the classic linear models, namely OLS and WLS.

Figures 2 and 3 present some limitations. In general, it is not possible to test the predictive influence of each of the subsystems within the established forecast horizon. To show this problem, Figure 4 presents a predictive comparison (MAPE) for each of the Brazilian electrical subsystems, considering the nine-hour predictive horizon. On the left is the technique with the best aggregation/disaggregation performance (BU) for the ARIMA model. On the right is the technique with the best average performance (MinT) for the ETS automatic selection model.

Figure 4 thus shows a negative influence of the "south" electrical subsystem in the global measures of accuracy, especially from a predictive horizon of three hours onward. This system should be analyzed more thoroughly to identify energy sources located in the "south" subsystem that contributed most to the predictive instability of this system. Simultaneously, the use of individualized predictive models for this "south" system can be a good strategy, since unique climatic conditions exist in southern Brazil.
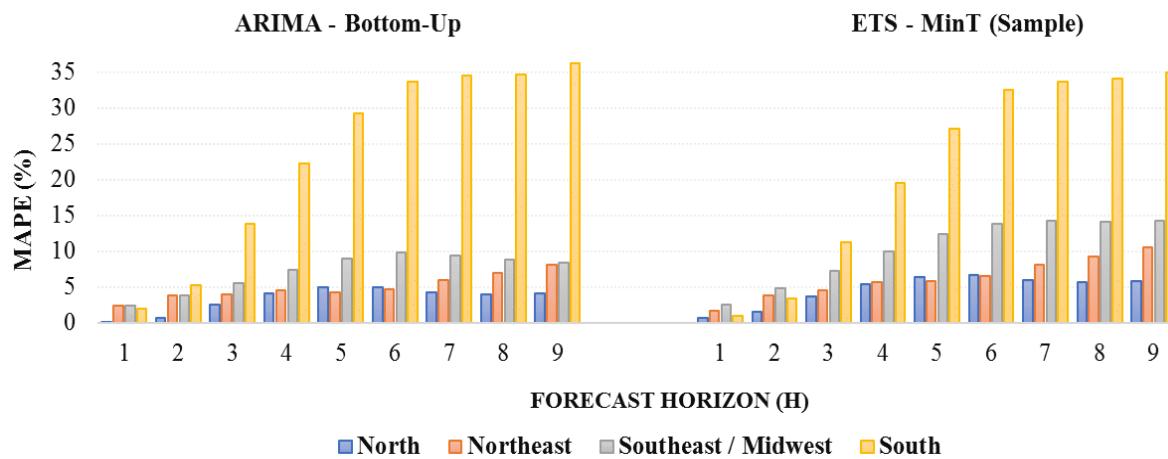
Figure 4. Hierarchical forecasting for power generation: electrical subsystem versus forecast horizon.

Figures A1 and A2 (Appendix A) present the accuracy measure of the ARIMA and ETS models in detail, considering energy sources versus electrical subsystems. These results reinforce those in Figure 4, indicating instability in the southern subsystem, especially wind energy data.

Finally, some limitations of the present paper are recognized here. First, predictive models are based on past information evaluable, so the presented results cannot be extrapolated for different contexts and other time periods. Additionally, it is necessary to incorporate other predictive models to make the results more robust. In future research, it is recommended that models which integrate high-frequency data, e.g., the Wavelet approach, be adopted.

4. Conclusions

Analysis of the energy market is complicated. It involves the relationship between forecasting models and uncertainty, distinctly regarding the stochastic behavior of variables. The present paper is aimed at policymakers, offering a forecasting tool that deals with grouped time series. It also proposes a new forecasting approach, based on hierarchical modeling of the energy generation in Brazil.

The present paper introduces the use of trace minimization procedures (MinT) to aggregate and disaggregate forecasts based on the ARIMA and ETS models. MinT models performed better than the classic linear approaches, such as OLS and WLS. The MinT models also have high reliability for short predictive horizons. It is noteworthy that both hierarchical procedures and forecasting methods influence the predictive values of power generation in Brazil. Despite its advantages, the optimal reconciliation approach also has some limitations. This method could be unduly influenced by the sample period, and thus its ranking might change for other periods.

Therefore, the use of other predictive models, such as those based on analogs, machine learning, and other hybrid techniques, for example, is recommended. For future research, fine-tuning forecasts of the "south" electrical subsystem, as well as testing the accuracy of the hierarchal methods by using new forecasting approaches, is also recommended.

Finally, the present study contributes to the energy planning processes of different agents, given that understanding energy generation patterns is singularly important for minimizing risks and supporting reliable production planning. Good forecasts for future energy generation can support operational arrangements since energy supply and demand impact spot market sales prices.

Abbreviations

The following abbreviations are used in this manuscript:

ARIMA	Autoregressive integrated moving average model
BU	Bottom-up
ETS	Error, trend, and seasonality model
GWh	Gigawatt hours
MAPE	Mean absolute percentage error
MinT	Minimum trace reconciliation
OLS	Ordinary least squares
ONS	Operator of the National System
TD	Top-down
TDFP	Top-down forecast proportions
TDGSA	Top-down Gross-Sohl method A
TDGSF	Top-down Gross-Sohl method F
WLS	Weighted least squares

Nomenclature

The following nomenclature is used in this manuscript:

k	Level of disaggregation
h	Forecast horizon
$\hat{b}_t$	m-dimensional vector of h-step-ahead forecasts
$\beta_{t+h|t}$	Unknown conditional mean of the most disaggregated series
ε_h	Error for each forecast horizon
$\sum_h$	Covariance matrix
p	Set of proportions in an m-dimensional vector
p_j	The average of the historical proportions
S	Summing matrix
$\hat{S}_{j,t}^{(l)}$	The sum of the h-step-ahead forecasts for TD
W_h	Covariance matrix of the corresponding h-step ahead base forecast errors
y_t	Total level of power generation
$\widetilde{y}_t$	an n-dimensional vector of h-step-ahead forecasts
$\hat{y}_{j,h}^{(l)}$	The h-step-ahead forecast for TD
$\widetilde{y}_{t+h|t}$	Reconciled forecasts
$\hat{\lambda}_D$	Shrinkage estimator

Appendix A

ARIMA
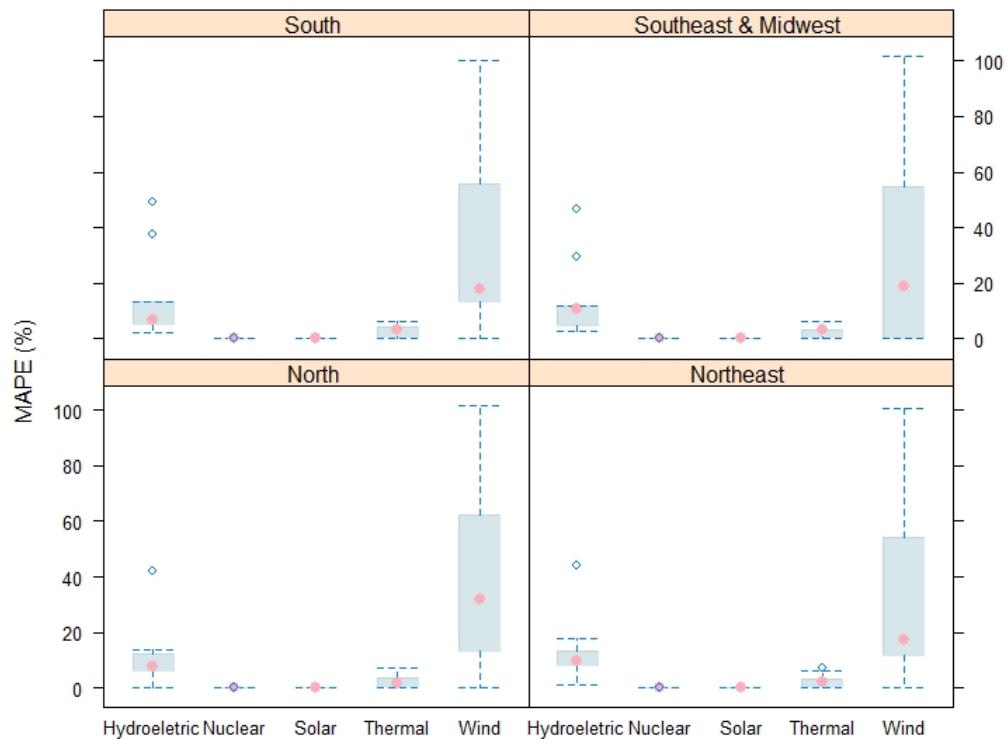
Figure A1. Hierarchical forecasting for power generation: electrical subsystem *versus* generating source.

ETS
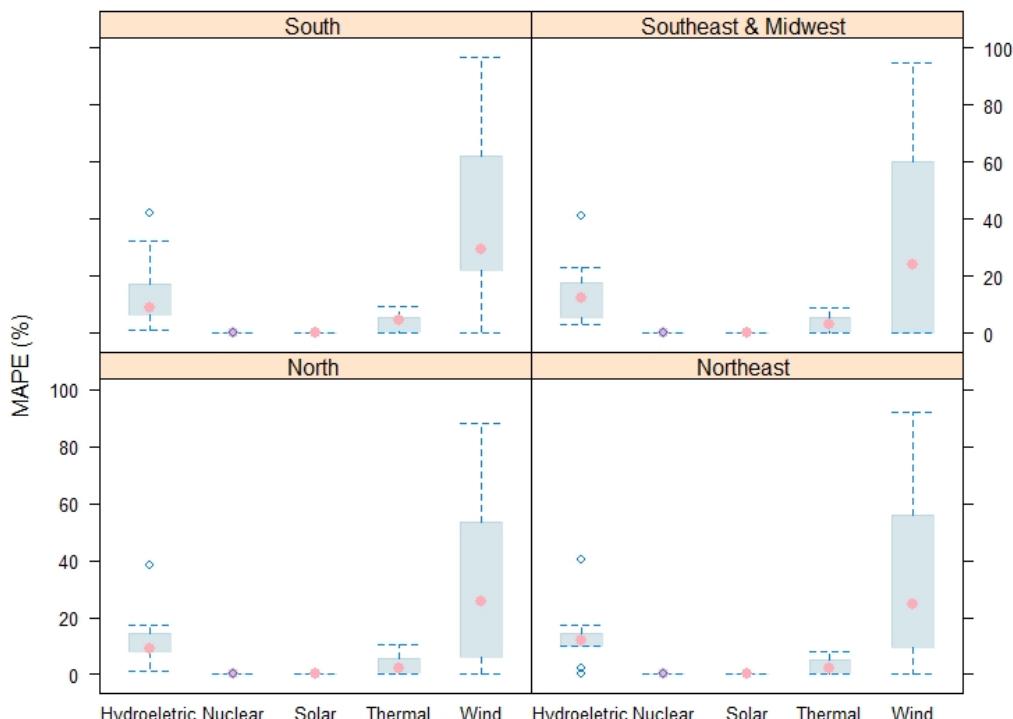
Figure A2. Hierarchical forecasting for power generation: electrical subsystem *versus* generating source.

		RMSE				MAE				MASE			
		Predictive model: Autoregressive integrated moving average (ARIMA)										
		Forecast horizon (h) - mean of the interval										
		[1-3]	[4-6]	[7-9]	Mean	[1-3]	[4-6]	[7-9]	Mean	[1-3]	[4-6]	[7-9]	Mean
		Hierarchical level 0 : Total - Brazil										
Method	BU	2.36	7.77	**8.50**	6.21	2.20	**6.46**	**7.14**	5.27	1.11	**3.27**	**3.62**	2.67
	TDGSA	2.47	8.07	9.33	6.62	2.35	7.11	8.35	5.93	1.19	3.60	4.23	3.01
	TDGSF	2.47	8.07	9.33	6.62	2.35	7.11	8.35	5.93	1.19	3.60	4.23	3.01
	TDFP	2.47	8.07	9.33	6.62	2.35	7.11	8.35	5.93	1.19	3.60	4.23	3.01
	OLS	2.41	7.97	9.12	6.50	2.28	6.95	8.05	5.76	1.15	3.52	4.08	2.92
	WLSv	2.34	7.82	8.75	6.31	2.19	6.66	7.49	5.45	1.11	3.37	3.80	2.76
	WLSs	2.32	7.79	8.67	6.26	2.16	6.59	7.36	5.37	1.09	3.34	3.73	2.72
	MintT (Sample)	**2.22**	**7.71**	8.56	**6.16**	**2.06**	6.49	7.22	**5.26**	**1.05**	3.29	3.65	**2.66**
	MinT (Shrink)	2.25	7.75	8.63	6.21	2.10	6.56	7.28	5.31	1.07	3.32	3.69	2.69
		Hierarchical level 1 : Electrical subsystems										
Method	BU	1.04	2.99	**3.18**	**2.40**	0.97	**2.52**	**2.66**	**2.05**	1.04	**2.74**	**2.97**	**2.25**
	TDGSA	5.48	6.88	7.06	6.47	5.46	6.72	6.88	6.35	6.30	7.44	7.45	7.06
	TDGSF	5.51	6.91	7.08	6.50	5.50	6.75	6.90	6.38	6.28	7.40	7.41	7.03
	TDFP	1.16	**2.95**	3.99	2.70	1.13	2.75	3.71	2.53	1.20	2.93	3.95	2.70
	OLS	1.08	3.20	3.65	2.64	1.04	2.89	3.27	2.40	1.12	3.14	3.64	2.64
	WLSv	1.05	3.10	3.44	2.53	1.00	2.73	3.01	2.24	1.06	2.91	3.26	2.41
	WLSs	1.05	3.10	3.44	2.53	1.00	2.73	3.00	2.24	1.07	2.95	3.33	2.45
	MintT (Sample)	**1.01**	3.07	3.41	2.50	**0.96**	2.70	2.95	2.20	**1.00**	2.87	3.18	2.35
	MinT (Shrink)	1.02	3.08	3.43	2.51	0.97	2.72	2.97	2.22	1.02	2.90	3.22	2.38
		Hierarchical level 2 : Energy sources										
Method	BU	0.89	2.32	**2.38**	1.87	**0.83**	1.98	**2.04**	**1.62**	1.22	2.70	2.95	2.29
	TDGSA	3.39	4.70	4.93	4.34	3.37	4.57	4.79	4.24	14.02	14.92	15.04	14.66
	TDGSF	3.36	4.68	4.90	4.32	3.35	4.55	4.76	4.22	14.79	15.66	15.77	15.41
	TDFP	0.98	**2.30**	2.97	2.08	0.95	2.15	2.77	1.95	1.62	3.57	4.91	3.37
	OLS	0.90	2.37	2.49	1.92	0.85	2.07	2.15	1.69	1.85	5.25	7.12	4.74
	WLSv	0.91	2.42	2.63	1.98	0.86	2.18	2.33	1.79	1.25	2.88	3.22	2.45
	WLSs	0.90	2.35	2.44	1.89	0.84	2.04	2.09	1.66	1.60	4.22	5.37	3.73
	MintT (Sample)	**0.88**	2.40	2.62	1.97	0.84	2.17	2.31	1.77	1.23	2.87	3.24	2.45
	MinT (Shrink)	0.89	2.41	2.64	1.98	0.85	2.18	2.33	1.79	1.24	2.89	3.26	2.46

Figure A3. Hierarchical forecasting for electricity generation based on the ARIMA procedure (RMSE, MAE, MASE). (Note: The performance was indicated into a color scale, where green means better values for calculated accuracy, and red means worse accuracy. The intermediate values are colored yellow.).

		Predictive model: Error, trend, seasonality (ETS)											
		RMSE				MAE				MASE			
		Forecast horizon (h) - mean of the interval											
		[1-3]	[4-6]	[7-9]	Mean	[1-3]	[4-6]	[7-9]	Mean	[1-3]	[4-6]	[7-9]	Mean
		Hierarchical level 0 : Total - Brazil											
Method	BU	3.02	9.14	11.26	7.81	2.98	8.62	10.57	7.39	1.51	4.37	5.35	3.74
	TDGSA	2.65	8.62	10.51	7.26	2.59	7.98	9.76	6.78	1.31	4.04	4.94	3.43
	TDGSF	2.65	8.62	10.51	7.26	2.59	7.98	9.76	6.78	1.31	4.04	4.94	3.43
	TDFP	2.65	8.62	10.51	7.26	2.59	7.98	9.76	6.78	1.31	4.04	4.94	3.43
	OLS	2.67	8.64	10.54	7.28	2.61	8.01	9.78	6.80	1.32	4.06	4.96	3.44
	WLSv	2.77	8.78	10.75	7.43	2.72	8.19	10.01	6.97	1.38	4.15	5.07	3.53
	WLSs	2.78	8.81	10.78	7.46	2.73	8.22	10.05	7.00	1.38	4.16	5.09	3.54
	MintT (Sample)	**2.61**	**8.56**	**10.43**	**7.20**	**2.55**	**7.91**	**9.66**	**6.71**	**1.29**	**4.01**	**4.89**	**3.40**
	MinT (Shrink)	2.62	8.58	10.45	7.21	2.56	7.92	9.68	6.72	1.30	4.01	4.90	3.40
		Hierarchical level 1 : Electrical subsystems											
Method	BU	1.30	3.66	4.48	3.15	1.29	3.47	4.18	2.98	1.44	3.84	4.66	3.31
	TDGSA	5.54	7.09	7.45	6.70	5.54	6.98	7.30	6.61	6.37	7.65	7.81	7.28
	TDGSF	5.58	7.12	7.48	6.73	5.57	7.00	7.32	6.63	6.34	7.62	7.77	7.24
	TDFP	**1.13**	**2.95**	**4.07**	**2.72**	**1.11**	**2.77**	**3.79**	**2.56**	**1.22**	**3.03**	**4.14**	**2.80**
	OLS	1.15	3.42	4.15	2.91	1.13	3.21	3.84	2.73	1.23	3.50	4.22	2.99
	WLSv	1.20	3.50	4.25	2.98	1.18	3.29	3.95	2.81	1.31	3.61	4.37	3.10
	WLSs	1.20	3.50	4.26	2.98	1.18	3.29	3.95	2.81	1.30	3.60	4.36	3.08
	MintT (Sample)	1.20	3.50	4.25	2.98	1.17	3.28	3.93	2.79	1.25	3.52	4.26	3.01
	MinT (Shrink)	1.18	3.47	4.22	2.96	1.16	3.26	3.91	2.77	1.25	3.51	4.25	3.00
		Hierarchical level 2 : Energy sources											
Method	BU	1.15	2.93	3.53	2.54	1.14	2.83	3.34	2.44	1.64	3.87	4.68	3.40
	TDGSA	3.43	4.85	5.21	4.50	3.42	4.75	5.08	4.42	14.21	15.56	16.08	15.28
	TDGSF	3.41	4.83	5.18	4.47	3.40	4.73	5.05	4.39	14.97	16.29	16.78	16.01
	TDFP	1.05	**2.47**	3.28	**2.27**	1.03	**2.36**	3.11	**2.17**	1.87	3.88	5.06	3.60
	OLS	1.13	2.89	3.48	2.50	1.11	2.79	3.29	2.40	2.64	5.55	6.80	5.00
	WLSv	1.07	2.81	3.36	2.41	1.05	2.69	3.17	2.30	1.55	3.74	4.50	3.26
	WLSs	1.13	2.91	3.50	2.51	1.12	2.80	3.31	2.41	2.31	4.99	6.09	4.46
	MintT (Sample)	**1.01**	2.71	**3.23**	2.32	**0.99**	2.59	**3.03**	2.20	**1.44**	**3.38**	**3.99**	**2.93**
	MinT (Shrink)	1.03	2.75	3.28	2.35	1.01	2.63	3.08	2.24	1.43	3.49	4.19	3.03

Figure A4. Hierarchical forecasting for electricity generation based on the ETS procedure (RMSE, MAE, MASE). (Note: The performance was indicated into a color scale, where green means better values for calculated accuracy, and red means worse accuracy. The intermediate values are colored yellow.).

References

1. Medojevic, M.; Medic, N.; Marjanovic, U.; Lalic, B.; Majstorovic, V. Exploring the impact of industry 4.0 concepts on energy and environmental management systems: Evidence from Serbian manufacturing companies. In Proceedings of the IFIP International Conference on Advances in Production Management Systems, Austin, TX, USA, 1–5 September 2019; pp. 355–362.
2. Alcácer, V.; Cruz-Machado, V. Scanning the industry 4.0: A literature review on technologies for manufacturing systems. *Eng. Sci. Technol. Int. J.* **2019**, *22*, 899–919. [CrossRef]
3. Bourdeau, M.; Zhai, X.Q.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustain. Cities Soc.* **2019**, *48*, 101533. [CrossRef]
4. Hammad, M.A.; Jereb, B.; Rosi, B.; Dragan, D. Methods and models for electric load forecasting: A comprehensive review. *Logist. Sustain. Transp.* **2020**, *11*, 51–76. [CrossRef]

5. Runge, J.; Zmeureanu, R. Forecasting energy use in buildings using artificial neural networks: A review. *Energies* **2019**, *12*, 3254. [CrossRef]

6. Choi, Y.B. *Paradigms and Conventions: Uncertainty, Decision Making, and Entrepreneurship*; University of Michigan Press: Ann Arbor, MI, USA, 1993.

7. Jiang, W.; Yan, Z.; Feng, D.H.; Hu, Z. Wind speed forecasting using autoregressive moving average/generalized autoregressive conditional heteroscedasticity model. *Eur. Trans. Electr. Power* **2012**, *22*, 662–673. [CrossRef]

8. Hao, C.H.E.N. A new method of load forecasting based on generalized autoregressive conditional heteroscedasticity model. *Autom. Electr. Power Syst.* **2007**, *15*, 012.

9. Stefenon, S.F.; Ribeiro, M.H.D.M.; Nied, A.; Mariani, V.C.; dos Santos Coelho, L.; da Rocha, D.F.M.; Grebogif, R.B.; de Barros Ruano, A.E. Wavelet group method of data handling for fault prediction in electrical power insulators. *Int. J. Electr. Power Energy Syst.* **2020**, *123*, 106269. [CrossRef]

10. Frizzo Stefenon, S.; Silva, M.C.; Bertol, D.W.; Meyer, L.H.; Nied, A. Fault diagnosis of insulators from ultrasound detection using neural networks. *J. Intell. Fuzzy Syst.* **2019**, *37*, 6655–6664. [CrossRef]

11. Gupta, S.; Srinivasan, D.; Reindl, T. Forecasting solar and wind data using dynamic neural network architectures for a micro-grid ensemble. In Proceedings of the 2013 IEEE Computational Intelligence Applications in Smart Grid (CIASG), Singapore, 16–19 April 2013; IEEE: New York, NY, USA, 2013; pp. 87–92.

12. Frizzo Stefenon, S.; Zanetti Freire, R.; dos Santos Coelho, L.; Meyer, L.H.; Bartnik Grebogi, R.; Gouvêa Buratto, W.; Nied, A. Electrical insulator fault forecasting based on a wavelet neuro-fuzzy system. *Energies* **2020**, *13*, 484. [CrossRef]

13. Moghaddam, A.A.; Seifi, A.R. Study of forecasting renewable energies in smart grids using linear predictive filters and neural networks. *IET Renew. Power Gener.* **2011**, *5*, 470–480. [CrossRef]

14. Barbounis, T.G.; Theocharis, J.B. A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation. *Neurocomputing* **2007**, *70*, 1525–1542. [CrossRef]

15. Xia, J.; Zhao, P.; Dai, Y. Neuro-fuzzy networks for short-term wind power forecasting. In Proceedings of the 2010 International Conference on Power System Technology, Hangzhou, China, 24–28 October 2010; IEEE: New York, NY, USA, 2010; pp. 1–5.

16. Dawan, P.; Sriprapha, K.; Kittisontirak, S.; Boonraksa, T.; Junhuathon, N.; Titiroongruang, W.; Niemcharoen, S. Comparison of power output forecasting on the photovoltaic system using adaptive neuro-fuzzy inference systems and particle swarm optimization-artificial neural network model. *Energies* **2020**, *13*, 351. [CrossRef]

17. Hyndman, R.J.; Kandahar, Y. Automatic time series forecasting: The forecast package for R. *J. Stat. Softw.* **2008**, *26*, 1–22.

18. Athanasopoulos, G.; Ahmed, R.A.; Hyndman, R.J. Hierarchical forecasts for Australian domestic tourism. *Int. J. Forecast.* **2009**, *25*, 146–166. [CrossRef]

19. Hyndman, R.J.; Ahmed, R.A.; Athanasopoulos, G.; Shang, H.L. Optimal combination forecasts for hierarchical time series. *Comput. Stat. Data Anal.* **2011**, *55*, 2579–2589. [CrossRef]

20. Almeida, V.; Ribeiro, R.; Gama, J. Hierarchical time series forecast in electrical grids. In *Information Science and Applications (ICISA)*; Springer: Singapore, 2016; pp. 995–1005.

21. Panamtash, H.; Zhou, Q. Coherent probabilistic solar power forecasting. In Proceedings of the 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Boise, ID, USA, 24–28 June 2018; IEEE: New York, NY, USA, 2018; pp. 1–6.

22. Abouarghoub, W.; Nomikos, N.K.; Petropoulos, F. On reconciling macro and micro energy transport forecasts for strategic decision making in the tanker industry. *Transp. Res. Part E Logist. Transp. Rev.* **2018**, *113*, 225–238. [CrossRef]

23. Auder, B.; Cugliari, J.; Goude, Y.; Poggi, J.M. Scalable clustering of individual electrical curves for profiling and bottom-up forecasting. *Energies* **2018**, *11*, 1893. [CrossRef]

24. Silva, F.L.; Souza, R.C.; Oliveira, F.L.C.; Lourenco, P.M.; Calili, R.F. A bottom-up methodology for long term electricity consumption forecasting of an industrial sector-Application to pulp and paper sector in Brazil. *Energy* **2018**, *144*, 1107–1118. [CrossRef]

25. Ghedamsi, R.; Settou, N.; Gouareh, A.; Khamouli, A.; Saifi, N.; Recioui, B.; Dokkar, B. Modeling and forecasting energy consumption for residential buildings in Algeria using bottom-up approach. *Energy Build.* **2016**, *121*, 309–317. [CrossRef]

26.   Kosiorowski, D.; Mielczarek, D.; Rydlewski, J. Forecasting of a hierarchical functional time series on example of macromodel for day and night air pollution in silesia region: A critical overview. *arXiv* **2017**, arXiv:1712.03797.

27.   Wickramasuriya, S.L.; Athanasopoulos, G.; Hyndman, R.J. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *J. Am. Stat. Assoc.* **2019**, *114*, 804–819. [CrossRef]

28.   National System Operator. Operation History (Report of Power Generation). 2020. Available online: http://www.ons.org.br/paginas/resultados-da-operacao/historico-da-operacao (accessed on 15 May 2020).

29.   R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020; Available online: https://www.R-project.org/ (accessed on 15 May 2020).

30.   Orcutt, G.H.; Watts, H.W.; Edwards, J.B. Data aggregation and information loss. *Am. Econ. Rev.* **1968**, *58*, 773–787.

31.   Gross, C.W.; Sohl, J.E. Disaggregation methods to expedite product line forecasting. *J. Forecast.* **1990**, *9*, 233–254. [CrossRef]

32.   Oliveira, J.M.; Ramos, P. Assessing the performance of hierarchical forecasting methods on the retail sector. *Entropy* **2019**, *21*, 436. [CrossRef]

33.   Yang, D.; Kleissl, J.; Gueymard, C.A.; Pedro, H.T.; Coimbra, C.F. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Sol. Energy* **2018**, *168*, 60–101. [CrossRef]

34.   Panigrahi, S.; Behera, H.S. A hybrid ETS–ANN model for time series forecasting. *Eng. Appl. Artif. Intell.* **2017**, *66*, 49–59. [CrossRef]

35.   Dong, Z.; Yang, D.; Reindl, T.; Walsh, W.M. Short-term solar irradiance forecasting using exponential smoothing state space model. *Energy* **2013**, *55*, 1104–1113. [CrossRef]

36.   Box, G.E. Jenkins. In *Time Series Analysis: Forecasting and Control*; Holden-Day Inc.: New York, NY, USA, 1976.

37.   Pegels, C.C. Exponential forecasting: Some new variations. *Manag. Sci.* **1969**, 311–315.

38.   Gardner, E.S., Jr. Exponential smoothing: The state of the art. *J. Forecast.* **1985**, *4*, 1–28. [CrossRef]

39.   Hyndman, R.J.; Koehler, A.B.; Snyder, R.D.; Grose, S. A state space framework for automatic forecasting using exponential smoothing methods. *Int. J. Forecast.* **2002**, *18*, 439–454. [CrossRef]

40.   Taylor, J.W. Exponential smoothing with a damped multiplicative trend. *Int. J. Forecast.* **2003**, *19*, 715–725. [CrossRef]

41.   Liu, Z.; Yan, Y.; Yang, J.; Hauskrecht, M. Missing value estimation for hierarchical time series: A study of hierarchical Web traffic. In Proceedings of the 2015 IEEE International Conference on Data Mining, Atlantic City, NJ, USA, 14–17 November 2015; IEEE: New York, NY, USA, 2015; pp. 895–900.

42.   Weiss, C. Essays in Hierarchical Time Series Forecasting and Forecast Combination. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 2018.

43.   Hong, T.; Xie, J.; Black, J. Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting. *Int. J. Forecast.* **2019**, *35*, 1389–1399. [CrossRef]