# COMPARAÇÃO DO DESEMPENHO ENTRE MODELOS LINEARES E REDES NEURAIS ARTIFICIAIS PARA PREVISÃO DE DEMANDA POR DE ENERGIA

**Carolina Deina**
Universidade Federal do Rio Grande do Sul – Campus Porto Alegre
Av. Osvaldo Aranha, 99, Porto Alegre - RS, 90035-190
caroldeina@gmail.com

**Flavio Trojan**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
trojan@utfpr.edu.br

**Flávio Sanson Fogliatto**
Universidade Federal do Rio Grande do Sul – Campus Porto Alegre
Av. Osvaldo Aranha, 99, Porto Alegre - RS, 90035-190
ffogliatto@gmail.com

**Matheus Henrique do Amaral Prates**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
matheushaprates22@gmail.com

**Hugo Valadares Siqueira**
Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa
Av. Monteiro Lobato, s/n - Km 04 – Ponta Grossa – PR, 84016-210
hugosiqueira@utfpr.edu.br

## RESUMO

Para que não haja produção excessiva de eletricidade ou abaixo da demanda, o planejamento de políticas eficientes é essencial. Para isso, torna-se necessário compreender o comportamento da demanda, pois uma produção que não à atende causa racionamentos; já a produção excessiva gera desperdícios e custos elevados. Esse trabalho teve como objetivo analisar o desempenho de métodos para previsão de demanda de energia, baseando-se em uma comparação entre métodos lineares e Redes Neurais Artificiais. São testados os métodos de Suavização Exponencial, modelo Auto Regressivo, Auto Regressivo Integrado de Médias Moveis, Perceptron de Múltiplas Camadas, Função de Base Radial (RBF) e Máquinas de Aprendizado Extremo (ELM). Para realizar a comparação entre os métodos, selecionou-se a série histórica de consumo de energia elétrica do Paraná. As previsões foram realizadas considerando 1, 3 e 6 passos à frente. As RNAs apresentaram desempenho superior em relação aos modelos lineares, destacando-se as redes ELM e RBF.

**PALAVRAS CHAVE. Consumo de Energia, Séries temporais, Modelos Computacionais.**

**Tópicos (Inteligência Computacional, Simulação, Metaheurísticas)**

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

**ABSTRACT**

Efficient planning and adoption of policies are essential to avoid excessive or scarce electricity production. To enable that it is necessary to model demand behavior, since producing below demand may lead to electricity rationing, while excessive production represent waste and high cost. This work aims at analyzing the performance of methods to forecast electricity demand, carrying out a comparison between linear methods and Artificial Neural Networks (ANNs). Methods investigated are Exponential Smoothing (SE), Auto Regressive (AR), Auto Regressive Integrated Moving Average (ARIMA), Multilayer Perceptron (MLP), Radial Basis Function (RBF) and Extreme Learning Machines (ELM). Comparison was carried out using historical series of electric energy consumption in the State of Paraná, Brazil. Forecasting methods were implemented to predict demand 1, 3 and 6 steps ahead. Results indicated that ANNs had the best performance in comparison to linear models, most notably ELM and RBF networks.

**KEYWORDS. Energy consumption, Time series, Computational models.**

**Paper topics (Computational Intelligence, Simulation, Metaheuristics)**

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

## 1. Introdução

O consumo de energia elétrica vem crescendo em uma escala global, representado hoje aproximadamente 21.771,69 bilhões de kWh, sendo que países desenvolvidos consomem 43,26% desse montante [CIA, 2018]. Grande parte desse crescimento ocorreu nos setores residencial, comercial e de serviços públicos. O setor industrial continua sendo o maior consumidor global de energia, mesmo com sua participação reduzindo ao longo dos anos em decorrência do aumento de eficiência energética [IEA, 2018]. Demais setores que consomem eletricidade, como agropecuário e de transportes, tem participação reduzida, exceto pelo setor transportes rodoviários, que apresenta expansão de demanda devida à proliferação de veículos elétricos em países desenvolvidos [IEA, 2018].

O bloco de países em desenvolvimento que integram o BRIC (Brasil, Rússia, Índia e China) representam hoje 38,39% do consumo global. Nesses países, a demanda industrial é predominante e responde por metade do consumo [CIA, 2018]. Países do BRIC enfrentam desafios contínuos no setor energético. Para que se mantenham em funcionamento, é necessário aperfeiçoar decisões já tomadas e incorporar questões futuras como, por exemplo, a expansão da matriz energética, garantia da segurança energética, atendimento aos objetivos do desenvolvimento sustentável e garantia da justiça tarifária [MME, 2019].

Como eletricidade não é estocável, desequilíbrios entre oferta e demanda devem ser evitados, pois implicam em custos que acabam sendo repassados aos consumidores [CHUJAI et al., 2013]. Tanto o excesso quanto a redução da produção de eletricidade são nocivos à economia de um país: enquanto a produção excessiva gera desperdícios, a escassez pode provocar racionamentos [TUTUN et al., 2015]. Nesse sentido, para que decisões sejam viáveis economicamente é essencial entender o comportamento da demanda, o que possibilita identificar tendências futuras e alocar corretamente investimentos no setor. Como a demanda por eletricidade pode ser prevista por diversos modelos matemáticos e/ou estatísticos, é preciso encontrar aquele que melhor se adeque à série temporal a ser prevista, a um custo e tempo de processamento adequados [WU et al., 2018].

Este trabalho tem por objetivo encontrar o modelo que melhor se adeque à série temporal de consumo de energia do estado do Paraná, de forma a prover subsídios que auxiliem no processo de decisão sobre políticas e investimentos no setor. Para tanto, comparam-se os desempenhos dos modelos de redes neurais Perceptron de Múltiplas Camadas (MLP), Função de Base Radial (RBF) e Máquina de Aprendizado Extremo (ELM) com os modelos lineares de Suavização Exponencial (SE), Auto Regressivo (AR) e Auto Regressivo Integrado de Médias Móveis (ARIMA).

O trabalho apresenta uma contribuição teórica à medida em que discute características dos modelos de previsão investigados, as quais podem se adequar à análise da série de demanda por energia elétrica. Também apresenta uma contribuição prática, ao prover subsídios para aperfeiçoar a tomada de decisão acerca de investimentos no setor de geração e distribuição de energia elétrica no estado do Paraná.

## 2. Referencial teórico

Nessa seção são apresentados os conceitos principais e processos de modelagem dos métodos lineares de previsão (SE, AR e ARIMA) e as arquiteturas de Redes Neurais Artificiais (MLP, RBF e ELM) investigados no trabalho.

### 2.1 Suavização Exponencial de Holt Winters (SEHW)

O modelo SEHW pode ser utilizado quando a série temporal é composta por nível, tendência e sazonalidade. O ajuste do modelo à série temporal se dá através da estimação das constantes de nível ($\alpha$), tendência ($\beta$) e sazonalidade ($\gamma$), que podem compor uma série [MAKRIDAKIS *et al.*, 1998]. Os valores das constantes variam entre 0 e 1 [MENEGHINI *et al.*, 2018].

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

Os modelos exponenciais podem ser do tipo multiplicativo (para séries em que a amplitude da variação sazonal aumenta ou reduz ao longo do tempo) ou aditivo (para séries em que a amplitude da variação sazonal é constante ao longo do tempo). Esses modelos estão apresentados nas equações (4) e (5) [MORETTIN; TOLOI, 2006].

$$Ciclo: \dot{Z}_t = \alpha\left(Z_t - \hat{G}_{t-s}\right) + (1-\alpha)\left(\hat{Z}_{t-1} + \hat{H}_{t-1}\right), \;\; 0 < \alpha < 1, \;\; t = s+1, \dots, N \quad (1)$$

$$Tendência: \hat{H}_t = \beta\left(\dot{Z}_t - \dot{Z}_{t-1}\right) + (1-\beta)\hat{H}_{t-1}, \quad 0 < \beta < 1, \;\; t = s+1, \dots, N \quad (2)$$

$$Sazonalidade: \hat{G}_t = \gamma(Z_t - \dot{Z}_t) + (1-\gamma)\hat{G}_{t-s}, \;\; 0 < \gamma < 1, \quad e\; t = s+1, \dots, N \quad (3)$$

$$\hat{Z}_{t+1}(h-1) = \left(\dot{Z}_{t+1} + (h-1)\hat{H}_{t+1}\right)\hat{G}_{t+1+h-2s}, \qquad h = s+2, \dots, 2s+1 \quad (4)$$

$$\hat{Z}_{t+1}(h-1) = \dot{Z}_{t+1} + (h-1)\hat{H}_{t+1} + \hat{G}_{t+1+h-2s}, \qquad h = s+2, \dots, 2s+1 \quad (5)$$

Nas equações acima, $\dot{Z}_t$, $\hat{H}_t$ *e* $\hat{G}_t$ representam as estimativas de nível, tendência e sazonalidade, $\alpha$, $\beta$ e $\gamma$ são as respectivas constantes de suavização, $s$ corresponde ao comprimento sazonal, e $\hat{Z}_{t+1}(h-1)$ é o valor previsto da série para $h$ passos à frente.

**2.2 Modelo Auto Regressivo (AR)**

O modelo preditivo AR é adequado para séries temporais com variância e média estáveis ao longo do tempo [HYNDMAN; ATHANASOPOULOS, 2018]. Sua representação matemática é dada pela eq. (6), em que $\varepsilon_t$ é um termo de erro ou ruído branco [MORETTIN; TOLOI, 2006].

$$y_t = \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \quad (6)$$

Neste modelo, os valores dos coeficientes livres $\varphi_p$ podem ser encontrados de forma determinística aplicando as equações de Yule-Walker, de forma a minimizar a soma do quadrado dos erros [BOX *et al.*, 2016].

**2.3 Modelo Autorregressivo Integrado de Médias Móveis (ARIMA)**

O ARIMA é um modelo linear de previsão que resulta da combinação dos componentes de autoregressão (AR), diferenciação (I) e média móvel (MA) [OSHODI *et al*. 2017]. Diferente de outros modelos consolidados (como AR e MA) que requerem a condição de estacionariedade, o ARIMA é uma alternativa para séries que possuam tendência [BOX *et al.*, 2016].

Os componentes do Modelo ARIMA tem as seguintes funções [Box *et al*., 2016]: AR assume o papel de modelar a relação dos dados atuais com os dados passados, assumindo uma ordem $p$; I corresponde ao número de diferenciações de ordem $d$ necessárias para remover a tendência e tornar a série estacionária; MA pressupõe que os valores previstos dependem dos erros de previsão de períodos anteriores de ordem $q$ [HYNDMAN; ATHANASOPOULOS, 2018]. O modelo ARIMA $(p, d, q)$ é representado pela seguinte equação:

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (7)$$

nesta equação $\varphi_0$ é uma constante, $\varphi_1, \dots, \varphi_p$ são ponderações para os valores passados de $y_t$, $\varepsilon_t$ são ruídos brancos da série e $\theta_1, \dots, \theta_q$ possibilitam escrever a série em função de choques passados. Uma das formas de determinar as ordens $p$ e $q$ do modelo ARIMA é através da função de autocorrelação parcial (PACF) [BOX *et al*., 2016].

### 2.4 Perceptron de múltiplas camadas (MLP)

Constituídas de uma ou mais camadas escondidas de neurônios (camadas intermediárias) entre a camada de entrada e saída, a rede MLP é considerada uma das arquiteturas mais versáteis quanto ao seu potencial de aplicabilidade [HAYKIN, 2008]. A rede tem seus pesos sinápticos e de limiar definidos através de um processo de treinamento supervisionado, que utiliza o algoritmo de aprendizado de retropropagação de erro, também conhecido como *backpropagation* [AYYANATHAN; KANNAMMAL, 2014].

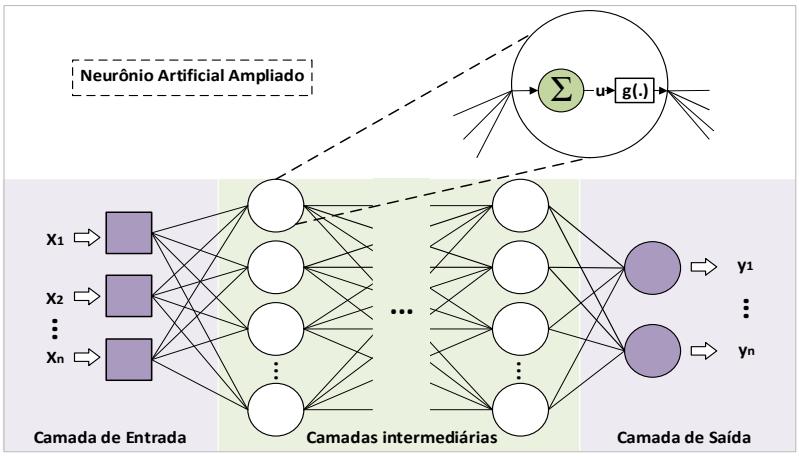



Figura 1 - Modelo genérico de uma rede neural artificial feedforward
Fonte: Adaptado de Silva, Spatti e Flauzino (2016)

O processo supervisionado utilizando *backpropagation* é realizado em duas fases. A primeira, conhecida por *forward*, é iniciada com a introdução e propagação dos dados da rede a fim de obter uma resposta de saída, com o valor dos pesos fixos. Na sequência, essa resposta é comparada com os valores desejados e uma medida de erro é obtida. Em cada iteração na fase de treinamento da rede, o objetivo é minimizar o erro [HAYKIN, 2008]. A segunda fase, conhecida por *backward*, executa os ajustes dos valores de pesos de acordo com a regra adotada de correção de erro, fazendo sucessivas iterações até que a saída dada pela rede se aproxime do valor desejado [HAYKIN, 2008]. Um algoritmo de primeira ordem, como o *backpropagation*, atualizará seus pesos sinápticos e os viéses de cada neurônio utilizando as eqs. (8) e (9) [SILVA et al., 2016].

$$w_{ij}^{m}(t+1) = w_{ij}^{m}(t) - \alpha \frac{\partial MSE(t)}{\partial w_{ij}^{m}(t)} \quad (8)$$

$$b_{j}^{m}(t+1) = b_{j}^{m}(t) - \alpha \frac{\partial MSE(t)}{\partial b_{j}^{m}(t)} \quad (9)$$

nesta equação $w_{ij}^{m}(t)$ é o peso sináptico da entrada $i$ do neurônio $j$ da camada $m$ na iteração $t$, e $b_{j}^{m}(t)$ é o viés do neurônio $j$ da camada $m$ na iteração $t$. A derivada da função $\partial MSE$ para cada camada é obtida pela regra da cadeia, a qual retropropaga o erro encontrado em direção à camada de entrada; $\alpha$ representa a etapa de otimização do algoritmo.

### 2.4 Função de Base Radial (RBF)

Devido a arquitetura de múltiplas camadas e padrão de interconexão *feedforward*, as redes RBF podem ser uma alternativa às redes MLPs, uma vez que possuem estrutura mais simplificada e processo de treinamento mais ágil [HAYKIN, 2008].

O treinamento das redes RBFs é realizado em duas etapas. A primeira, calcula os pesos dos neurônios da camada intermediária através de um método de aprendizagem não supervisionado [HAYKIN, 2008]. Considerando cada neurônio da camada intermediária representado por uma

função gaussiana [eq. (10)], com centro definido pelos pesos sinápticos dos neurônios, o objetivo é posicionar os centros das gaussianas da forma mais adequada possível [SILVA et al., 2016]. Após o cálculo dos centros das funções gaussianas, calcula-se suas variâncias, utilizando a eq. (11).

$$g(u) = -e^{-\frac{(u-c)^2}{2\sigma^2}} \qquad (10)$$

$$\sigma_j^2 = \frac{1}{|\Omega^j|} \sum_{x^k \epsilon \Omega^j} \sum_{i=1}^{n} (x_i^k - w_{ji}^1)^2 \qquad (11)$$

em que $\sigma_j^2$ representa a variância da função gaussiana do neurônio *j*, $|\Omega^j|$ é o tamanho do conjunto do neurônio *j*, $x^k$ é a amostra *k* do conjunto $\Omega$ , $x_i^k$ é a entrada *i* da amostra *k*, e $w_{ji}^1$ é o peso do neurônio *j* para a camada *i*. A saída de cada neurônio *j* da camada intermediária é expressa na eq. (12) [HAYKIN, 2008].

$$g_j^{(1)}(u_j^1) = g_j^{(1)}(x) = e^{-\frac{\sum_{i=1}^{n}(x_i - w_{ji}^{(1)})^2}{2\sigma_j^2}}, \qquad j = 1, \dots, n_1 \qquad (12)$$

Após o término do treinamento da camada intermediária, aplica-se a segunda etapa de treinamento, que consiste no ajuste dos pesos da camada de saída. Essa etapa ocorre como na rede MLP, podendo também ser utilizado o algoritmo de retropropagação do erro [SILVA et al., 2016].

**2.5 Máquinas de Aprendizado Extremo (ELM)**

A rede ELM foi proposta por Huang et al. (2004) com o objetivo de reduzir gargalos no processo de treinamento e aumentar a eficiência de aprendizagem das RNAs até então conhecidas. Para tanto, propõe-se uma rede com arquitetura semelhante à MLP, com padrão de interconexão *feedforward*, mas com apenas uma camada intermediária, que se diferencia pelo seu processo de treinamento [SIQUEIRA et al., 2012], no qual os pesos da camada intermediária são determinados de maneira aleatória e permanecendo sem ajuste, enquanto apenas os pesos da camada de saída são atualizados durante o processo de treinamento. Essas características conferem baixo custo computacional ao processo de treinamento e grande agilidade no processo de convergência dos resultados, se comparado às outras arquiteturas disponíveis [HUANG et al., 2004]. As saídas da camada intermediária são utilizadas como entradas para a camada de saída e seu sinal $x_t^h$ é dado pela eq. (13) [SILVA et al., 2016].

$$\mathbf{x}_\mathrm{t}^\mathrm{h} = \mathrm{f}^\mathrm{h}\left(\mathbf{W}^\mathrm{h}\mathbf{u}_\mathbf{t} + \mathrm{b}\right) \qquad (13)$$

em que $\mathrm{f}^\mathrm{h}(.)$ é a função de ativação dos neurônios da camada intermediária; $\mathbf{u_t} = [\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_{\mathrm{t-k+1}}]^T$ é o vetor das entradas, que tem seus componentes passados à camada intermediária; $\mathbf{W}^\mathrm{h} \epsilon \Re^{\mathrm{N_h K}}$ são pesos gerados de forma aleatória e *b* é o viés de cada neurônio.

Após encontradas as saídas da camada intermediária, estas são multiplicadas pela matriz $\mathbf{W}^{\mathrm{out}}$ de pesos sinápticos da camada de saída, para obter o valor de saída da rede através da eq. (14) [SILVA et al., 2016]. Dessa forma, o treinamento para a rede ELM consistirá em encontrar a matriz de pesos ($\mathbf{W}^{\mathrm{out}}$) com o menor erro para o vetor *d*, que contém todas as saídas desejadas. O operador de matriz pseudo-inversa de *Moore-Penrose* (Equação 15) é uma opção para resolver a tarefa de treinamento [SIQUEIRA *et al.*, 2012].

$$y_\mathrm{t} = \mathbf{W}^{\mathrm{out}}\mathbf{x}_\mathrm{t}^\mathrm{h} \qquad (14)$$

$$\mathbf{W}^{\mathrm{out}} = (\mathbf{x}_\mathrm{h}^\mathrm{T}\mathbf{x_h})^{-1}\,\mathbf{x}_\mathrm{h}^\mathrm{T}\mathbf{d} \qquad (15)$$

em que $\mathbf{x}^{\mathrm{h}} \in \Re^{\mathrm{T_s N_k}}$ representa a matriz que contem a saída da camada intermediária, $T_s$ é o número de amostras do treinamento e $(\mathbf{x}_{\mathrm{h}}^{\mathrm{T}}\mathbf{x}_{\mathbf{h}})^{-1}\,\mathbf{x}_{\mathrm{h}}^{\mathrm{T}}$ é a pseudo-inversa de $\mathbf{x}_{\mathbf{h}}$.

## 3. Metodologia

Para atingir os objetivos deste trabalho, uma aplicação numérica foi realizada através da coleta de dados brutos sobre o consumo de energia elétrica no Estado do Paraná, correspondente a 186 dados mensais, de janeiro de 2004 a junho de 2019. Os dados foram coletados na base disponível no site da Empresa de Pesquisa Energética, a qual é vinculada ao Ministério de Minas e Energia [EPE, 2019].

A primeira etapa da metodologia correspondeu ao pré-processamento desses dados coletados. Para tanto, realizou-se a análise estatística e gráfica da série temporal com o objetivo de identificar tendência, nível e sazonalidade. Conforme Morettin e Toloi (2006), a presença de tendência pode ser verificada pelo teste de Cox-Stuart; já sazonalidade pode ser verificada pelo teste de Friedman. Se for confirmada a presença de tendência, faz-se uma ou mais diferenciações [eq. (16)] até torná-la estacionária.

$$y'_t = y_{t+1} - y_t \qquad (16)$$

nesta equação $y'_t$ é a diferença entre i valor atual da série menos no valor anterior. Se um componente sazonal for identificado, um processo estatístico conhecido como dessazonalização [eq. (14)] é aplicado aos dados [BALLINI, 2000], isso é expresso na Equação (17).

$$z_i = \frac{x_i - \mu_m}{\sigma_m} \qquad (17)$$

em que $\mu_m$ é a média, $\sigma_m$ é o desvio padrão para cada ano. Tendo os dados estacionários para os modelos que exigem essa condição, seguiu-se para a etapa de previsão de demanda.

A série temporal do consumo de energia elétrica no Paraná foi particionada em três conjuntos. Uma porção (**a**) de treino (dados de Janeiro de 2004 a dezembro de 2015, totalizando 144 observações), uma porção (**b**) de validação (dados de janeiro de 2016 a dezembro de 2017, totalizando 24 observações) e uma porção (**c**) de teste (dados de janeiro de 2018 a junho de 2019, totalizando 18 observações). Após a partição dos dados, seguiu-se para a etapa de previsão. As previsões dos modelos lineares SEHW, AR e ARIMA foram obtidas através do software Minitab®, enquanto as previsões das redes MLP, RBF e ELM foram obtidas por programação em Python®.

Cada modelo foi simulado 30 vezes para todas as configurações possíveis, sendo retido o melhor resultado (menor soma dos quadrados dos resíduos, MSE) junto ao conjunto de validação. Como os dados eram mensais, foram feitas previsões considerando horizontes de previsão de 1, 3 e 6 meses ($P$ = 1, 3, 6). Ao final, os resultados das previsões pelos modelos propostos foram comparados utilizando as seguintes medidas de erro: Erro Quadrático Médio (MSE), Erro Absoluto Médio (MAE) e Erro Percentual Absoluto Médio (MAPE).

## 4. Resultados e Discussão

### 4.1 Análise e tratamento dos dados

Os dados da série temporal foram plotados (Figura 2), tendo sido identificados visualmente componentes de tendência (de crescimento) e sazonalidade, os quais foram verificados através dos testes não-paramétricos de Cox-Stuart e Friedman (Tabela 1). Considerando os *p*-valores inferiores a 0,05 em ambos os testes, comprovou-se estatisticamente a presença de tendência e sazonalidade na série, componentes estes removidos através da diferenciação [eq. (13)] e dessazonalização [eq. (14)] da série, conforme evidenciado pelo resultado dos testes estatísticos (Tabela 1) e pela análise visual (Figura 3). Os valores da série na Figura 2 variam entre +3 e -3, centrados em zero, indicando a condição de estacionariedade.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

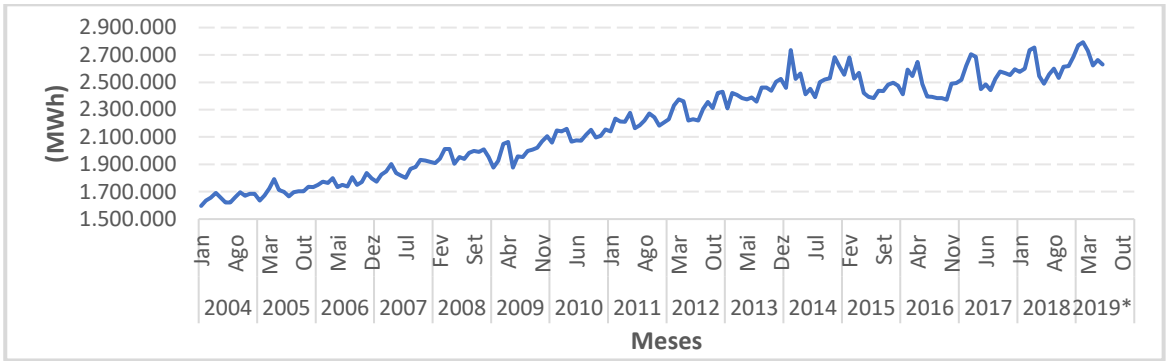



Figura 2 - Consumo mensal de energia no Estado do Paraná no período de janeiro de 2004 a junho de 2019
Fonte: Empresa de Pesquisa Energética (2019)

Tabela 1 – Testes não-paramétricos para análise de tendência e sazonalidade

| Testes | Série Original *p*-valor | Série diferenciada e dessazonalizada *p*-valor |
|---|---|---|
| Cox-Stuart | 0 | 0,251322103 |
| Friedman | 3,49E-15 | 0,994761543 |

Fonte: Autor

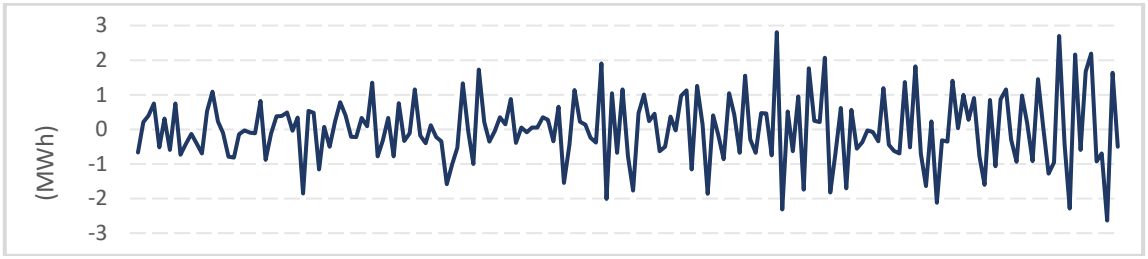



Figura 3 – Série diferenciada e dessazonalizada de consumo mensal de energia no Estado do Paraná
Fonte: Autor

Na sequência, definiram-se os parâmetros de cada modelo. O modelo SEHW é utilizado para prever séries que apresentam componentes de nível, tendência e sazonalidade [MENEGHINI et al., 2018]. O modelo apresenta constantes que suavizam esses componentes (α, β e γ, respectivamente). Os melhores resultados foram obtidos para o modelo aditivo, com $\alpha = 0{,}63$, $\beta = 0{,}12$, e $\gamma = 0{,}45$. Os parâmetros dos modelos AR($p$) e ARIMA($p$, $d$, $q$) foram estimados a partir da análise da função de autocorrelação parcial (FACP) para a série estacionária (Figura 4).

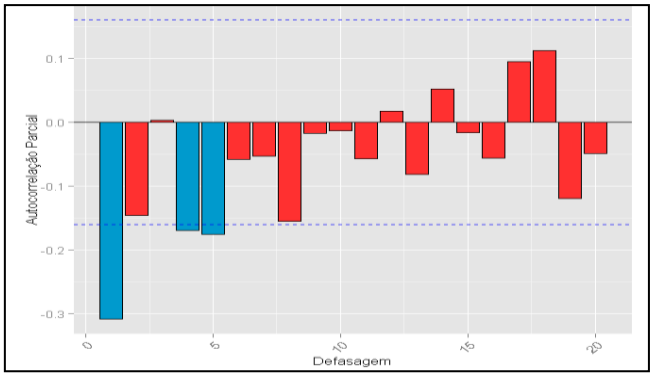



Figura 4 - Gráfico FACP para a série temporal estacionária de consumo de energia elétrica no Paraná
Fonte: Autor

O parâmetro d foi definido como 1, correspondendo ao número de diferenciações necessárias para tornar a série estacionária. Pela análise da Figura 3, percebe-se que a série apresenta coeficiente de autocorrelação parcial significativo até o quinto atraso, com um grau de confiança de 95%. Com os testes realizados considerando as possíveis configurações de 1 a 5 atrasos para p e q, os modelos AR (2) e ARIMA (5,1,5) foram os que apresentaram os menores erros de previsão. Para ajustar os modelos à série, utilizaram-se estimadores de máxima verossimilhança.

Para as RNAs, variou-se empiricamente a quantidade de neurônios na camada intermediária, iniciando com 5 neurônios e usando incrementos de 5, até atingir 200 neurônios. Para cada variação aplicou-se o método *wrapper*, que fez uma varredura progressiva considerando até 6 atrasos. As configurações das RNAs que obtiveram os melhores resultados de previsão foram: MLP de 5 neurônios na camada intermediária e com o terceiro atraso como entrada para previsão; RBF com 10 neurônios e atraso 6; ELM com 25 neurônios e atrasos 1 e 6.

### 4.2 Previsão do consumo de energia

Com os parâmetros definidos na seção anterior, foram feitas previsões para 1 passo à frente ($P = 1$) para a série de consumo de energia. Na Tabela 2 são apresentados os resultados das medidas MSE com valores observados, MSE (*d*) com valores diferenciados e dessazonalizados, MAE, MAE (*d*), e MAPE (por tratar-se de porcentagem, a medida não é sensível ao tratamento dos dados). Os comparativos usaram os dados da porção de teste.

Tabela 2 – Previsões para 1 passo à frente e medidas MSE, MAE e MAPE

| 1 passo à frente | | | | | | |
|---|---|---|---|---|---|---|
| **Modelo** | **Parâmetros do modelo** | **MSE** | **MSE (*d*)** | **MAE** | **MAE (*d*)** | **MAPE** |
| **SEHW** | $\alpha = 0{,}63$, $\beta = 0{,}12$, $\gamma = 0{,}45$ | 5679573386,83 | 0,2566 | 62124,83 | 0,4286 | 2,3265 |
| **AR** | p = 2 | 5247636249,96 | 0,2371 | 60064,66 | 0,4144 | 2,2765 |
| **ARIMA** | p = 5, d = 1, q = 5 | 5860003521,64 | 0,2647 | 65994,23 | 0,4553 | 2,5035 |
| **MLP** | neurônios = 5, atraso = 3 | 6557391893,34 | 0,3152 | 70231,17 | 0,4796 | 2,6589 |
| **RBF** | neurônios = 10, atraso = 6 | 4634913071,81 | 0,2094 | 55409,00 | 0,3823 | 2,0994 |
| **ELM** | neurônios = 25, atrasos = 1 e 6 | 4599769468,03 | 0,1945 | 48604,73 | 0,3463 | 1,8538 |

Fonte: Autor

Os resultados da Tabela 2 evidenciam que a melhor previsão foi obtida pela rede ELM, seguida da rede RBF, para todas as medidas de erro analisadas. O pior resultado de previsão foi apresentado pela rede MLP. A Figura 5 apresenta as previsões dadas pelos modelos e os dados reais separados para teste.

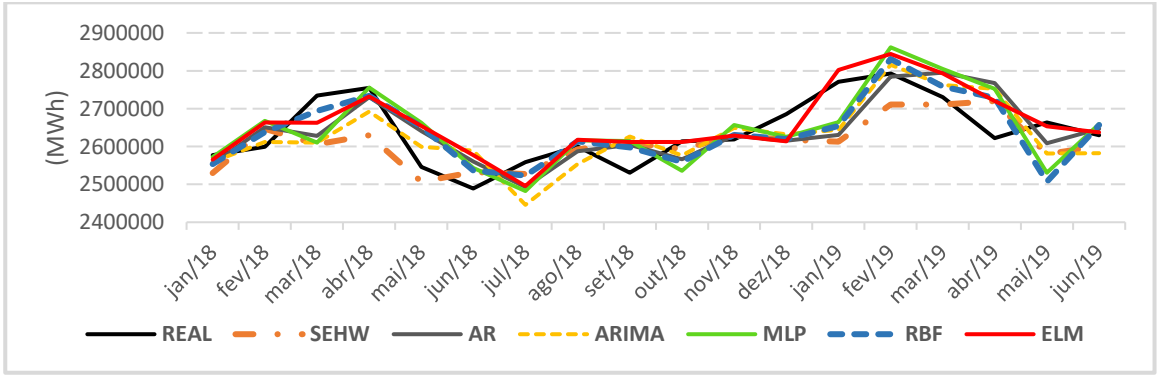



Figura 5 - Comparação entre dados reais e valores previstos para 1 passo à frente, referentes aos 18 meses dos anos de 2018 e 2019, utilizando os modelos SEHW, AR, ARIMA, MLP, RBF e ELM
Fonte: Autor

Embora visualmente os resultados na Figura 4 pareçam próximos, a série analisada tem consumo mensal na casa de milhões de MWh. A diferença entre o valor real e previsto, que está aproximadamente na casa do milhar (MWh), quando elevado ao quadrado e dividido pelo número de previsões (cálculo do MSE) pode chegar a casa de bilhões de MWh por mês, para todo o estado, já que o MSE penaliza erros maiores se comparado às outras medidas de desempenho. Considerando a escala dos valores a serem previstos, pode-se concluir que o modelo ELM apresentou um resultado bastante superior.

Na sequência, fez-se também previsões considerando o horizonte de 3 passos à frente, tendo seus resultados comparados pelas medidas MSE, MSE(d), MAE, MAE(d) e MAPE (Tabela 3). A Figura 6 apresenta as previsões realizadas pelos modelos e os dados reais separados para teste.

Tabela 3 – Previsões para 3 passos à frente e medidas MSE, MAE e MAPE

| **3 passos à frente** | | | | | | |
|---|---|---|---|---|---|---|
| **Modelo** | **Parâmetros do modelo** | **MSE** | **MSE (*d*)** | **MAE** | **MAE (*d*)** | **MAPE** |
| **SEHW** | $\alpha = 0,63$, $\beta = 0,12$, $\gamma = 0,45$ | 7485077440,82 | 0,3292 | 72273,26 | 0,4987 | 2,7328 |
| **AR** | p = 2 | 7013687861,15 | 0,3247 | 70897,07 | 0,4836 | 2,6770 |
| **ARIMA** | p = 5, d = 1, q = 5 | 8924931364,90 | 0,4853 | 80795,54 | 0,5770 | 3,0485 |
| **MLP** | neurônios = 5, atraso = 3 | 6573580787,03 | 0,3162 | 70331,85 | 0,4806 | 2,6625 |
| **RBF** | neurônios = 10, atraso = 6 | 6528243517,99 | 0,2871 | 62256,53 | 0,4296 | 2,3556 |
| **ELM** | neurônios = 25, atrasos = 1 e 6 | 6366837641,56 | 0,2740 | 64652,19 | 0,4296 | 2,4487 |

Fonte: Autor

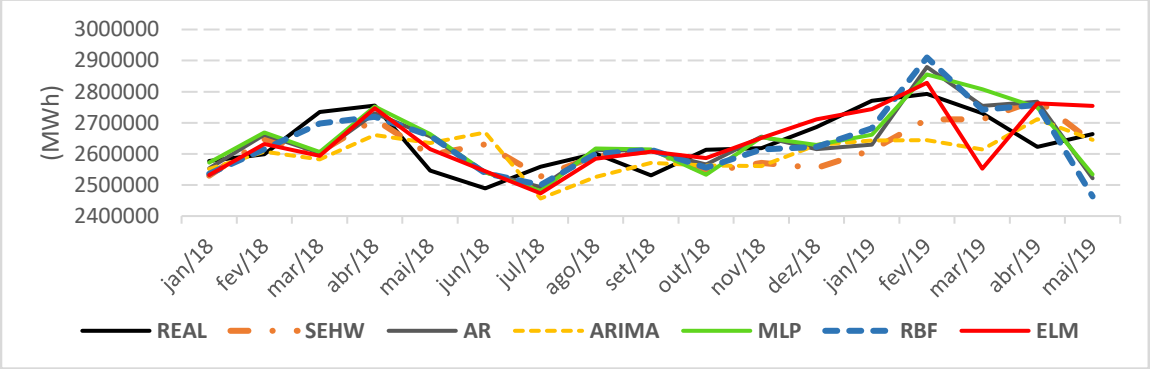



Figura 6 - Comparação entre dados reais e valores previstos para 3 passos à frente, referentes aos 18 meses dos anos de 2018 e 2019, utilizando os modelos SEHW, AR, ARIMA, MLP, RBF e ELM
Fonte: Autor.

Pela análise da Tabela 3, a rede ELM novamente apresentou desempenho de previsão superior as demais, considerando o MSE. Considerando MAE e MAPE, a rede RBF apresenta o melhor desempenho. Cabe ressaltar que MAE e MAPE não fazem distinção entre erros pequenos e grandes quando comparadas à MSE, que eleva os erros ao quadrado. Se comparado o MSE no domínio real, o erro da rede RBF em relação a ELM em termos percentuais foi de 4,62%, não muito distante da rede MLP em relação à ELM, que foi de 5,91%. Já os modelos AR, SEHW e ARIMA obtiveram um desempenho muito inferior se comparados à rede ELM.

Por fim, foram obtidas previsões considerando o horizonte de 6 passos à frente, tendo seus resultados comparados pelas medidas MSE, MSE(d), MAE, MAE(d) e MAPE, na Tabela 4. A Figura 7 traz as previsões realizadas pelos modelos e os dados reais separados para teste. Percebe-se que as redes ELM e RBF novamente apresentaram os melhores desempenhos, enquanto o modelo de SEHW apresentou o pior desempenho. Quando se compara o MSE do domínio real

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

da rede ELM em relação aos modelos RBF, MLP, AR, ARIMA e SEHW, o desempenho é reduzido respectivamente em 5,38%, 7,99%, 12,39%, 50,55% e 53,66%.

Tabela 3 – Previsões para 6 passos à frente e medidas MSE, MAE e MAPE

| | | 6 passos à frente | | | | |
|---|---|---|---|---|---|---|
| **Modelo** | **Parâmetros do modelo** | **MSE** | **MSE (*d*)** | **MAE** | **MAE (*d*)** | **MAPE** |
| **SEHW** | $\alpha = 0{,}63$, $\beta = 0{,}12$, $\gamma = 0{,}45$ | 8283400542,03 | 0,3916 | 76488,21 | 0,5268 | 2,8931 |
| **AR** | p = 2 | 6840287879,73 | 0,3175 | 71062,50 | 0,4815 | 2,6849 |
| **ARIMA** | p = 5, d = 1, q = 5 | 8174710719,13 | 0,3865 | 73181,97 | 0,5041 | 2,7387 |
| **MLP** | neurônios = 5, atraso = 3 | 6686402960,82 | 0,3168 | 69819,32 | 0,4766 | 2,6426 |
| **RBF** | neurônios = 10, atraso = 6 | 6595147324,83 | 0,3118 | 68586,05 | 0,4724 | 2,5950 |
| **ELM** | neurônios = 25, atrasos = 1 e 6 | 6406990346,37 | 0,3020 | 63890,27 | 0,4487 | 2,4310 |

Fonte: Autor

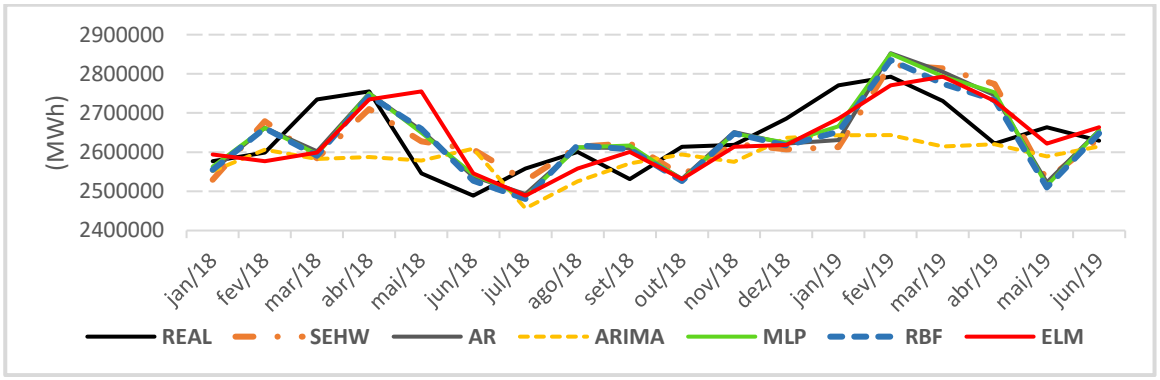



Figura 7 - Comparação entre dados reais e valores previstos para 3 passos à frente, referentes aos 18 meses dos anos de 2018 e 2019, utilizando os modelos SEHW, AR, ARIMA, MLP, RBF e ELM
Fonte: Autor

## Conclusões

Quando comparadas as previsões entre os modelos lineares e RNAs, considerando a série temporal de consumo de energia elétrica, as RNAs obtiveram desempenhos superiores, em uma análise geral, para todos os horizontes de previsão, exceto para a rede MLP em previsões de 1 passo à frente. À medida em que se aumentaram os horizontes para 3 e 6 passos à frente, a rede MLP, mesmo com desempenho inferior, teve desempenho superior aos modelos lineares. O desempenho superior das RNAs em relação aos modelos lineares já era esperado, visto que as redes, por serem não-lineares, ajustam-se melhor a séries temporais que possuem maior variabilidade.

Outro ponto a ser destacado é que as redes RBF e ELM, que obtiveram resultados superiores em relação a rede MLP, possuem custo computacional mais baixo em relação à esta. No caso da RBF, isso decorre de sua arquitetura; no caso da ELM, além de uma arquitetura mais simplificada, o processamento mais rápido também decorre da camada de saída tem seu treinamento baseado em um método determinístico.

## Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro ao projeto de pesquisa.

## Referências

AYYANATHAN, N.; KANNAMMAL, A. Share price time series forecasting for effective supply chain information exchange. **International Journal of Logistics Systems and Management**. v. 18, p. 139-158, 2014.

BALLINI, R. **Análise e previsões de vasões utilizando modelos de series temporais, redes neurais e redes neurais nebulosas**. 2000. Tese (Doutorado). Programa de Pós-Graduação em Engenharia Elétrica, Universidade Estadual de Campinas, Campinas, 2000.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Times Series Analysis** – Forecasting and Control. 5 ed. New Jersey: Wiley, 2016.

CHUJAI, P, KERDPRASOP N, KERDPRASOP K. Time series analysis of household electric consumption with ARIMA and ARMA models. **In: Proc. IMECS Conf.** Hong Kong, 2013.

CIA, Central Intelligence Agency. **Eletricidade, mapa de consumo mundial.** USA: CIA World Factbook, 2018. Disponível em: < https://www.cia.gov/library/publications/the-world-factbook/>. Acesso em: 10 ago. 2019.

EPE, Empresa de Pesquisa Energética. **Publicações/dados abertos**. Rio de Janeiro: Ministério de Minas e Energia, 2019. Disponível em: < http://www.epe.gov.br/pt/publicacoes-dados-abertos>. Acesso em: 11 jun. 2019.

HUANG, G. B.; ZHU Q. Y.; SIEW, C. K. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. **Proceedings of International Joint Conference on Neural Networks**. p. 985-990, 2004.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting:** principles and practice. Australia: Melbourne, 2018.

IEA, Agência Internacional de Energia. **Estatísticas de eletricidade.** 2018. Disponível em: < https://www.iea.org/statistics/electricity/>. Acesso em: 11 jun. 2019.

MAKRIDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R. J. **Forecasting methods and applications**. 3 ed. New York: John Wiley & Sons, 1998.

MENEGHINI, M.; ANZANELLO, M.; KAHMANN, A.; TORTORELLA, G. L. Ajuste de previsão de demanda quantitativa com base em fatores qualitativos: estudo de caso em um restaurante fast food. **Sistemas & Gestão.** v. 13, p. 68-80, 2018.

MME, Ministério de Minas e Energia. Estatísticas do setor energético brasileiro. 2019. Disponível em: < http://www.mme.gov.br/>. Acesso em: 11 jun. 2019.

MONTGOMERY, D.; JOHNSON, L.; GARDINER, J. **Forecasting and Time Series Analysis**. New York: McGraw-Hill, 1990.

MORETTIN, P.A.; TOLOI, C.M.C., **Análise de Séries Temporais**. Ed. Blucher, 2006.

OSHODI, O. S.; EJOHWOMU, O. A.; FAMAKIN, I. O.; CORTEZ, P. Comparing univariate techniques for tender price index forecasting: Box-Jenkins and neural network model. **Construction Economics and Building**. v. 17, p. 109–123, 2017.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais Para Engenharia e Ciências Aplicadas**: Fundamentos Teóricos e Aspectos Práticos. 2 ed. São Paulo: Artliber, 2016.

SIQUEIRA, H. V.; BOCCATO, L.; ATTUX, R. R. F.; LYRA FILHO, C. Echo state networks in seasonal streamflow series prediction. **Learning and Nonlinear Models**, v. 10, p. 181–191, 2012.

TUTUN, S.; CHOU, C. A.; CANYLMAZ, E. A new forecasting framework for volatile behavior in net electricity consumption: A case study in Turkey. **Energy**. v. 93, p. 2406-2422, 2015.

WU, L.; GAO, X.; XIAO, Y.; YANG, Y.; CHEN, X. Using a novel multi variable grey model to forecast the electricity consumption of Shandong Province in China. **Energy.** v. 157, p. 327-335, 2018.

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings