



Comparison of Adaptive Neuro-Fuzzy Inference System (ANFIS) and Machine Learning Algorithms for Electricity Production Forecasting

Elen Y.A. Rodríguez , Alexander A.R. Gamboa , Elias C.A. Rodríguez , Aneirson F. da Silva , Paloma M.S.R. Rizol , and Fernando A.S. Marins 

Abstract—Combined cycle power plants (CCPP) are popular in the energy sector for the production of electricity, and are the union of two thermodynamic cycles, corresponding to the steam turbine and the gas turbine. This paper presents the application of several machine learning (ML) techniques and the adaptive neuro-fuzzy inference system (ANFIS) to predict the hourly electricity production in a CCPP. The models were developed using 5-fold cross-validation with the collected features of temperature, exhaust pressure, relative humidity, ambient pressure, and electricity production per hour (the target feature). The hyperparameters of the tested models were optimized. The correlation and determination coefficients of the models were higher than 92%, showing a significant performance. The ANFIS ($r = 98\%$ e $R^2 = 95\%$) model shows the lowest values in the evaluated error metrics, compared to the other ML models. Finally, the results showed the effectiveness of ANFIS in predicting the hourly production of electricity in CCPP.

Index Terms—Electricity, Power generation, Fuzzy neural networks, Machine learning, Predictive models.

I. INTRODUÇÃO

O aumento do uso de equipamentos eletrônicos tem levado a uma necessidade crescente de fornecer energia elétrica confiável para atender à progressiva demanda de energia. Para a geração de eletricidade existem diferentes tipos de usinas, incluindo usinas tradicionais que utilizam combustíveis fósseis que produzem cerca de 80% da produção mundial de energia, e também usinas que utilizam fontes de energia renovável [1]. As usinas de ciclo combinado (CCPP - *Combined Cycle Power Plants*) são usinas tradicionais, compostas por uma combinação de turbina a gás e turbina a vapor, em que a energia térmica é transformada em eletricidade [1], [2].

As CCPP caracterizam-se por terem um baixo impacto ambiental em comparação com outras tecnologias de combustíveis fósseis, custos relativamente baixos e uma eficiência superior a 60% [2], [3]. No entanto, a previsão da produção de eletricidade tornou-se um problema relevante para reduzir custos operacionais, melhorar o aproveitamento da energia produzida [4], e maximizar os megawatts-hora disponíveis [5].

Na literatura, diversos trabalhos relacionados à previsão da produção de energia elétrica têm sido publicados, sendo observado um interesse crescente nos últimos anos na aplicação de

técnicas de aprendizado de máquina (ML - *Machine Learning*) [6], [7] e sistema de inferência neuro-fuzzy adaptativo (ANFIS - *Adaptive Neuro-Fuzzy Inference System*) [8].

Alguns dos trabalhos relevantes da literatura testaram diferentes algoritmos, como Tüfekci [5], que comparou a regressão linear (RL), *multilayer perceptron* (MLP), máquina de vetor de suporte (MVS), *K-nearest neighbors* e *REPTree* para a previsão de energia elétrica de uma CCPP utilizando variáveis termodinâmicas ambientais e validação cruzada *5x2-fold*, na qual o algoritmo *REPTree* superou os demais modelos com erro médio absoluto (*EMA*) de 2,818 e raiz do erro quadrático médio (*REQM*) de 3,787. Em [9] utilizaram a técnica de redução de dimensionalidade com os algoritmos de RL, LASSO, *ridge*, *elasticnet*, *random forest* e *gradient boost*, sendo obtido o menor erro com o algoritmo *gradient boost* (*REQM* = 6,084 e *EMA* = 4,742).

Wenyue *et al.* [10] realizaram uma abordagem diferente de [5], combinando a técnica *Granular Computing* com ANFIS e dividindo o conjunto de dados 50:50 (treino:teste), gerando um modelo com *REQM* = 4,223. Na sequência, Yeom e Kwak [11], semelhante do ANFIS, combinaram o raciocínio difuso de Takagi-Sugeno com o algoritmo de máquina de aprendizado extremo, dividindo o conjunto de dados em 60:40, na qual o melhor modelo foi desenvolvido com 100 épocas e *REQM* = 3,970. Em [12] foram utilizadas as redes neurais artificiais; nos experimentos o conjunto de dados foi dividido em 75:25 e foram comparados os algoritmos de Levenberg-Marquardt e regularização bayesiana, o melhor modelo foi gerado com a regularização bayesiana com 37 épocas e *REQM* = 5,003.

Além das variáveis termodinâmicas, outras características também foram consideradas para a previsão de energia elétrica, como em [13] que utilizaram um conjunto de dados com 24 variáveis relacionadas a turbinas a vapor e gás, e geradores de vapor com recuperação de calor; os algoritmos testados foram RL, LASSO, árvore de decisão e *random forest*, sendo obtido o menor erro com RL com *REQM* = 1,118 e *EMA* = 0,185. O autor de [14] usou os algoritmos de rede neural artificial, RL, MVS e ANFIS para estimar carga de eletricidade a longo e curto prazo usando variáveis ambientais, temporais e econômicas, os dados foram divididos em 70:30, na qual o melhor modelo para previsão de curto prazo foi a rede neural artificial (*REQM* = 7,670) e para longo prazo foi a MVS (*REQM* = 26,440). Já em [15] utilizaram ANFIS e variáveis de temperatura e pressão ambiente de uma turbina a gás de CCPP, obtendo um modelo com coeficiente

Elen Y.A. Rodríguez, Elias C.A. Rodríguez, Aneirson F. da Silva, Paloma M.S.R. Rizol, Fernando.A.S Marins, Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, Brasil.

Alexander A.R. Gamboa, Instituto Tecnológico de Aeronáutica, Brasil.
Corresponding author: Elen Y.A. Rodríguez (elen.aguirre@unesp.br)

de determinação (R^2) igual a 0,985.

Deste modo, a potência de uma usina depende principalmente de parâmetros ambientais, pois essa relação influencia no seu desempenho, dado que o aumento da temperatura ambiente gera uma diminuição na produção de energia [5], [15]. Portanto, é importante desenvolver um modelo preditivo com resultados confiáveis sobre a geração de eletricidade, considerando as variações existente nos parâmetros ambientais.

Neste contexto, o objetivo geral deste trabalho foi desenvolver um modelo de previsão para a produção de energia elétrica de uma CCPP. Os dados utilizados são de variáveis termodinâmicas e foram obtidos do repositório UCI [16], que foram coletados por meio de sensores quando a planta foi configurada para operar a plena carga. Além disso, os resultados referentes ao desempenho preditivo dos modelos gerados com as técnicas de ML e ANFIS foram comparados e analisados.

O artigo está estruturado da seguinte forma: a Seção II apresenta os materiais e métodos deste estudo. A Seção III descreve a avaliação e os resultados, detalhando os modelos ML e ANFIS aplicados. Por fim, a Seção IV apresenta as conclusões, seguidas das referências bibliográficas.

II. MATERIAL E MÉTODOS

A Fig. 1 apresenta o diagrama de fluxo do modelo de previsão para a produção de eletricidade, desde a obtenção do conjunto de dados até ao desenvolvimento do modelo preditivo e a obtenção dos resultados. Além disso, neste estudo, o modelo híbrido ANFIS foi desenvolvido no MATLAB® e os modelos de ML foram desenvolvidos no programa R v.4.0.3.



Fig. 1. Diagrama de fluxo do modelo de previsão para a produção de eletricidade.

A. Coleta do Banco de Dados

O conjunto de dados deste estudo foi coletado de uma CCPP [5], que tem a capacidade de geração de energia de 480 MW (ver Fig. 2)

A usina é composta por uma turbina a vapor ABB 160MW, dois geradores de vapor com recuperação de calor de dupla pressão (HRSG), e duas turbinas a gás ABB 13E2 160MW. A carga da turbina a vapor é sensível à pressão de vapor de exaustão, e a turbina a gás é sensível às condições ambientais, tais como a temperatura, pressão ambiente e umidade relativa, conforme a Fig. 2.

Os dados foram coletados quando a CCPP foi configurada para trabalhar a carga total, e coletaram-se 9.568 instâncias, correspondentes aos dados médios por hora recebidos por sensores.

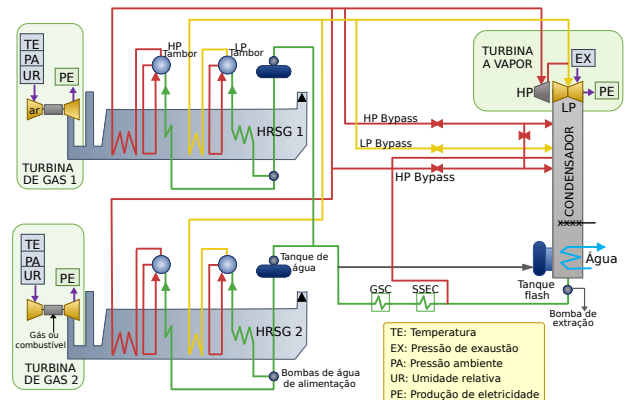


Fig. 2. Esquema da planta de ciclo combinado (CCPP).

O conjunto de dados utilizado é composto por quatro variáveis independentes (variáveis de entrada): temperatura (TE : 1,81 - 37,11 °C), pressão de exaustão (EX : 25,36 - 81,56 cmHg), pressão ambiente (PA : 992,89 - 1.033,30 mBar), umidade relativa (UR : 25,56 - 100,16%); uma variável dependente (variável de saída): produção de eletricidade por hora (PE : 420,26 - 495,76 MW) da usina. A Fig. 3 ilustra a série temporal dos dados coletados da energia elétrica produzida por hora, recebidos pelos sensores.

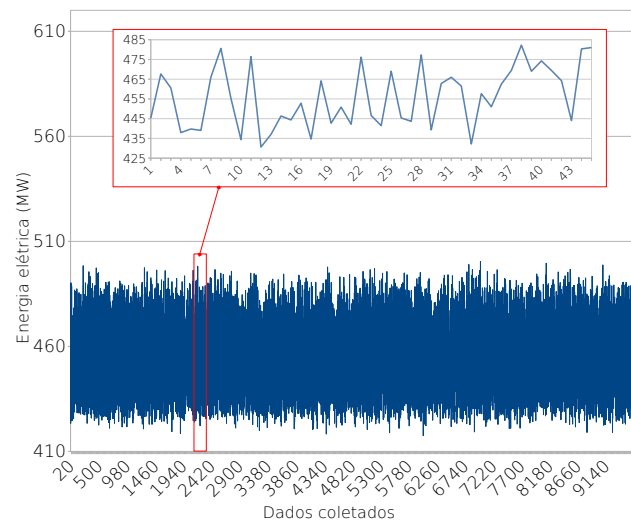


Fig. 3. Dados da energia elétrica produzida por hora.

B. Pré-processamento

A limpeza e transformação de dados é importante na etapa de pré-processamento para garantir a qualidade dos dados antes de ser usados para o treinamento do modelo preditivo. O conjunto de dados não apresentou dados faltantes, porém, os valores atípicos do conjunto de treino foram removidos. O método MIN-MAX foi utilizado para dimensionar os dados no intervalo [0 - 1], usando os valores mínimo e máximo da variável transformada, conforme Equação (1):

$$x'_{i,j} = \frac{x_{i,j} - x_{min,j}}{x_{max,j} - x_{min,j}} \quad (1)$$

sendo $x'_{i,j}$ o valor normalizado de $x_{i,j}$, que varia entre 0 e 1 [17], enquanto $x_{max,j}$ e $x_{min,j}$ são os valores máximo e mínimo, respectivamente, do atributo selecionado.

C. Desenvolvimento de Modelos

Neste estudo foram comparados cinco modelos, usando os algoritmos de: ANFIS, árvore modelo M5P, regressão linear, máquina de vetor de suporte e multilayer perceptron. Para minimizar o problema de sobreajuste (*overfitting*) dos modelos, o método de validação cruzada *5-fold* foi utilizado [5].

1) *Sistema de inferência neuro-fuzzy adaptativo*: O ANFIS é uma técnica de previsão não linear multivariada [4], que combina o melhor das redes neurais artificiais com o raciocínio difuso de Takagi-Sugeno do Sistema de Inferência Fuzzy (FIS - *Fuzzy Inference System*) [1], em um sistema híbrido [8]. O ANFIS combina as redes neurais com a abordagem de inferência Takagi-Sugeno usando as regras fuzzy de se-então (*if-then*), conforme ilustrado na Fig. 4.

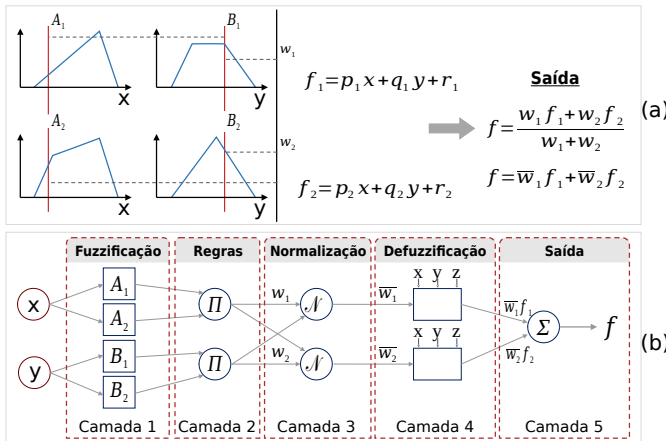


Fig. 4. (a) Raciocínio Fuzzy Tipo 3 (b) Arquitetura do sistema de inferência neuro-fuzzy adaptativo.

O ANFIS é composto de cinco camadas, cada um com uma finalidade específica, conforme mostrado na Fig. 4(b). As camadas são descritas a seguir:

Camada 1 – Fuzzificação: Os valores inseridos pelos nós de entrada x e y são convertidos em valores difusos por meio das funções de pertinência, ou seja o valor obtido da função de A específica o grau em que o valor do nó de entrada satisfaz o quantificador A_i , conforme está em (2):

$$\mu_{A_i}(x) = f(x) \wedge \mu_{B_i}(y) = f(y) \quad , i = 1, 2 \quad (2)$$

Camada 2 – Regras: Recebem as informações de seus respectivos neurônios difusos conectados, e a relação entre os valores de entrada é estabelecida calculando a intensidade da regra w_i , conforme está em (3):

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \quad , i = 1, 2 \quad (3)$$

Camada 3 – Normalização: As saídas desta camada são chamadas de força de disparo normalizado, e em \bar{w}_i é cal-

culada a intensidade da regra normalizada, conforme está em (4):

$$\bar{w}_i = \frac{w_i}{w_1 + w_2} \quad , i = 1, 2 \quad (4)$$

Camada 4 – Desfuzzificação: O valor ponderado de cada regra pode ser calculado por (5):

$$\bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad , i = 1, 2 \quad (5)$$

sendo que p_i , q_i e r_i são os parâmetros da camada de defuzzificação que são calculados por meio da fase de treinamento.

Camada 5 – Saída: Esta camada soma todas as saídas do neurônio da desfuzzificação, até atingir o valor final ANFIS (f), podendo ser calculada por (6):

$$f = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (6)$$

Além disso, as redes neurais artificiais são usadas para reduzir os erros estimados, ajustando os parâmetros da função de pertinência, por meio de um método híbrido que combina o método de *backpropagation* e mínimos quadrados [8].

2) *Regressão linear*: A RL é um método usado tanto na estatística quanto no ML, e tem o objetivo de investigar a influência que as variáveis independentes (regressoras) têm sobre a variável de resposta, a fim de modelar a relação entre as variáveis [6], [7]. Matematicamente, a RL é representada por (7):

$$y = \beta_0 + \sum_{i=1}^d \beta_i x_i + \epsilon \quad , i = 1, 2, \dots, d \quad (7)$$

sendo β_i os coeficientes de regressão e $d \geq 1$ é a dimensão do vetor de variáveis regressoras. Além disso, os mínimos quadrados é o método mais utilizado na RL, que consiste na busca pelo valor mínimo da soma residual dos quadrados (SRQ) [7]. O SRQ é calculado por (8):

$$SRQ = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

O erro quadrático médio (EQM) também é amplamente usado pelo RL, e representa uma função de perda que avalia a qualidade da função de previsão [6], [7]. O EQM é calculado por (9):

$$EQM = \frac{1}{n} \times SRQ \quad (9)$$

3) *Árvore modelo M5P*: A árvore de decisão é uma técnica muito popular em ML, por sua simplicidade e eficiência, além de ter uma estrutura simples e de fácil interpretação. A árvore modelo M5P permite criar modelos muito eficientes, com uma estrutura em forma de árvore, que possui modelos lineares em seus nós de previsão [18].

O critério de divisão no M5P, é baseado principalmente na seleção da variável x_i que maximiza a redução do erro, por meio do desvio padrão dos valores y que chegam a um nó específico. A redução do desvio padrão (RDP) é calculado por (10):

$$RDP = sd(T) - \sum_i \frac{T_i}{T} \times sd(T_i) \quad (10)$$

sendo T o conjunto de instâncias que atingem o nó antes da divisão, T_i é um subconjunto de dados de um nó com base em uma variável escolhida que maximiza a redução do erro esperado, $sd(T)$ é o desvio padrão do conjunto T e $sd(T_i)$ é o desvio padrão dos valores alvo de T_i , considerado como uma medida de erro [19]. A divisão da árvore termina quando a redução do erro é inferior a 5% do desvio padrão de todo o conjunto de dados, e em cada folha da árvore é construído um modelo de RL.

4) *Máquina de vetor de suporte*: A MVS mapeia os dados de entrada em um espaço não linear de alta dimensão, e pode ser formulado por (11):

$$f(x) = \langle w, x \rangle + b, \quad w \in X, b \in R \quad (11)$$

sendo X o espaço vetorial de dados de entrada, $\langle \cdot, \cdot \rangle$ denota o produto escalar em X , w é um peso e b é um viés [20].

Assim, busca-se estimar os valores de b e um pequeno peso w , sendo que o problema de regressão é transformado em um problema de minimização, e pode ser escrito como um problema de otimização convexa, conforme está em (12):

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i^- + \xi_i^+) \quad (12)$$

$$\text{Sujeito a } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi_i^- \\ \langle w, x_i \rangle + b \leq \epsilon + \xi_i^+ \\ \xi_i^-, \xi_i^+ \geq 0, \quad i = 1, \dots, l \end{cases}$$

sendo ξ_i^- e ξ_i^+ variáveis de folga, $C > 0$ é constante e representa o fator de regularização que calcula o *trade-off* entre o modelo e a quantidade de desvios maiores que ϵ são permitidos [20].

5) *Multilayer perceptron*: O MLP é o tipo de rede neural artificial mais comum, cuja estrutura de rede consiste em uma camada de entrada, uma ou mais camadas ocultas, e um nó na camada de saída [18], [21]. A saída da rede MLP pode ser expressa conforme está em (13):

$$f(x) = v_o^T z = \sum_{h=1}^H v_{ih} z_h + v_{i0} \quad (13)$$

sendo v_{i0} os pesos dos vieses e z_h os valores dos neurônios da camada oculta, na qual é aplicada a função $g(\cdot)$ sigmóide não linear dada por (14):

$$z_h = g(w_h^T x) = \frac{1}{1 + \exp \left[- \left(\sum_{j=1}^d w_{hj} x_j + w_{h0} \right) \right]} \quad (14)$$

sendo $h = 1, \dots, H$, e $w = [w_0, w_1, \dots, w_d]^T$ são os pesos sinápticos, que são gerados a partir das entradas x_j na fase de treinamento [22].

Por outro lado, o desempenho dos modelos testados depende da configuração de seus hiperparâmetros. Os hiperparâmetros dos modelos foram otimizados utilizando a função *expand.grid* com diferentes valores de configuração.

Assim, para avaliar a precisão do modelo preditivo, foi utilizado o *EMA*, *REQM*, erro relativo absoluto (*ERA*), erro quadrado relativo (*EQR*), coeficiente de correlação (r) e

R^2 , dado que, embora existam várias métricas de avaliação, ainda não foi estabelecida uma métrica padrão para medir a eficiência de um modelo de previsão [23]. A seguir, em (15) - (20), são apresentadas as formulas das métricas adotadas.

$$EMA = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (15)$$

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (16)$$

$$ERA = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{\sum_{i=1}^n |\bar{y}_i - y_i|} \quad (17)$$

$$EQR = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2}} \quad (18)$$

$$r = \frac{n(\sum y_i \hat{y}_i) - (\sum y_i)(\sum \hat{y}_i)}{\sqrt{[n \sum y_i^2 - (\sum y_i)^2][n \sum \hat{y}_i^2 - (\sum \hat{y}_i)^2]}} \quad (19)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (20)$$

sendo que, em cada equação o \hat{y}_i é o valor previsto pelo modelo para a observação $i = 1, 2, \dots, n$, y_i é o valor real da observação i , e a \bar{y} é a média de y . Observe-se que, os menores valores das medidas de erro e os maiores valores em r e R^2 indicam que o modelo selecionado tem o melhor desempenho na previsão [4].

Adicionalmente, o teste-t pareado foi utilizado para comparar as médias, com base na hipótese nula ($H_0 : \mu_x = \mu_y$) e intervalo de confiança de 95%, dos dados reais e os resultados obtidos com os modelos de previsão, a fim de comparar as previsões e identificar o modelo com melhor desempenho [24].

III. AVALIAÇÃO E RESULTADOS

Nesta seção são apresentados os resultados obtidos com os modelos de previsão para a produção de energia elétrica por hora, utilizando-se o método ANFIS e os quatro modelos de ML selecionados.

Dos modelos MVS, M5P, MLP, RL e ANFIS, desenvolvidos e otimizados por validação cruzada 5-fold no processo de desenvolvimento de modelos deste trabalho têm-se:

- Da MVS, os hiperparâmetros foram: *kernel* RBF (*radial basis function*), parâmetro de regularização $C = 100$, grau = 3, $\epsilon = 0.15$ e $\xi = 1$.
- Da M5P seus hiperparâmetros otimizados foram: profundidade máxima igual a 5 e mínimo 6 instância em cada nó interno antes de ser dividido.
- No MLP, seus hiperparâmetros foram: quatro neurônios na camada de entrada, um neurônio na camada de saída, 50 neurônios na camada oculta, com função de ativação da camada oculta ReLU (unidade linear retificada), otimizador de peso baseado no gradiente estocástico, taxa de aprendizado constante com valor inicial igual a 0,001 e decaimento de peso igual a 0,9.
- A função obtida com o modelo de RL foi $y = \beta_0 + \sum_{i=1}^4 \beta_i x_i$, com $\beta_0 = 502,365$, $\beta_1 = -69,867$, $\beta_2 = -13,085$, $\beta_3 = 2,552$ e $\beta_4 = -11,503$.

- No ANFIS, a configuração de seus hiperparâmetros foi: particionamento em grade (*grid partition*), duas funções de pertinência para as variáveis *TE* e *PA*, quatro funções de pertinência para as variáveis *EX* e *UR* (conforme Tabela I). Variável de saída do tipo linear, e foi escolhido o método híbrido que combina o método de *backpropagation* e mínimos quadrados, uma tolerância de erro mais próximo a zero, e número épocas de treinamento igual a 48.

TABELA I
PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA.

Variáveis	V.Ling. ^a	Tipo	Parâmetros
<i>TE</i>	In1mf1	gbellmf ^b	[0, 484; 1, 999; -0, 014]
	In1mf2	gbellmf	[0, 468; 2, 003; 1, 023]
<i>EX</i>	In2mf1	gaussmf ^c	[0, 142; 0, 031]
	In2mf2	gaussmf	[0, 093; 0, 484]
	In2mf3	gaussmf	[0, 108; 0, 584]
	In2mf4	gaussmf	[0, 095; 0, 986]
<i>PA</i>	In3mf1	gaussmf	[0, 271; -0, 044]
	In3mf2	gaussmf	[0, 199; 1, 069]
<i>UR</i>	In4mf1	gbellmf	[0, 280; 1, 987; 0, 042]
	In4mf2	gbellmf	[0, 103; 2, 004; 0, 393]
	In4mf3	gbellmf	[0, 137; 2, 00; 0, 759]
	In4mf4	gbellmf	[0, 187; 2, 001; 0, 978]

(a)V.Ling: Variável Lingüística, (b)gbellmf: Função de Pertinência Bell-shaped, (c)gaussmf: Função de Pertinência Gaussian.

Na sequência, a Fig. 5 ilustra o gráfico de dispersão de cada modelo de previsão testado, dos valores reais e os valores previstos, além de mostrar o valor do R^2 de cada modelo.

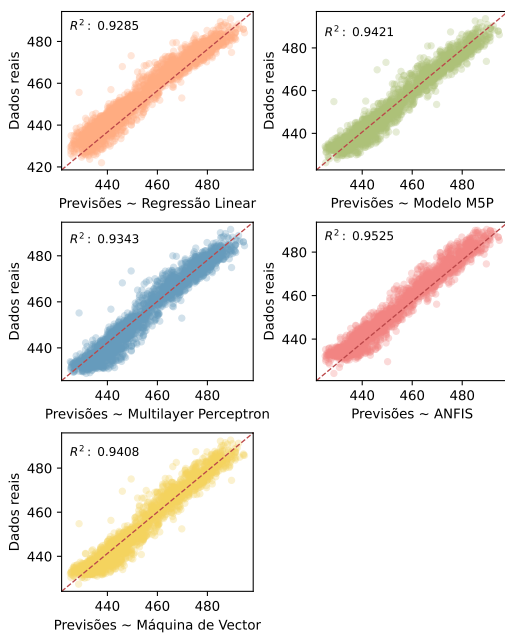


Fig. 5. Gráfico de dispersão dos dados reais e previstos.

Ainda na Fig. 5 pode-se observar que todos os modelos de previsão provaram ser eficientes ($R^2 > 92\%$), porém,

visualmente o ANFIS apresentou um melhor comportamento na comparação entre os valores previsto e os valores reais, superando os demais modelos testados.

A Fig. 6 apresenta os valores médios das métricas estatísticas de *EMA*, *REQM*, *ERA*, *EQR*, *r* e R^2 , utilizadas para medir a taxa de erro das previsões e analisar o desempenho dos modelos de MVS ($EMA = 3,172$, $REQM = 4,149$, $ERA = 0,214$, $EQR = 0,243$, $r = 0,970$ e $R^2 = 0,941$), M5P ($EMA = 3,150$, $REQM = 4,104$, $ERA = 0,212$, $EQR = 0,241$, $r = 0,971$ e $R^2 = 0,942$), MLP ($EMA = 3,430$, $REQM = 4,369$, $ERA = 0,231$, $EQR = 0,256$, $r = 0,967$ e $R^2 = 0,934$), RL ($EMA = 3,629$, $REQM = 4,561$, $ERA = 0,245$, $EQR = 0,267$, $r = 0,964$ e $R^2 = 0,928$) e ANFIS ($EMA = 2,912$, $REQM = 3,717$, $ERA = 0,196$, $EQR = 0,218$, $r = 0,976$ e $R^2 = 0,953$).

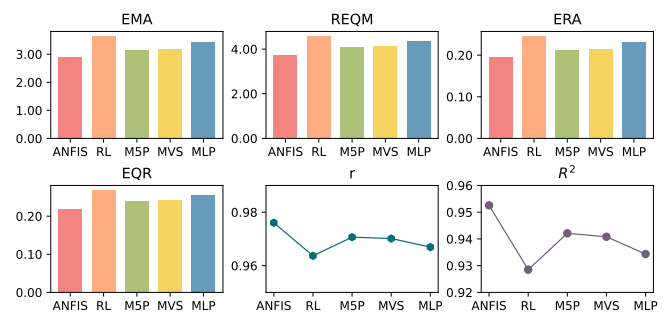


Fig. 6. Resultados das métricas de desempenho.

Na Fig. 6, constata-se que todos os modelos avaliados têm um coeficiente de correlação superior a 96% ($r > 0,96$). Além disso, os modelos MVS, M5P e ANFIS mostraram-se ser os mais eficientes, com valores de R^2 igual a 0,941 0,942 e 0,952, respectivamente. O modelo ANFIS superou os demais modelos, pois seus coeficientes de correlação e determinação foram superiores aos dos outros modelos avaliados, além de ter os menores valores nas métricas de erro de *EMA*, *REQM*, *ERA* e *EQR*. Na Fig. 7 apresenta-se o gráfico de 35 erros de predição selecionados aleatoriamente dos modelos MVS, M5P e ANFIS.

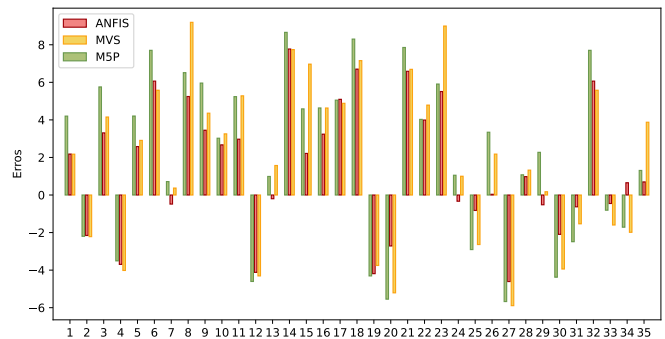


Fig. 7. Gráfico dos erros dos modelos MVS, M5P e ANFIS.

Os resultados apresentados na Fig. 7 evidenciam que os valores obtidos com o modelo ANFIS apresentam uma maior aproximação com os valores reais quando comparado com os modelos MVS e M5P.

Um teste-t pareado foi realizado para comparar estatisticamente o desempenho de cada modelo, conforme mostrado na Tabela II, principalmente do MVS (média = 454.437 e desvio padrão = 16.789), M5P (média = 454.321 e desvio padrão = 16.624) e ANFIS (média = 454.253 e desvio padrão = 16.789), sendo que a média dos valores reais foi 454,158.

TABELA II
RESULTADO DO TESTE-T PAREADO.

Modelo	valor-t	$t_{critico}$	p-valor	Observação
M5P	1,687	1,961	0,092	Aceitamos a H_0
ANFIS	1,108	1,961	0,268	Aceitamos a H_0
MVS	2,876	1,961	0,004	Rejeitar a H_0

(*) Hipótese nula ($H_0 : \mu_x = \mu_y$). Valores em negrito indicam o melhor valor.

Os resultados do teste-t pareado (Tabela II) revelam $valor-t < t_{critico}$ e $p\text{-valor} > 0,05$ para os modelos ANFIS e M5P, por conseguinte, a hipótese nula é aceita, podendo-se dizer que há semelhança entre as médias analisadas dos valores reais e os modelos ANFIS e M5P. Porém, o $valor-t$ do ANFIS (1,108) ainda é menor que o $valor-t$ do modelo M5P (1,687), indicando que as previsões do modelo ANFIS possuem maior similaridade com os valores reais, e melhor desempenho preditivo que o M5P. Portanto, esses valores reforçam os resultados obtidos com as métricas de desempenho, do que o modelo ANFIS apresentou as melhores previsões.

Finalmente, na Fig. 8 apresenta-se o gráfico de linhas de cem dados reais, a fim de permitir sua comparação com os dados previstos. Este gráfico evidencia que existe considerável semelhança entre os dados reais e os dados previstos pelo modelo ANFIS.

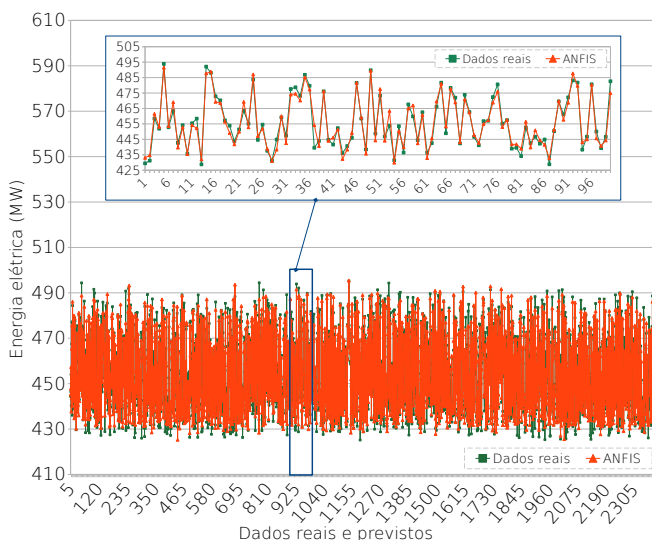


Fig. 8. Gráfico de linha de valores reais e previstos com ANFIS.

O modelo ANFIS ($REQM = 3,717$) apresentou um desempenho melhor que os demais modelos testados neste trabalho, além de ser melhor que os demais modelos relacionados na literatura, como *gradient boost* ($REQM = 6,084$) [9], redes neurais artificiais ($REQM = 5,003$) [12], *Granular Computing* com ANFIS ($REQM = 4,223$) [10], Takagi-Sugeno

com máquina de aprendizado extremo ($REQM = 3,970$) [11] e *REPTree* ($REQM = 3,787$) [5]. Assim, mostrou-se que as regras fuzzy do modelo ANFIS desempenham um papel importante no cálculo de valores contínuos, como neste caso, prever a produção de energia elétrica por hora de uma CCPP quando é configurada para funcionar a carga completa.

IV. CONCLUSÕES

O rápido avanço da tecnologia e o armazenamento de dados de diferentes transações ou processos possibilitaram o desenvolvimento de modelos inteligentes que podem processar grandes quantidades de dados e identificar padrões entre os dados para fazer previsões com rapidez e precisão.

A estimativa tradicional da energia elétrica gerada é calculada por meio de leis termodinâmicas e equações semi-empíricas que demandam maior tempo e esforço computacional para obter os valores produzidos [5], [9]. Neste trabalho foi apresentada uma alternativa para estimar a energia elétrica produzida por hora de uma CCPP, utilizando variáveis termodinâmicas ambientais e aplicando técnicas de ML e ANFIS, para obter resultados confiáveis no menor tempo.

Os resultados mostraram que o modelo desenvolvido com o ANFIS mostrou-se ser o melhor em relação dos outros modelos usados neste trabalho, mesmo com outros modelos testados na literatura [5], [9]–[12]. Os valores obtidos nas métricas de desempenho do ANFIS foram $R^2 = 0,95$, $r = 0,98$, $EMA = 2,91$, $REQM = 3,72$, $ERA = 0,20$ e $EQR = 0,22$.

O modelo desenvolvido neste trabalho limita-se à configuração e condições operacionais da CCPP de 480 MW, e à qualidade dos dados coletados pelos sensores (incerteza de medição) do CCPP. Por fim, como trabalhos futuros, pretende-se testar outras técnicas de ML para serem comparadas com o modelo ANFIS, dado que a utilização do ANFIS na previsão da produção de energia elétrica ainda é incipiente. Além disso, outros conjuntos de dados relacionados a outros ciclos termodinâmicos serão testados e avaliados.

AGRADECIMENTOS

Este estudo foi parcialmente financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Código de Financiamento 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) processo No. 304197/2021-1, 303090/2021-9, 381252/2020-5, e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) processo No. 2016/10274-9.

REFERENCES

- [1] M. Karacor, A. Uysal, H. Mamur, G. Sen, M. Nil, M. Z. Bilgin, H. DoGan, and C. Sahin, "Life performance prediction of natural gas combined cycle power plant with intelligent algorithms," *Sustainable Energy Technologies and Assessments*, vol. 47, p. 101398, 2021.
- [2] J. M. Alemany, D. Moitre, and F. Magnago, "Power system reliability considering combined cycle plants," *IEEE Latin America Transactions*, vol. 8, no. 5, pp. 547–556, 2010.
- [3] J. Kotowicz and M. Brzeczek, "Analysis of increasing efficiency of modern combined cycle power plant: A case study," *Energy*, vol. 153, pp. 90–99, 2018.

- [4] W. Dong, Q. Yang, and X. Fang, "Multi-step ahead wind power generation prediction based on hybrid machine learning techniques," *Energies*, vol. 11, no. 8, 2018.
- [5] P. Tüfekci, "Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods," *International Journal of Electrical Power and Energy Systems*, vol. 60, pp. 126–140, 2014.
- [6] S. Marsland, *Machine Learning: An Algorithmic Perspective*. 6000 Broken Sound Parkway NW, Suite 300, Boca Raton: Taylor and Francis Group, LLC, 2015.
- [7] R. Bali, D. Sarkar, B. Lantz, and C. Lesmeister, *R: Unleash Machine Learning Techniques*. 35 Livery Street, Birmingham B3 2PB, UK: Packt Publishing, Ltd, 2016.
- [8] J.-S. Jang, "Anfis: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [9] N. S. Santarisi and F. S. S., "Prediction of combined cycle power plant electrical output power using machine learning regression algorithms," *Eastern-European Journal of Enterprise Technologies*, vol. 6, no. 8, 2021.
- [10] W. Sun, J. Zhang, and R. Wang, "Predicting electrical power output by using granular computing based neuro-fuzzy modeling method," in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, pp. 2865–2870, 2015.
- [11] C.-U. Yeom and K.-C. Kwak, "A design of tsf-based elm for prediction of electrical power in combined cycle power plant," pp. 226–229, 10 2018.
- [12] E. Elfaki and A. H. A. Hassan, "Prediction of electrical output power of combined cycle power plant using regression ann model," *Journal of Power and Energy Engineering*, vol. 6, pp. 17–38, 12 2018.
- [13] M. G. Rabby Shuvo, N. Sultana, L. Motin, and M. R. Islam, "Prediction of hourly total energy in combined cycle power plant using machine learning techniques," in *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, pp. 170–175, 2021.
- [14] D. Solyali, "A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in cyprus," *Sustainability*, vol. 12, no. 9, 2020.
- [15] T. K. Ibrahim, R. M. Mohammed, and F. Basrawi, "Statistical analysis and optimum performance of the gas turbine power plant," *International Journal of Automotive and Mechanical Engineering*, vol. 13, no. 1, pp. 3215–3225, 2016.
- [16] UCI, "Machine learning repository," Irvine, CA: University of California, School of Information and Computer Science [Online]. Available from: <http://archive.ics.uci.edu/ml>, 2022.
- [17] M. M. Suarez-Alvarez, D.-T. Pham, M. Y. Prostov, and Y. I. Prostov, "Statistical approach to normalization of feature vectors and clustering of mixed datasets," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 468, no. 2145, pp. 2630–2651, 2012.
- [18] L. K. Moayedi, Hossein ADN Foong and H. Nguyen, "Soft computing method for predicting pressure drop reduction in crude oil pipelines based on machine learning methods," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 42, no. 11, 2020.
- [19] A. Etemad-Shahidi and M. Taghipour, "Predicting longitudinal dispersion coefficient in natural streams using m5p model tree," *Journal of Hydraulic Engineering*, vol. 138, no. 6, pp. 542–554, 2012.
- [20] M. Nait Amar and N. Zeraibi, "Application of hybrid support vector regression artificial bee colony for prediction of mmp in co2-eor process," *Petroleum*, vol. 6, no. 4, pp. 415–422, 2020.
- [21] B. Kaczorowski, M. Kleina, M. Augusto Mendes Marques, and W. de Assis Silva, "Artificial intelligence and the multivariate approach in predictive analysis of the small cap index of the brazilian stock exchange," *IEEE Latin America Transactions*, vol. 19, no. 11, pp. 1924–1932, 2021.
- [22] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, Massachusetts, London, England: The MIT Press, 2004.
- [23] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," *Applied Energy*, vol. 87, no. 7, pp. 2313–2320, 2010.
- [24] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*. John Wiley and Sons, 6th ed., 2013.



Elen Yanina Aguirre Rodríguez Graduada em Engenharia de Sistemas pela Universidade Nacional de Trujillo, Perú, mestrado em Engenharia Produção na Universidade Estadual Paulista (UNESP). Está atualmente cursando Doutorado em Engenharia de Mecânica (Área de Produção) na UNESP. Possui experiência na área de Engenharia de Produção e Ciência da Computação com ênfase em otimização, Machine learning, sistemas de informação, mineração de dados e reconhecimento de padrões.



Alexander Alberto Rodríguez Gamboa Graduado em Engenharia Mecânica pela Universidade Nacional de Trujillo, Perú, mestrado, doutorado em Engenharia Mecânica (Área de Energia) na Universidade Estadual Paulista (UNESP) e pós-doutorado no Instituto Tecnológico de Aeronáutica. Possui experiência na área de Combustão de combustíveis líquidos não convencionais como o óleo pirolítico de pneus.



Elias Carlos Aguirre Rodríguez Graduado em Ciências Estatísticas pela Universidade Nacional de Trujillo, Perú. Está atualmente cursando o Mestrado em Engenharia (Área de Produção) na Universidade Estadual Paulista (UNESP). Possui experiência na área de Probabilidade e Estatística com ênfase em estatística aplicada, análise multivariada, demografia, ciência dos dados, machine learning, técnicas de amostragem e análise de dados.



Aneirson Francisco da Silva Graduado em administração pelo Centro Universitário do Triângulo, mestrado em Engenharia de Produção pela Universidade Federal de Itajubá, doutorado em Engenharia Mecânica pela Universidade Estadual Paulista (UNESP) e pós-doutorado na Universidade de São Paulo. Tem experiência na área de Engenharia de Produção, com ênfase em Pesquisa Operacional nos temas de Programação Linear, Não-Linear, Mista e Dinâmica, Design of Experiments, Teoria dos Grafos, Métodos Multicriteriais, Goal Programming e Otimização

Multiobjetivo.



Paloma Maria Silva Rocha Rizol Graduada em Engenharia Elétrica pela Universidade Estadual Paulista (UNESP), mestrado e doutorado em Engenharia Eletrônica e Computação pelo Instituto Tecnológico de Aeronáutica. Atualmente é professora assistente da UNESP. Tem experiência na área de Eletrônica, com ênfase em Projeto de Circuitos Integrados Analógicos, atuando principalmente nos temas de lógica fuzzy, fuzzy tipo-2 e fuzzy hardware.



Fernando Augusto Silva Marins Graduado em Engenharia Mecânica pela Universidade Estadual Paulista (UNESP), mestrado em Pesquisa Operacional pelo Instituto Tecnológico de Aeronáutica, doutorado em Engenharia Elétrica pela Universidade Estadual de Campinas e pós-doutorado na Brunel University - UK. Atualmente é professor titular no Departamento de Produção da UNESP. Tem experiência na área de Engenharia de Produção, com ênfase em Pesquisa Operacional e Logística, atuando principalmente nos temas de Logística e Supply Chain Management, Pesquisa Operacional, DEA, AHP e Simulação.