

Combined Forecast Model for Wind Generation in Brazilian Monthly Dispatch Scheduling

G. Godinho, P. Maçaira, and R. Souza

Abstract—The fast growing of wind generation in Brazil has brought benefits to the country's energy supply, especially under critical hydrological conditions, as observed between 2014 and 2017. However, deviations in wind generation predictions may lead to the depletion of reservoirs and expensive thermal generation. In order to improve the accuracy of wind generation forecasts for the monthly dispatch scheduling, four different models based on exponential smoothing (Holt-Winters) and ARIMA (Box & Jenkins) approach were proposed for Rio do Fogo wind farm, located in the Northeast of Brazil. The results show that the variance-weighted combination of an ARIMA (1,0,0) x (1,1,1)₁₂ and a multiplicative Holt-Winters model provides the best fit for the studied series. In addition, the mean absolute percentage error found is 14% lower than the one provided by the current forecast model (moving averages). Thus, it is possible to apply the ARIMA/Holt-Winters variance-weighted combination model with greater accuracy for forecasting the monthly generation of Rio do Fogo wind farm, during the dispatch planning carried out by the Brazilian independent system operator (ONS for short).

Index Terms—Seasonal ARIMA, Combined Forecast, Wind Power, Holt-Winters, Energy Planning, Dispatch Scheduling.

I. INTRODUÇÃO

NO passado recente, a geração eólica tem se mostrado essencial para o atendimento à carga do Sistema Interligado Nacional (SIN), principalmente sob condições hidrológicas críticas, conforme observado entre os anos de 2014 e 2017. No Nordeste, por exemplo, a geração eólica foi responsável pelo atendimento de 83% da carga deste subsistema em base horária (6.689 MWh às 8:00 de 10/09/2017), e 60% em base mensal (5.750 MWmed em setembro/2017) [1].

A expressiva geração de energia eólica no Nordeste tem possibilitado o replecionamento de importantes reservatórios da região. Além dos benefícios citados, a rápida expansão dessa fonte, crescendo de 2,1 GW em 2013 para 12,3 GW instalados em 2017, também tem trazido desafios para a operação em tempo real, dentre os quais pode-se destacar: alta variabilidade em intervalos curtos de tempo (rampas e rajadas) e difícil previsibilidade, não só na etapa da programação diária (D-1), como durante a operação em tempo real.

O Operador Nacional do Sistema Elétrico (ONS) é responsável pela realização dos estudos de otimização e simulação da operação do SIN no Programa Mensal da Operação (PMO). Estes estudos estabelecem as políticas de

geração térmica e intercâmbios inter-regionais ótimas para os períodos analisados, fornecendo metas e diretrizes a serem seguidas pela Programação Diária da Operação Eletroenergética e pela Operação em Tempo Real [2].

Atualmente, a geração eólica mensal prevista para o PMO é realizada de acordo com as diretrizes da Resolução Normativa ANEEL nº 440/2011 [3]. Segundo a Resolução, a previsão de geração eólica das usinas em operação corresponde à média da geração mensal verificada nos últimos 5 anos. Esta previsão é atualizada uma vez por ano e é responsável por estabelecer os montantes de geração eólica que serão simulados nos modelos energéticos de formação de preço.

Desvios na previsão mensal de geração eólica do PMO podem ser responsáveis por uma operação energética desotimizada, podendo haver deplecionamento de reservatórios e geração térmica fora da ordem de mérito. Desta forma, objetiva-se neste trabalho ajustar um modelo de previsão de fatores de capacidade eólicos mensais que tenha desempenho superior ao modelo simplificado vigente (médias móveis) utilizado nas simulações do PMO.

Para isso, este trabalho analisa a série temporal de fatores de capacidade da usina eólica Rio do Fogo, ajustando previsões realizadas pelos modelos ARIMA e Holt-Winters, e avaliando o desempenho da combinação entre esses dois métodos de previsão. A seção II traz a análise descritiva da série temporal escolhida. A seção III detalha a metodologia utilizada no ajuste dos modelos ARIMA, Holt-Winters e combinado, enquanto a seção IV traz os principais resultados das previsões utilizando os modelos propostos. Por fim, a seção V resume as conclusões do trabalho.

II. LEVANTAMENTO DOS DADOS E ANÁLISE DESCRITIVA

A fim de se obter uma série temporal de geração eólica com um maior número de observações, foi feito um levantamento do histórico de geração eólica mensal verificada, disponibilizado pelo ONS [1]. Pôde-se verificar que a usina eólica Rio do Fogo, localizada no Rio Grande do Norte, é aquela com maior histórico de geração verificada, em um período que compreende de maio/2006 a abril/2018 (145 observações).

Como a geração eólica de usinas futuras é estimada através de fatores de geração (fatores de capacidade) no âmbito do PMO [3], optou-se por trabalhar com os dados de geração normalizados pela capacidade instalada. O fator de capacidade é a geração média dividida pela capacidade instalada, conforme representado em (1).

$$FC = \frac{G_{media}}{C.Instalada} \quad (1)$$

G. C. Godinho is with the Electrical Engineering Department, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ, Brazil (email: gabriel.godinho@gmail.com).

P. M. Maçaira and R. C. Souza are with the industrial Engineering Department, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ, Brazil (email: paulamacaira@gmail.com; reinaldo@puc-rio.br).

A Fig. 1 apresenta os fatores de capacidade da usina eólica (UEE) Rio do Fogo, calculados para o horizonte de 2007-2017. A Fig. 2 ilustra os fatores de capacidade médios mensais, de onde é possível se observar a sazonalidade mensal da geração eólica dessa usina.

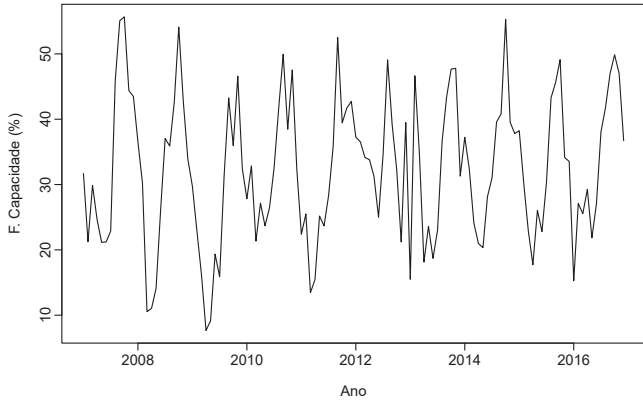


Fig. 1. Fatores de capacidade mensais UEE Rio do Fogo.

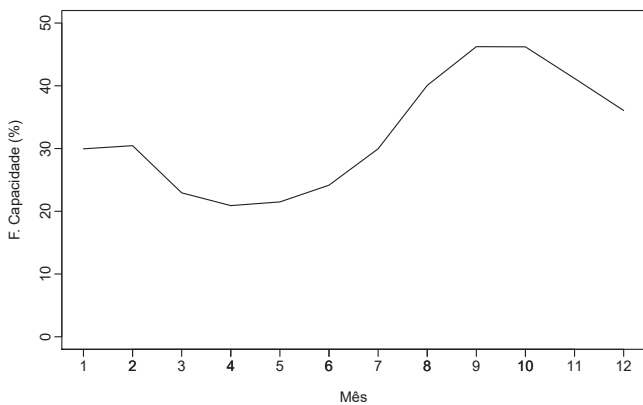


Fig. 2. Fatores de capacidade médios mensais da UEE Rio do Fogo.

Com base nos dados levantados, selecionou-se o horizonte de 2007-2016 para servir como ajuste do modelo de previsão a ser proposto. Dessa maneira, os dados verificados de 2017 podem ser utilizados na validação e escolha do melhor modelo de previsão. Não foram utilizados os dados do ano de 2006 pois estes apresentavam valores discrepantes relativos ao período de início da operação comercial da usina.

A Fig. 3 apresenta o histograma dos fatores de capacidade em comparação com uma distribuição normal, podendo-se observar que a distribuição da geração eólica mensal da UEE Rio do Fogo se aproxima de uma normal. Ao realizar-se o teste Qui-Quadrado de Pearson comparando as distribuições, obteve-se um p -valor = 0,253, acima do nível de significância de 5%. Dessa maneira, não é possível rejeitar a hipótese nula (H_0) de associação entre as variáveis, corroborando que a série de estudo possui distribuição Normal com média 32,3 e variância 123,8.

A análise do gráfico da série temporal, apresentado na Fig. 1, leva a crer que se trata de uma série estacionária com média e variância aproximadamente constantes. Para confirmar a hipótese de estacionariedade da série, foi aplicado o teste da raiz unitária (Augmented Dickey-Fuller), no qual

foi possível rejeitar a hipótese nula de não-estacionariedade (p -valor = 0,01), confirmando que a série é estacionária e por isso não possui qualquer tendência positiva ou negativa significativa.

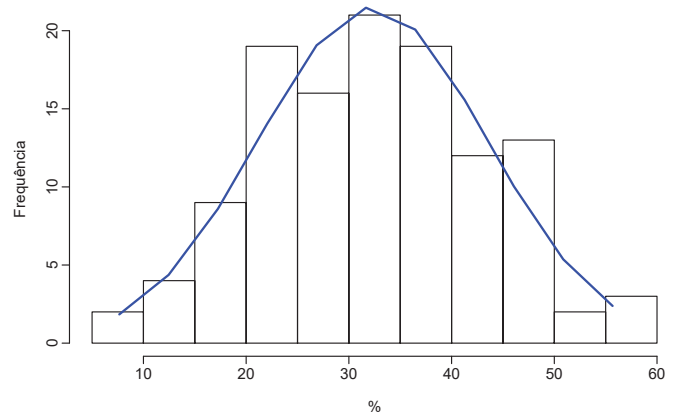


Fig. 3. Histograma da série histórica de fatores de capacidade.

Por fim, a Fig. 4 apresenta os gráficos da função de autocorrelação (FAC ou ACF) e autocorrelação parcial (FACP ou PACF). A análise da FAC permite concluir que a série em estudo apresenta decaimento exponencial e uma importante componente sazonal no lag 12, ou seja, os fatores de capacidade apresentam perfil de sazonalidade anual. Esta sazonalidade é evidentemente compreensível, uma vez que a geração eólica na região Nordeste do Brasil depende de processos físicos ligados à climatologia e estes possuem sazonalidade anual, assim como demais fenômenos naturais. A análise da FACP, por sua vez, demonstra um corte brusco no lag 2, indicando uma componente autoregressiva de ordem 1.

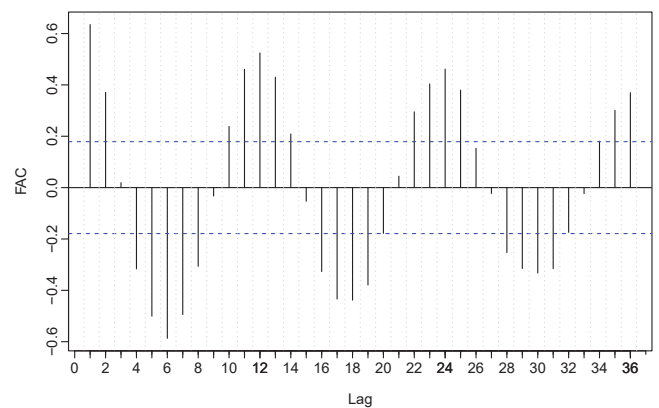


Fig. 4. Função de autocorrelação (FAC) e função de autocorrelação parcial (FACP) da série de fatores de capacidade.

III. METODOLOGIA E APLICAÇÃO AO CASO DE ESTUDO

O ajuste e previsão da série temporal de fatores de capacidade mensais da UEE Rio do Fogo podem ser originados por diferentes métodos, como ARIMA e amortecimento exponencial. Estes métodos são consagrados na literatura, sendo utilizados para previsão de diferentes classes de séries temporais.

Maçaira *et.al* [4] utilizaram métodos de amortecimento exponencial de Pegels, obtendo resultados satisfatórios na previsão de séries de consumo de energia elétrica até 2050.

Recentemente, Velásquez *et.al* [5] obtiveram resultados satisfatórios na combinação de previsões feitas por metodologias de amortecimento exponencial e SARIMA. Os autores provaram que a combinação utilizando neurônio multiplicativo único generalizado (GSMN), é preferível à combinação pela média como método de previsão para séries temporais.

Já Cardoso e Cruz [6] obtiveram melhores resultados de previsão de consumo de gás utilizando um modelo híbrido entre o método ARIMA e redes neurais artificiais, em vez de utilizar apenas o método ARIMA.

As seções a seguir irão descrever as principais características dos métodos de amortecimento exponencial e ARIMA, bem como as metodologias consideradas para ajuste e combinação dos modelos de previsão utilizados neste trabalho.

A. Box & Jenkins para Séries Sazonais (SARIMA)

Os modelos propostos por Box & Jenkins aplicados a séries temporais que possuem componentes sazonais podem ser denominados modelos ARIMA (*Autoregressive Integrated Moving Average*) multiplicativos, ou, simplesmente, SARIMA. De acordo com [7], esses modelos podem ser aplicados para a maioria das séries sazonais reais, ou seja, realizações de processos que possuem correlação serial intra e inter períodos sazonais.

A identificação dos modelos SARIMA pode ser feita com a definição dos parâmetros “ p , d , q ” e “ P_{12} , D_{12} , Q_{12} ” através das FACs e FACP. Avaliando novamente a Fig. 4 é possível identificar, além da sazonalidade no lag 12, picos não-decrescentes nos lags múltiplos de 12 para as autocorrelações estimadas, indicando que a série deve ser diferenciada pelo menos uma vez, no lag sazonal, para eliminar o efeito da não-estacionariedade sazonal. A Fig. 5 apresenta a série temporal diferenciada no lag 12, bem como os gráficos da FAC e FACP.

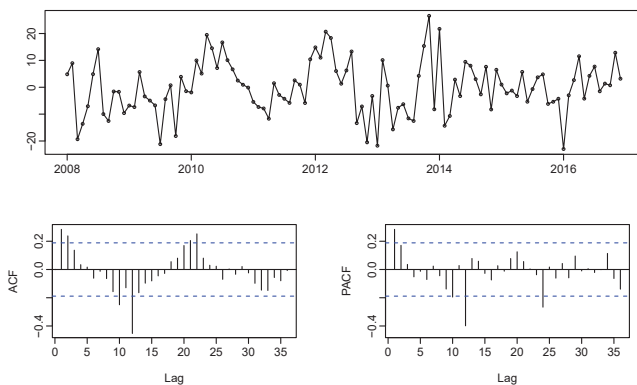


Fig. 5. Série diferenciada de fatores de capacidade, FAC e FACP.

Como dito anteriormente, dentro da metodologia Box&Jenkins é essencial a interpretação da FAC e FACP para definir quais modelos serão testados. Assim, nota-se na

Fig. 5 que os lags 1, 2, 10, 12, 23 e 24 são significantes, isto é, ultrapassam a linha tracejada que representam o intervalo de confiança, havendo um corte brusco nos lags sazonais. Da mesma forma, a análise do gráfico das autocorrelações parciais, mostra que existem componentes significantes nos lags 1, 12 e 24.

Logo, após análise inicial da FAC e FACP, os principais modelos ARIMA elegíveis para o ajuste da série são:

- I. ARIMA (1,0,2) x (2,1,2)₁₂
- II. ARIMA (1,0,2) x (1,1,2)₁₂;
- III. ARIMA (1,0,2) x (2,1,1)₁₂;e
- IV. ARIMA (1,0,2) x (1,1,1)₁₂, admitindo parcimônia.

Com auxílio do pacote *forecast* [8], disponível no *softwareR* [9], foram estimados inicialmente os parâmetros dos modelos elegíveis para o ajuste da série. Os coeficientes resultantes do ajuste destes modelos podem ser observados na Tabela I.

TABELA I
COEFICIENTES RESULTANTES DO AJUSTE DOS MODELOS ELEGÍVEIS

	σ^2 estimado	AIC	BIC
Modelo I (1,0,2) x (2,1,2)₁₂	0,07154	48,30	69,76
Modelo II (1,0,2) x (1,1,2)₁₂	0,06949	46,36	65,14
Modelo III (1,0,2) x (2,1,1)₁₂	0,06788	46,35	65,12
Modelo IV (1,0,2) x (1,1,1)₁₂	0,06915	44,36	60,46

A análise dos coeficientes permite concluir que, apesar do modelo III apresentar menor variância estimada dos resíduos, o modelo IV se ajusta melhor à série, uma vez que possui menor AIC (Critério de Informação de Akaike) e BIC (Critério de Informação Bayesiano). Segundo [10], os modelos com menor BIC são preferíveis, uma vez que esse critério penaliza a quantidade de parâmetros no modelo e a variância estimada dos resíduos. O critério BIC é baseado em parte na função de verossimilhança e está relacionado ao AIC.

Com o objetivo de verificar se o número de parâmetros escolhido no modelo IV é o mais correto, foram realizados testes de subfixação e sobrefixação com modelos ARIMA de diferentes ordens. Segundo [7], o teste de sobrefixação consiste na elaboração de um modelo com um maior número de parâmetros que o modelo fixado, a fim de se corrigir algum erro que possa ter sido cometido na identificação do modelo.

A Tabela II apresenta os modelos testados com a comparação das variâncias estimadas dos resíduos e dos critérios AIC e BIC obtidos pela estimação dos parâmetros de cada modelo. De acordo com os resultados apresentados, apesar da variância estimada de resíduos do modelo IX ser superior à do modelo IV, inicialmente escolhido, os critérios AIC e BIC são sensivelmente menores. Neste caso, o modelo IX aparenta melhor se ajustar à série, uma vez que possui os menores AIC e BIC dentre os modelos testados.

TABELA II
RESULTADOS DOS TESTES DE SUBFIXAÇÃO E SOBREFIXAÇÃO

		AR(1)	MA(1)	MA(2)	SAR(1)	SAR(2)	SMA(1)	SMA(2)	σ^2 estimado	AIC	BIC
Modelo I	Coef.	0,4027	-0,0423	0,1511	-0,7833	-0,0872	-0,1712	-0,5901	0,07154	48,30	69,76
(1,0,2) x (2,1,2) ₁₂	E. Padrão	0,2744	0,2744	0,1414	1,4315	0,1784	1,4221	1,2849			
Modelo II	Coef.	0,3916	-0,0314	0,1536	-0,1472	-	-0,8273	-0,0643	0,06949	46,36	65,14
(1,0,2) x (1,1,2) ₁₂	E. Padrão	0,2795	0,2793	0,1425	0,9557	-	0,9113	0,9179			
Modelo III	Coef.	0,3889	-0,0284	0,1534	-0,0614	0,0216	-0,9363	-	0,06788	46,35	65,12
(1,0,2) x (2,1,1) ₁₂	E. Padrão	0,2838	0,2838	0,1432	0,2077	0,1795	0,6691				
Modelo IV	Coef.	0,3941	-0,034	0,1531	-0,0795	-	0,8898	-	0,06915	44,36	60,46
(1,0,2) x (1,1,1) ₁₂	E. Padrão	0,2774	0,2771	0,1423	0,1433	-	0,2538	-			
Modelo V	Coef.	0,5740	-0,1930	-	-0,0822	-	-0,8687	-	0,07012	43,415	56,825
(1,0,1) x (1,1,1) ₁₂	E. Padrão	0,1543	0,1777	-	0,1434	-	0,2171	-			
Modelo VI	Coef.	0,5771	-0,1952	-	-0,7856	-0,0935	-0,1530	-0,5729	0,07234	47,332	66,107
(1,0,1) x (2,1,2) ₁₂	E. Padrão	0,1542	0,1776	-	1,1694	0,1756	1,1652	1,0228			
Modelo VII	Coef.	0,5734	-0,1923	-	-0,1471	-	-0,8078	-0,0607	0,07053	45,41	61,503
(1,0,1) x (1,1,2) ₁₂	E. Padrão	0,1548	0,1781	-	0,8959	-	0,8629	0,8379			
Modelo VIII	Coef.	0,5709	-0,1899	-	-0,0593	0,0267	-0,9170	-	0,06903	45,391	61,484
(1,0,1) x (2,1,1) ₁₂	E. Padrão	0,1564	0,1795	-	0,2070	0,1762	0,5069	-			
Modelo IX	Coef.	0,4137	-	-	-0,0978	-	-0,8550	-	0,07051	42,459	53,187
(1,0,0) x (1,1,1) ₁₂	E. Padrão	0,0882	-	-	0,1414	-	0,2012	-			
Modelo X	Coef.	0,4125	-	-	-0,0557	0,0476	-0,9510	-	0,06739	44,386	57,796
(1,0,0) x (2,1,1) ₁₂	E. Padrão	0,0884	-	-	0,2092	0,1771	0,8500	-			

A Fig. 6 apresenta a análise gráfica e estatística dos resíduos do modelo IX, podendo-se interpretar que a variância é aproximadamente constante para os resíduos ao longo do tempo. Além disso, o histograma dos resíduos mostra que é plausível assumir que a distribuição dos erros de previsão se aproxima de uma normal, com média zero e variância constante. O teste Qui-Quadrado de Pearson forneceu um p-valor = 0,022, não sendo possível rejeitar a hipótese-nula de associação entre distribuições com nível de 1% de significância.

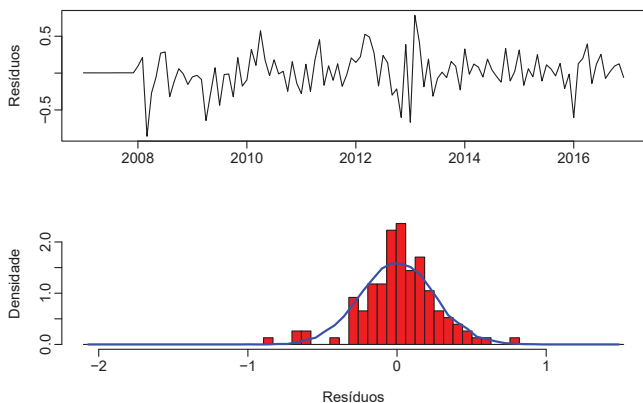


Fig. 6. Análise dos resíduos obtidos com o ajuste do modelo IX.

A aplicação do teste de Ljung-Box nos resíduos do modelo IX retornou p-valor = 0,74 para lags 1-50. O resultado deste teste, juntamente com a análise do gráfico da FAC dos resíduos do modelo IX (Fig. 7), permite concluir que não há evidência de dependência nas autocorrelações dos resíduos, ou seja, trata-se de ruído branco, como desejado.

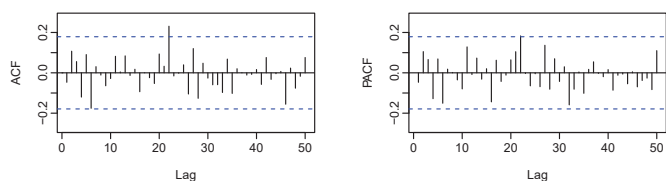


Fig. 7. FAC e FACP dos resíduos do modelo IX.

B. Modelo Holt-Winters (Amortecimento Exponencial)

Assim como os modelos SARIMA, os modelos de alisamento exponencial Holt-Winters também podem ser utilizados no ajuste de séries temporais que apresentam padrões de sazonalidade [11]. Estes modelos podem ter sazonalidade de efeito multiplicativo ou aditivo, conforme descrito em [12].

- Holt-Winters Multiplicativo:

$$\hat{y}_{t+h|t} = (\ell_t + hb_t) s_{t-m} \tag{2}$$

$$\ell_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1-\alpha) (\ell_{t-1} + b_{t-1}) \tag{3}$$

$$b_t = \beta (\ell_t - \ell_{t-1}) + (1-\beta) b_{t-1} \tag{4}$$

$$s_t = \gamma \left(\frac{y_t}{\ell_{t-1} + b_{t-1}} \right) + (1-\gamma) s_{t-m} \tag{5}$$

- Holt-Winters Aditivo:

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m} \tag{6}$$

$$\ell_t = \alpha (y_t - s_{t-m}) + (1-\alpha) (\ell_{t-1} + b_{t-1}) \tag{7}$$

$$b_t = \beta (\ell_t - \ell_{t-1}) + (1-\beta) b_{t-1} \tag{8}$$

$$s_t = \gamma (y_t - \ell_{t-1} + b_{t-1}) + (1-\gamma) s_{t-m} \tag{9}$$

As equações (2) e (6) descrevem os modelos de previsão, onde: $\hat{y}_{t+h|t}$ representa a previsão $t+h$ passos à frente no instante t ; ℓ representa o parâmetro de nível, calculado em (3) e (7); b representa o parâmetro de tendência, calculado em (4) e (8); e s representa o parâmetro de sazonalidade, calculado em (5) e (9). Os hiperparâmetros de α , β e γ são utilizados no cálculo das componentes de nível, tendência e sazonalidade, respectivamente.

Fazendo uso da função *hw* [8] no software R [9], foram ajustados dois modelos Holt-Winters à série de fatores de capacidade mensais. Conforme observado na Tabela III, os modelos ajustados apresentam valores bastante reduzidos para os hiperparâmetros α , β e γ . Isso significa que as estimativas de nível, tendência e sazonalidade de ambos modelos são baseadas, principalmente, em observações de um passado mais distante, conforme as equações (2) a (9).

TABELA III

COEFICIENTES RESULTANTES DO AJUSTE DOS MODELOS HOLT-WINTERS

	α	β	γ	σ^2 estimado	AIC	BIC
Holt-Winters Multiplicativo	0,0262	0,0001	0,0001	0,06131	1084,50	1131,89
Holt-Winters Aditivo	0,1141	0,0001	0,0001	51,69	1064,77	1112,16

Ainda de acordo com a Tabela III, apesar do modelo Holt-Winters aditivo possuir menores AIC e BIC que o modelo multiplicativo, sua variância estimada dos resíduos foi significativamente superior. Pode-se concluir que, neste caso, o modelo Holt-Winters multiplicativo é mais adequado para a previsão da série mensal de fatores de capacidade eólicos.

A metodologia de análise dos resíduos do modelo Holt-Winters foi feita analogamente àquela realizada no ajuste do modelo SARIMA. Pôde-se verificar que a distribuição de erros se aproxima de uma Normal com média igual a 0. O teste

Qui-Quadrado de Pearson forneceu um p-valor = 0,2578, não sendo possível rejeitar a hipótese-nula de associação entre distribuições. Porém, após avaliação da FAC e aplicação do teste de Ljung-Box, foi possível rejeitar a hipótese-nula de que os resíduos são independentes e identicamente distribuídos, indicando que, possivelmente, nem todas informações da série temporal estão sendo capturadas pelo modelo ajustado.

C. Modelo Combinado (SARIMA + Holt-Winters)

As previsões com os modelos SARIMA e Holt-Winters de 1 a 12 passos à frente ajustados em III-A e III-B, respectivamente, foram comparadas à série verificada de fatores de capacidade de 2017 (período *out-of-sample* para ajuste dos modelos). Para avaliar a precisão dos modelos, utilizou-se como métrica o erro absoluto médio percentual (MAPE), apresentada em [13]. O MAPE é a média da diferença absoluta entre os valores previstos e os verificados, expressa em porcentagem dos valores verificados, conforme demonstrado em (10).

$$MAPE = \sum_{t=1}^n \left| \frac{(y_t - \hat{y}_t) / y_t}{n} \right| \cdot 100 \quad \text{se } y_t \neq 0 \quad (10)$$

A Fig. 8 apresenta a série prevista e o MAPE calculado de 1 a 12 passos à frente para os modelos ajustados SARIMA e Holt-Winters multiplicativo. Apesar da indicação do teste de Ljung-Box de que o modelo Holt-Winters possivelmente não é o modelo mais adequado para ajuste da série, ele possui resultados mais satisfatórios que o modelo SARIMA para previsões 4, 6, 7, 9 e 10 passos à frente, ou seja, para quase metade do horizonte de previsão de 12 meses. O MAPE médio calculado para os 12 períodos foi de 11,23% para SARIMA e 11,02% para o modelo Holt-Winters. Além disso, ambos modelos fornecem previsões que se aproximam em nível dos valores verificados e acompanham o movimento da série, respeitando as componentes de sazonalidade, sem que houvesse um modelo que se destacasse nas previsões de curto ou longo prazo.

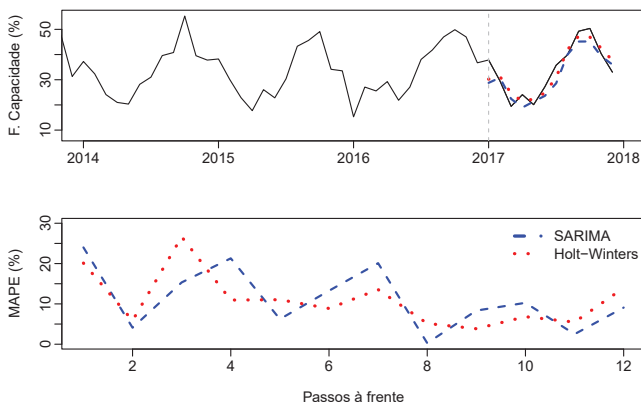


Fig. 8. Série prevista e MAPE de 1 a 12 passos à frente para os modelos ajustados SARIMA e Holt-Winters.

Por conta do bom ajuste de ambas metodologias (MAPE reduzido), testou-se a combinação entre as previsões dos modelos propostos. Desta forma, foram avaliadas duas abordagens descritas em [14] para a combinação dos modelos de previsão:

- Média simples das previsões:

$$Comb_{t+h|t}^{Med} = \frac{\hat{y}_{t+h|t}^{Pior} + \hat{y}_{t+h|t}^{Melhor}}{2} \quad (11)$$

- Combinação baseada na variância:

$$Comb_{t+h|t}^{Var} = k_h \hat{y}_{t+h|t}^{Melhor,h} + (1-k_h) \hat{y}_{t+h|t}^{Pior,h} \quad (12)$$

$$k_h = 1 - \frac{\min \{ \mathbb{V} [e_h^{BJ}], \mathbb{V} [e_h^{HW}] \}}{\mathbb{V} [e_h^{BJ}] + \mathbb{V} [e_h^{HW}]} \quad (13)$$

Enquanto a primeira abordagem, média simples, associa o mesmo peso para a previsão de cada modelo em (11), a combinação baseada na variância dá um maior peso aos modelos que apresentam menor variância entre a média dos valores amostrais *in-sample* (2007-2016) e os valores previstos para cada um dos modelos, de 1 a 12 passos à frente. Na equação (13), o peso k_h é calculado através do modelo que fornece a menor variância dos resíduos em cada passo h . Em seguida, os pesos são aplicados em (12) de forma que o modelo que teve melhor desempenho em cada passo tenha maior ponderação na combinação baseada na variância.

A Tabela IV apresenta os pesos calculados para cada modelo na combinação baseada na variância.

TABELA IV
PESOS APLICADOS AOS MODELOS NA COMBINAÇÃO BASEADA NA VARIÂNCIA

Passo	SARIMA	Holt-Winters
1	0,87155	0,12845
2	0,89898	0,10102
3	0,69009	0,30991
4	0,21685	0,78315
5	0,84708	0,15292
6	0,93277	0,06723
7	0,78482	0,21518
8	0,99346	0,00654
9	0,77257	0,22743
10	0,78856	0,21144
11	0,31500	0,68500
12	0,92799	0,07201

IV. RESULTADOS

Para avaliar o desempenho dos modelos ajustados, foram realizadas previsões de 1 a 12 passos à frente, que foram comparadas com os fatores de capacidade verificados no ano de 2017. De acordo com [3], o cálculo dos fatores de geração das usinas eólicas em expansão deve ser realizado uma vez ao ano, para consideração nos estudos de planejamento energético. Os fatores mensais de geração de usinas eólicas existentes (em operação comercial) são calculados através da média de geração mensal de cada usina nos últimos 5 anos. Esses fatores típicos mensais são projetados no horizonte de planejamento, repetindo-se o comportamento típico mensal ao longo de 5 anos. Assim, para avaliação das condições de atendimento energético e definição de preço da energia no mercado de curto-prazo, a geração de usinas eólicas existentes é projetada como uma média mensal da geração dos últimos cinco anos.

De posse dos fatores de capacidade verificados e da previsão oficial (média móvel) feita para a UEE Rio do Fogo em 2017, foram comparados os modelos de previsão propostos neste trabalho, a fim de se constatar se algum destes modelos apresenta melhores resultados que o modelo oficial. Para isso, foram avaliadas a raiz do erro quadrático médio (RMSE) e MAPE.

A RMSE, expressa em (14), representa o desvio padrão amostral das diferenças entre os valores previstos e observados e, assim como o MAPE, é amplamente usada para expressar a precisão de previsões numéricas [15].

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \tag{14}$$

A Fig. 9 apresenta os resultados do MAPE de 1 a 12 passos à frente para os modelos ajustados, em comparação com o modelo oficial. A Tabela V apresenta os resultados de MAPE e RMSE totais para o período de validação. Ressalta-se que como a previsão oficial é atualizada uma vez ao ano, foram utilizados dados de previsão de até 12 passos à frente.

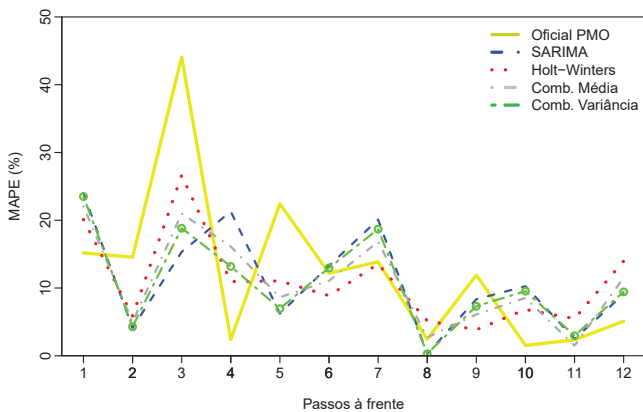


Fig. 9. MAPE de 1 a 12 passos à frente para os modelos ajustados em comparação ao modelo oficial utilizado no PMO.

TABELA V
PRECISÃO DOS MODELOS AJUSTADOS NO PERÍODO DE VALIDAÇÃO

Modelo	RMSE	MAPE
Oficial PMO (Médias móveis)	4,290	12,33%
Box & Jenkins (SARIMA)	4,463	11,23%
Holt-Winters Multiplicativo	3,815	11,02%
Combinação pela média (BJ+HW)	4,035	10,92%
Combinação pela variância (BJ+HW)	4,188	10,66%

Através da análise dos resultados pôde-se observar:

- Apesar do modelo Oficial PMO ter o maior MAPE, à exceção das previsões para os meses de janeiro, abril, outubro e dezembro (1, 4, 10 e 12 passos à frente, respectivamente), ele se ajusta na mesma faixa de erro que os demais modelos, exibindo RMSE inferior ao modelo SARIMA.

- Individualmente, o modelo Holt-Winters multiplicativo se ajusta melhor à série do que o modelo SARIMA.
- O modelo que melhor se ajustou aos valores verificados em 2017 foi a combinação das previsões SARIMA e Holt-Winters, ponderada pela variância.
- A combinação pela variância garantiu erros médios 14% menores (-1,7 p.p) que o modelo oficial.

V. CONCLUSÃO

A crescente expansão da geração eólica na região Nordeste do Brasil tem trazido benefícios no atendimento energético do país e manutenção dos volumes armazenados nos principais reservatórios. Todavia, uma previsão mensal de geração eólica que se afasta muito da geração verificada pode trazer sinais incoerentes para o atendimento sistêmico, afetando os montantes programados de geração térmica, resultando em encargos para o consumidor.

Com o objetivo de propor um modelo que estime a geração eólica mensal com maior precisão que o modelo vigente, foi estudada a série de geração da UEE Rio do Fogo, usina com maior histórico verificado disponibilizado pelo ONS. Através da caracterização dos fatores de capacidade mensais dessa série no horizonte de 2007-2016, foram propostos quatro modelos distintos de previsão para o ano de 2017.

O modelo encontrado que fornecia o melhor ajuste combinava as previsões de um modelo ARIMA (1,0,0) x (1,1,1)₁₂ e um modelo Holt-Winters multiplicativo, ponderando as previsões de cada um deles pela variância. Ao comparar-se os resultados deste modelo frente aos fatores de capacidade mensais verificados de 2017, obteve-se um MAPE de 10,66% para as previsões de 1 a 12 passos à frente. O erro absoluto médio percentual encontrado é 14% menor que aquele fornecido pelo modelo de previsão vigente.

Dessa maneira, conclui-se que a combinação SARIMA/Holt-Winters ponderada pela variância pode ser aplicada com maior segurança na previsão dos fatores típicos de geração mensal da UEE Rio do Fogo no âmbito do Programa Mensal da Operação e de demais estudos energéticos realizados pelo ONS. Sugere-se que sejam avaliados como trabalhos futuros:

- Ajuste de modelos de previsão de geração eólica mensal análogos ao apresentado, para as demais usinas eólicas do SIN;
- Estudo aprofundado da correlação entre a geração eólica do Nordeste e os cenários hidrológicos considerados no planejamento da operação energética; e
- Análise do impacto da sazonalidade da geração eólica mensal de usinas do litoral e do interior do Nordeste, dado o perfil diferenciado de geração horária destas regiões.

REFERÊNCIAS

[1] ONS, “Histórico da Operação,” June 2018. [Online]. Available: http://ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/geracao_energia.aspx

[2] —, “Energia no Futuro - Programação da Operação,” June 2018. [Online]. Available: <http://ons.org.br/paginas/energia-no-futuro/programacao-da-operacao>

[3] ANEEL, “Resolução Normativa Aneel nº 440/2011,” July 2011. [Online]. Available: <http://www2.aneel.gov.br/cedoc/ren2011440.pdf>

- [4] P. M. Maçaira, R. C. Souza, and F. L. C. Oliveira, "Forecasting Brazil's electricity consumption with Pegels Exponential Smoothing Techniques," *IEEE Latin America Transactions*, vol. 14, no. 3, pp. 1252–1258, Abril 2016.
- [5] J. D. Velasquez, C. O. Zambrano, and C. J. Franco, "Forecast Combining Using a Generalized Single Multiplicative Neuron," *IEEE Latin America Transactions*, vol. 12, no. 4, pp. 713–717, July 2014.
- [6] C. A. V. Cardoso and G. L. Cruz, "Forecasting Natural Gas Consumption using Arima Models and Artificial Neural Networks," *IEEE Latin America Transactions*, vol. 14, no. 5, pp. 2233–2238, Agosto 2016.
- [7] R. C. Souza and M. E. Camargo, *Análise e Previsão de Séries Temporais: Os Modelos ARIMA*, vol. 2004.
- [8] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: the forecast package for R," *Journal of Statistical Software*, vol. 26, no. 3, pp. 1–22, 2008.
- [9] R. F. for Statistical Computing, R. . A. Language, and E. for Statistical Computing, 2018, vienna.
- [10] G. E. Schwarz, *Estimating the dimension of a model*, 1978.
- [11] P. A. Morettin and C. M. d. C. Toloí, *Previsão de Séries Temporais*, 1985.
- [12] R. J. Hyndman and G. Athanasopoulos, *Forecasting Principles and Practice*, October 2018. [Online]. Available: <https://otexts.org/fpp2/holt-winters.html>
- [13] P. M. Swamidass, *Encyclopedia of Production and Manufacturing Management, Boston*, Boston, 2000.
- [14] C. W. J. Granger, *Forecasting in Business and Economics*. New York: Academic Press, Inc, 1980.
- [15] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," November 2005. [Online]. Available: <https://robjhyndman.com/papers/mase.pdf>



Reinaldo Castro Souza é graduado em Engenharia Elétrica pela UFJF em 1971, mestrado em Engenharia de Sistemas pela PUC-Rio em 1975, doutorado em estatística pela Warwick University, Coventry, Inglaterra em 1979 e pós-doutorado em modelos estatísticos e econométricos na London School of Economics, Inglaterra em 1986. Atualmente é professor titular da PUC-Rio. É especialista em análise e previsão de séries temporais (tema de sua tese de doutorado) e tem desenvolvido modelos estatísticos e de séries temporais em diversas áreas, tais como, pesquisas de mercado, previsões eleitorais, aplicações no setor energético e modelos econométricos. Tem também interesse na área de eficiência energética e uso racional de energia, com trabalhos de pesquisas para várias distribuidoras de energia elétrica, para o PROCEL/Eletrobrás, para o CEPEL, ONS, ANEEL e ANP.



Gabriel Campos Godinho é graduado em Engenharia de Energia pela Pontifícia Universidade Católica de Minas Gerais em 2012, e está cursando Mestrado em Engenharia Elétrica na Pontifícia Universidade Católica do Rio de Janeiro. Atualmente é Engenheiro de Sistemas de Potência no Operador Nacional do Sistema Elétrico (ONS). Tem experiência no planejamento da operação energética atuando na análise de riscos de déficit, custos marginais de operação, prospecção de armazenamentos de usinas hidroelétricas, perfil de geração de fontes intermitentes (eólica e solar), dentre outras condições de atendimento ao Sistema Interligado Nacional.



Paula Medina Maçaira possui graduação em Estatística pela Escola Nacional de Ciências Estatísticas (ENCE) em 2012, Mestrado em Engenharia Elétrica (Métodos de Apoio à Decisão) pela Pontifícia Universidade Católica do Rio de Janeiro (2015) e está cursando Doutorado em Engenharia de Produção na Pontifícia Universidade Católica do Rio de Janeiro. Atualmente é Pesquisadora do Projeto de P&D "Sistema Integrado de Apoio à Elaboração de Cenários de Projeção de Demanda e Planejamento da Geração" da ANEEL em colaboração com a PUC-Rio. Tem experiência na área de Probabilidade e Estatística, atuando principalmente nos seguintes temas: Séries Temporais, Previsão e Simulação.