

Forecasting hierarchical time series in power generation: modeling and analysis in Brazil

Tiago Silveira Gontijo

Economista, mestre em Engenharia de Produção pela UFMG e doutorando em Engenharia de Produção pela UFMG. Departamento de Engenharia de Produção da Universidade Federal de Minas Gerais - UFMG.
E-mail: tsgontijo@hotmail.com

Marcelo Azevedo Costa

Engenheiro Eletricista pela UFMG, doutor em Engenharia Elétrica pela UFMG, pós-doutorado pela Harvard Medical School e Harvard Pilgrim Health Care (EUA) e pós-doutorado pela Universidade de Linköping (Suécia). Professor do Departamento de Engenharia de Produção/Universidade Federal de Minas Gerais - UFMG.
E-mail: macosta@ufmg.br

RESUMO

O estudo de séries temporais hierárquicas tem despertado a atenção da literatura. Especialmente no setor elétrico, diversas são as aplicações onde as informações podem ser organizadas em uma estrutura hierárquica. Este trabalho analisou a geração horária de energia no Brasil (2018-2020), que foi agrupada em função de cada um dos subsistemas elétricos e de suas respectivas fontes geradoras de energia. O objetivo geral foi calcular a acurácia das principais medidas para agregar e desagregar as previsões dos modelos ARIMA e ETS. Especificamente, analisou-se as abordagens hierárquicas: (i) Bottom-Up (BU); (ii) Top-Down (TD) e (iii) Reconciliação Ótima. Os modelos de Reconciliação Ótima apresentaram o melhor desempenho médio, levando-se em consideração as primeiras janelas preditivas. Verificou-se ainda que as previsões de energia no subsistema Sul indicaram maior grau de imprecisão frente aos demais, o que sinaliza a necessidade de modelos individualizados para este subsistema.

PALAVRAS CHAVE. Geração de Energia, Subsistemas Elétricos, Séries Temporais

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 ARIMA and 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 [Medojevic et al., 2019]. In high-level data analysis, forecasting models allow the extraction of behavior patterns, as well as the prediction of future scenarios for the collected data set [Alcácer & Cruz-Machado, 2019].

There are several forecasting applications relevant to power generation, the object of the present study, including: (i) classical time series models [Chen et al., 2013], (ii) pre-processing techniques [Malvoni, De Giorgi & Congedo, 2017], (iii) machine learning approaches [Sharifzadeh, Sikinioti-Lock & Shah, 2019], among others. Additionally, an alternative class known as hierarchical forecasting [Athanasopoulos, Ahmed & Hyndman, 2009] 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 application of hierarchical forecasting models in the energy sector. Some examples of applications are: electrical grids [Almeida, Ribeiro & Gama, 2016], forecasting models for air pollution [Kosiorowski, Mielczarek & Rydlewski, 2017], solar power generation [Panamtaash & Zhou, 2018], energy transport [Abouarghoub, Nomikos & Petropoulos, 2018], among others.

The papers identified above have calibrated the forecasts using only the Bottom-up, Top-down and OLS assumptions [Hyndman et al., 2011]. Thus, the following research question is formulated: how is it possible to make hierarchical predictions using advanced linear regression models with regularization? We can obtain more reliable forecasts if we rewrite 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 [Wickramasuriya, Athanasopoulos & Hyndman, 2019].

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 predictive models ARIMA and ETS were used to test the performance of these reconciliation methods.

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. Material 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 [ONS, 2020], 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 17521 hours.

Based on Hyndman et al. [2011], 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 $k = 2$).

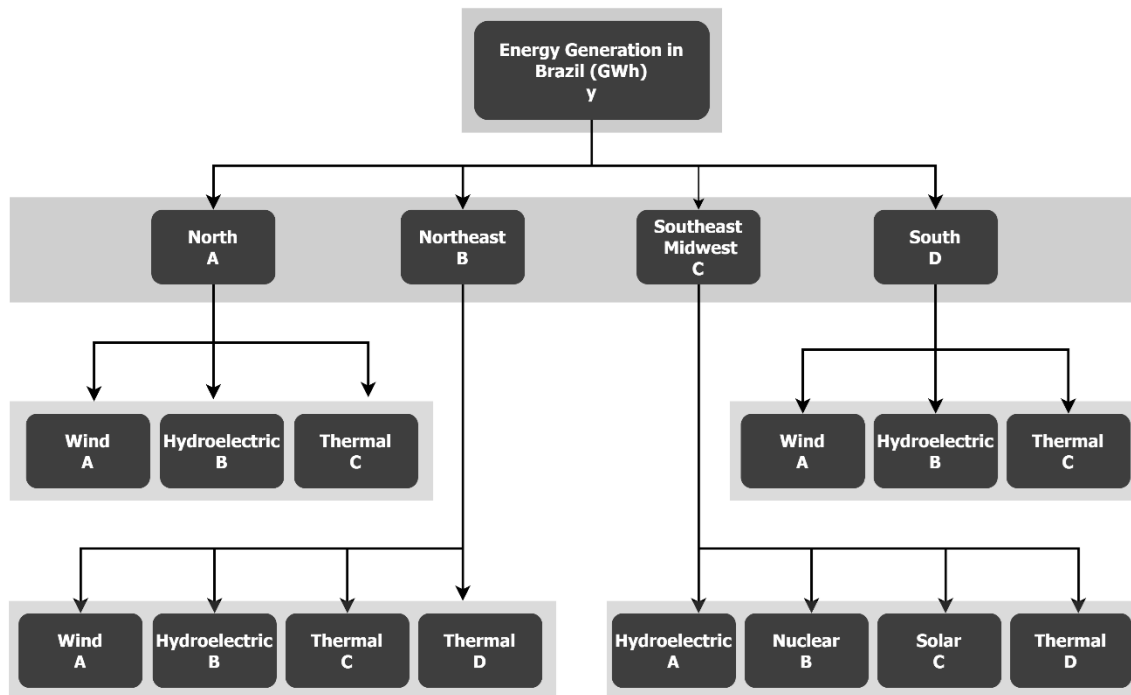


Fig. 1. Hierarchical aggregation structure for the energy generation in Brazil

Table 1 shows the proportions of electricity generation in Brazil, 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.

Table 1. Proportions of electricity generation in Brazil

Subsystem / Source		Wind	Hydro	Thermal	Solar	Nuclear	Total (GWh - Subsystem)	%
North	(A)	2688	125182	31489	0	0	159359	14.3%
Northeast	(B)	85377	37705	36699	4626	0	164407	14.7%
Southeast / Midwest	(C)	0	518714	73555	2437	31805	626511	56.1%
South	(D)	11326	135914	19472	0	0	166712	14.9%
Total (GWh - Source)		99391	817516	161215	7063	31805	1116989	100%
%		8.9%	73.2%	14.4%	0.6%	2.8%	100%	-

In the next section, we present the hierarchical reconciliation models used in the present paper.

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 (Orcutt, Watts & Edwards, 1969). 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 [Wickramasuriya, Athanasopoulos & Hyndman, 2019]. 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:

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

According to Hyndman et al. [2011], it is possible to arrange the equations expressed in (2) into an algebra notation. Below is a complete notation for this problem:

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

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:

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

where $\tilde{y}_t$ 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. [Hyndman & Khandakar, 2007].

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 [Hyndman & Khandakar, 2007]. Let $p_1, \dots, p_m$ 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 $p_1, \dots, p_{14}$:

$$\begin{aligned}\tilde{y}_{AA,t} &= p_1 \hat{y}_t, \tilde{y}_{AB,t} = p_2 \hat{y}_t, \tilde{y}_{AC,t} = p_3 \hat{y}_t \\ \tilde{y}_{BA,t} &= p_4 \hat{y}_t, \tilde{y}_{BB,t} = p_5 \hat{y}_t, \tilde{y}_{BC,t} = p_6 \hat{y}_t, \tilde{y}_{BD,t} = p_7 \hat{y}_t \\ \tilde{y}_{CA,t} &= p_8 \hat{y}_t, \tilde{y}_{CB,t} = p_9 \hat{y}_t, \tilde{y}_{CC,t} = p_{10} \hat{y}_t, \tilde{y}_{CD,t} = p_{11} \hat{y}_t \\ \tilde{y}_{DA,t} &= p_{12} \hat{y}_t, \tilde{y}_{DB,t} = p_{13} \hat{y}_t, \tilde{y}_{DC,t} = p_{14} \hat{y}_t\end{aligned}\quad (5)$$

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

$$\begin{aligned}\tilde{b}_t &= p_j \hat{y}_t \\ \tilde{y}_t &= S p_j \hat{y}_t\end{aligned}\quad (6)$$

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 Eq. 7 can be obtained from Gross & Sohl [1990] and Athanasopoulos, Ahmed & Hyndman [2009].

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	
$p_j = \frac{1}{T} \sum_{t=1}^T \frac{y_{j,t}}{y_t}$ for $j = 1, \dots, m$. Each proportion p_j reflects the average of the historical proportions of the bottom-level series $y_{j,t}$, t over the period $t = 1, \dots, T$ relative to the total aggregate y_t.	$p_j = \sum_{t=1}^T \frac{y_{j,t}}{T} / \sum_{t=1}^T \frac{y_t}{T}$ for $j = 1, \dots, m$. 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.	$p_j = \prod_{l=0}^{K-1} \frac{\hat{y}_{j,t}^{(l)}}{\hat{S}_{j,t}^{(l+1)}}$ where $j = 1, \dots, m$, $\hat{y}_{j,t}^{(l)}$ 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.	(7)

2.3. The optimal reconciliation approaches

The optimal reconciliation approach proposed by Hyndman, Ahmed, et al. [2011] 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 Oliveira & Ramos [2019], we can write the base prediction as:

$$\hat{y}_{t+h|t} = S \beta_{t+h|t} + \varepsilon_h \quad (8)$$

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 Σ_h . If Σ_h were known, the estimator of $\beta_{t+h|t}$ would lead to the following weighted least squares, producing reconciled forecasts, as follows:

$$\tilde{y}_{t+h|t} = S \hat{\beta}_{t+h|t} = S(S' \Sigma_h^{-1} S)^{-1} S' \Sigma_h^{-1} \hat{y}_{t+h|t} = S P \hat{y}_{t+h|t} \quad (9)$$

where $P = (S' \Sigma_h^{-1} S)^{-1} S' \Sigma_h^{-1} S$. If the base forecasts $\hat{y}_{t+h|t}$ are unbiased, then the reconciled forecasts $\tilde{y}_{t+h|t}$ will be unbiased, provided that $SPS = S$ [Hyndman, Ahmed, et al., 2011]. 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, Wickramasuriya et al. [2019] proved that, in general, Σ_h is not known and not identifiable. The covariance matrix of the h -step-ahead reconciled forecast errors is given by the following expression:

$$Var_{(y_{t+h} - \tilde{y}_{t+h|t})} = SPW_h P' S' \quad (10)$$

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 $Var_{(y_{t+h} - \tilde{y}_{t+h|t})}$. Finally, Wickramasuriya et al. [2019] demonstrated that the optimal reconciliation matrix P that minimizes the trace of $SPW_h P' S'$, such that $SPS = S$, and the optimal reconciled forecasts, respectively, are given by:

$$P = (S' W_h^{-1} S)^{-1} S' W_h^{-1} \\ \tilde{y}_{t+h|t} = S(S' W_h^{-1} S)^{-1} S' W_h^{-1} \hat{y}_{t+h|t} \quad (11)$$

which is introduced as the MinT (Minimum Trace) estimator. The next step consists of estimating W_h , a matrix of order n . Wickramasuriya et al. [2019] proposed the following procedures (Table 3) to obtain the matrix:

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

OLS	$\mathbf{W}_h = k_h \mathbf{I}, \forall h$ where $k_h > 0$. This is the most simplifying premise, and collapses the MinT estimator to the OLS estimator, proposed by Hyndman et al. [2011]. This is optimal when the base forecast errors are uncorrelated and equivariant.	(12)
WLSv	$\mathbf{W}_h = k_h \text{diag}(\widehat{\mathbf{W}}_1), \forall h$ where $k_h > 0$ and: $\widehat{\mathbf{W}} = \frac{1}{T} \sum_{t=1}^T \hat{e}_t(1) \hat{e}_t(1)'$ is the unbiased sample covariance estimator of the insample one-step-ahead base forecast errors. In this case, we can describe MinT as a WLS estimator applying variance scaling [Athanasopoulos et al., 2017].	(13)
WLSs	$\mathbf{W}_h = k_h \mathbf{\Lambda}, \forall h$ where $k_h > 0$ and $\mathbf{\Lambda} = \text{diag}(\mathbf{S1})$ with $\mathbf{1}$ 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 $\mathbf{\Lambda}$ matrix receives the number of forecast error variances contributing to that aggregation level [Athanasopoulos et al., 2017]. This estimator depends only on the grouping structure of the hierarchy.	(14)
MinT (Sample)	$\mathbf{W}_h = k_w \widehat{\mathbf{W}}_1, \forall h$ where $k_h > 0$, the unrestricted sample covariance estimator for $h = 1$ [Wickramasuriya et al., 2019]. In the results section, we denote this as MinT (Sample).	(15)
MinT (Shrink)	$\mathbf{W}_h = k_w \widehat{\mathbf{W}}_{1,D}^*, \forall h; k_h > 0; \mathbf{W}_{1,D}^* = \lambda_D \widehat{\mathbf{W}}_{1,D} + (1 - \lambda_D) \widehat{\mathbf{W}}_1$ is a shrinkage estimator with diagonal target, $\widehat{\mathbf{W}}_{1,D}$, which is a diagonal matrix comprising the diagonal entries of $\widehat{\mathbf{W}}_1$, and λ_D is the shrinkage intensity parameter. Thus, off-diagonal elements of $\widehat{\mathbf{W}}_1$ are shrunk toward zero and diagonal elements (variances) remain unchanged [Wickramasuriya et al., 2019]. Schäfer and Strimmer (2005) proposed a scale and location invariant shrinkage estimator by parameterizing the shrinkage in terms of variances and correlations: $\hat{\lambda}_D = \frac{\sum_{i \neq j} \widehat{\text{var}}(\hat{r}_{ij})}{\sum_{i \neq j} \hat{r}_{ij}^2}$ where $\hat{r}_{ij}$ is the ij th element of $\widehat{\mathbf{R}}_1$, the 1-step-ahead sample correlation matrix to shrink it toward an identity matrix.	(16)

Source: adapted by authors from: Wickramasuriya et al. [2019].

2.4. Evaluating forecast accuracy

According to Almeida, Ribeiro & Gama [2016], the literature proposes a family of techniques, such as mean absolute error (MAE), mean square error (MSE) or mean absolute percentage error (MAPE) to evaluate the performance of point prediction methods.

Specifically, in studies of hierarchical time series, the MAPE indicator appears the most frequently in the literature [Liu et al., 2015; Weis, 2018; Hong, Xie & Black, year?]. MAPE was also the selected metric for the present paper, defined as follows:

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (17)$$

The values of the MAPE statistic were obtained using a weighted average, with proportions from Table 1. We chose two forecasting models, ARIMA [Box & Jenkins, 1970] and ETS [Hyndman & Khandakar, 2007], to demonstrate the application of the methods analyzed in 2.1 to 2.3.

3. Results and Discussion

Figure 2, below, shows the predictive result obtained, using the ARIMA model, considering a predictive window of nine hours ($h = 1, \dots, 9$). 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 the Table presents the average performance for each forecast horizon (h) for each hierarchical approach.

As pointed by Wickramasuriya, Athanasopoulos & Hyndman [2019], 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.

Predictive model: Autoregressive integrated moving average (ARIMA)											
MAPE		Forecast horizon (h)									Mean
		1	2	3	4	5	6	7	8	9	
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

Fig. 2. Hierarchical forecasting for electricity generation based on the ARIMA procedure

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, OLS and WLS.

Predictive model: Error, trend, seasonality (ETS)											
MAPE		Forecast horizon (h)									Mean
		1	2	3	4	5	6	7	8	9	
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

Fig. 3. Hierarchical forecasting for electricity generation based on the ETS procedure

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 9-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 3 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 there exist unique climatic conditions southern Brazil.

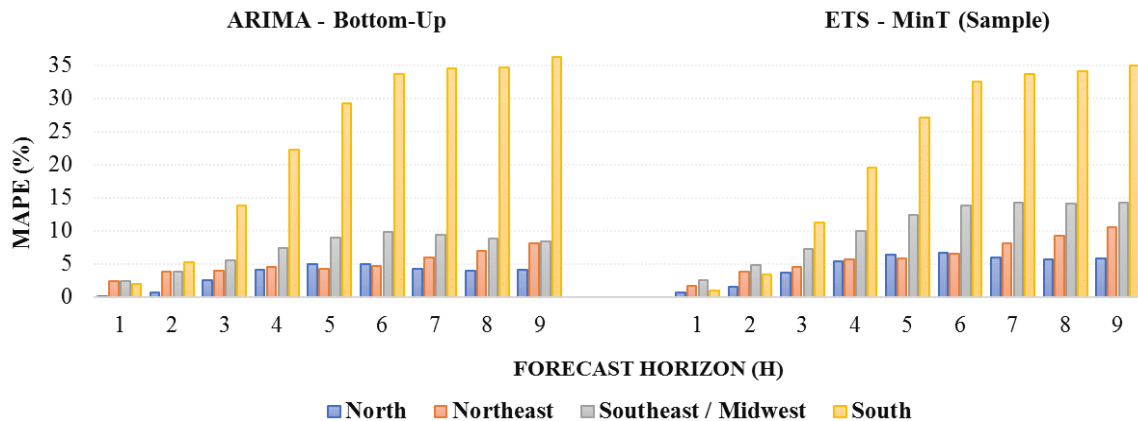


Fig. 4. Hierarchical forecasting for electricity generation based on the ETS procedure

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. 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 on spot market sales prices.

Acknowledgments

Authors would like to thank the National Council for Scientific and Technological Development (CNPq) and *Companhia Energética Integrada* (CEI) for its financial support on the project (Grant number 141740/2019-1).

References

- Abouarghoub, W., Nomikos, N. K., & Petropoulos, F. (2018). On reconciling macro and micro energy transport forecasts for strategic decision making in the tanker industry. *Transportation Research Part E: Logistics and Transportation Review*, 113, 225-238.
- Alcácer, V., & Cruz-Machado, V. (2019). Scanning the industry 4.0: A literature review on technologies for manufacturing systems. *Engineering Science and Technology, an International Journal*.

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

Athanasopoulos, G., Ahmed, R. A., & Hyndman, R. J. (2009). Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting*, 25(1), 146-166.

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & Petropoulos, F. (2017). Forecasting with temporal hierarchies. *European Journal of Operational Research*, 262(1), 60-74.

Box, G. E., & Jenkins, G. M. (1970). Time series analysis: Forecasting and control Holden-Day. *San Francisco*, 498.

Chen, H., Wan, Q., Li, F., & Wang, Y. (2013, July). GARCH in mean type models for wind power forecasting. In *2013 IEEE Power & Energy Society General Meeting* (pp. 1-5). IEEE.

Gross, C. W., & Sohl, J. E. (1990). Disaggregation methods to expedite product line forecasting. *Journal of forecasting*, 9(3), 233-254.

Hong, T., Xie, J., & Black, J. (2019). Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting. *International Journal of Forecasting*, 35(4), 1389-1399.

Hyndman, R. J., & Khandakar, Y. (2007). *Automatic time series for forecasting: the forecast package for R* (No. 6/07). Clayton VIC, Australia: Monash University, Department of Econometrics and Business Statistics.

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational statistics & data analysis*, 55(9), 2579-2589.

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

Liu, Z., Yan, Y., Yang, J., & Hauskrecht, M. (2015, November). Missing value estimation for hierarchical time series: A study of hierarchical Web traffic. In *2015 IEEE International Conference on Data Mining* (pp. 895-900). IEEE.

Malvoni, M., De Giorgi, M. G., & Congedo, P. M. (2017). Forecasting of PV Power Generation using weather input data-preprocessing techniques. *Energy Procedia*, 126, 651-658.

Medojevic, M., Medic, N., Marjanovic, U., Lalic, B., & Majstorovic, V. (2019, September). Exploring the Impact of Industry 4.0 Concepts on Energy and Environmental Management Systems: Evidence from Serbian Manufacturing Companies. In *IFIP International Conference on Advances in Production Management Systems* (pp. 355-362). Springer, Cham.

Oliveira, J. M., & Ramos, P. (2019). Assessing the Performance of Hierarchical Forecasting Methods on the Retail Sector. *Entropy*, 21(4), 436.

Orcutt, G. H., Watts, H. W., & Edwards, J. B. (1968). Data aggregation and information loss. *The American Economic Review*, 58(4), 773-787.

Panamtash, H., & Zhou, Q. (2018, June). Coherent probabilistic solar power forecasting. In *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)* (pp. 1-6). IEEE.

Sharifzadeh, M., Sikinioti-Lock, A., & Shah, N. (2019). Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression. *Renewable and Sustainable Energy Reviews*, 108, 513-538.

Wickramasuriya, S. L., Athanasopoulos, G., & Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526), 804-819.