

# Um Sistema Baseado Em Combinação de Modelos para Previsão de Velocidade do Vento

*A Model Combination Based System for Wind Speed Forecast*

**Antonio Assunção<sup>1,3</sup>**

 [orcid.org/0000-0001-5727-2427](https://orcid.org/0000-0001-5727-2427)

**Eduardo Vasconcelos<sup>1,3</sup>**

 [orcid.org/0000-0001-5727-2427](https://orcid.org/0000-0001-5727-2427)

**Paulo S. G. de Mattos Neto<sup>2</sup>**

 [orcid.org/0000-0002-2396-7973](https://orcid.org/0000-0002-2396-7973)

<sup>1</sup> Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: ([ada.ecav@ecomppoli.br](mailto:ada.ecav@ecomppoli.br))

<sup>2</sup> Centro de Informática, Universidade Federal de Pernambuco, Recife, Brasil. E-mail: [psgmn@cin.ufpe.br](mailto:psgmn@cin.ufpe.br).

<sup>3</sup> Programa de Pós-graduação em Inteligência Artificial, Escola Politécnica de Pernambuco, Brasil.

**DOI: 10.25286/rep.v7i2.2206**

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Antonio Assunção; Eduardo Vasconcelos; Paulo S. G. de Mattos Neto. Um Sistema Baseado Em Combinação de Modelos para Previsão de Velocidade do Vento. Revista de Engenharia e Pesquisa Aplicada, v.7, n. 2, p. 1-11, 2022.

## RESUMO

O mundo vem, há anos, sofrendo efeitos decorrentes das mudanças climáticas. O aumento dos incêndios, inundações, secas e a diminuição do tamanho da calota polar são exemplos dessas mudanças. Nações em todo mundo têm, de forma recorrente, buscado formas sustentáveis de geração de energia. As fontes de energia de origem renovável, como o Sol, água e vento têm, cada vez mais, sido usadas em detrimento dos combustíveis fósseis. A energia eólica vem crescendo substancialmente devido à sua capacidade de geração em larga escala, mas, no entanto, não é das mais fáceis de ser implantada. Previsões consistentes da velocidade do vento, são essenciais para evitar perdas econômicas e aumentar a eficiência operacional desses sistemas, tornando cada projeto um desafio à parte. A análise e previsão do comportamento não estacionário e os padrões não lineares dos ventos é um dos problemas de pesquisa mais relevantes da atualidade. A literatura apresenta diversas abordagens, sendo parte delas baseadas no uso individual de sistemas lineares ou outras técnicas de Inteligência Artificial (IA). Este artigo apresenta o uso de um sistema híbrido para a previsão de curto prazo da velocidade do vento.

**PALAVRAS-CHAVE:** Velocidade do vento; Previsão; Sistemas Híbridos; Redes Neurais Artificiais;

## ABSTRACT

The world has, for years, been suffering from the effects of climate change. Increased fires, floods, droughts and the shrinking size of the ice cap are examples of these changes. Nations around the world have recurrently pursued sustainable forms of energy generation. Renewable energy sources, such as the sun, water and wind, have increasingly been used at the expense of fossil fuels. Wind energy has been growing substantially due to its large-scale generation capacity, however, it is not the easiest to implement. Consistent wind speed predictions are essential to avoid economic losses and increase the operational efficiency of these systems, making each project a challenge in its own right. The analysis and prediction of non-stationary behavior and non-linear wind patterns is one of the most relevant research problems today. The literature presents several approaches, part of which are based on the individual use of linear systems or other Artificial Intelligence (AI) techniques. This article presents the use of a hybrid system for short-term wind speed prediction.

**KEYWORDS:** Wind speed, Forecasting; Hybrid Systems; Artificial Neural Networks;

### 1 INTRODUÇÃO

A geração de energia a partir de fontes renováveis e limpas é imprescindível para a sustentabilidade do planeta. O uso de fontes renováveis, além de aumentar a confiabilidade dos sistemas de geração de energia, reduz as emissões de carbono [1]. Por este motivo, sistemas baseados em fontes sustentáveis são preferidos pelas empresas do setor elétrico [2]. O vento é uma das principais fontes de energia limpa, uma vez que é um recurso inesgotável e está disponível de forma abundante. A otimização do processo de conversão da energia eólica, cada vez mais, torna-se um processo especializado.

O vento pode ser considerado uma variável estocástica e, dessa forma, a correta previsão de sua velocidade pode beneficiar sobremaneira os operadores das plantas de geração de energia eólica. Por ter um comportamento não estacionário e apresentar padrões não lineares, as tomadas de decisão relacionadas à operação e gerenciamento desses sistemas se torna mais desafiadora, assim como a criação de abordagens que objetivem prever seu comportamento [2].

O comportamento intermitente do vento é o maior desafio à sua implementação como uma fonte confiável e autônoma de energia elétrica. A previsão de sua velocidade envolve erros, de forma similar ao que ocorre no processamento da relação entre a velocidade e potência previstas e a velocidade e potência reais, desempenhando essas estimativas um papel crucial na seleção de um algoritmo de previsão apropriado [3]. A geração de energia eólica está relacionada, diretamente, à velocidade do vento. Desta forma, para que a velocidade do vento possa ser usada em larga escala em sistemas de geração de energia, deve-se, antes, obter respostas para alguns problemas, tais como: qualidade da energia gerada, operações de rede em tempo real, padrões de interconexão com a rede de distribuição, capacidade do sistema, estabilidade e confiabilidade do sistema de transmissão.

Obter a melhor relação entre os valores previstos e a velocidade real do vento traz vantagens econômicas e técnicas significativas. Em [4], é proposto um método probabilístico para calcular as perdas de energia associadas a erros de previsão para geradores eólicos. Neste trabalho, os estudos de caso mostram que os custos de previsão de erro podem chegar a até 10% da receita total da venda da energia gerada [2]. De forma complementar, a

correta previsão da velocidade do vento pode ajudar a desenvolver o bom funcionamento dos mercados de hora à frente ou do dia seguinte.

As atividades de previsão da velocidade do vento podem ser classificadas, em termos da escala de tempo, como previsão de muito curto, curto, médio, longo e muito longo prazo [5]. Cada uma dessas escalas tem aplicações específicas, como, por exemplo, o uso das previsões de curto prazo em atividades operacionais, enquanto as previsões de mais longo prazo são empregadas no planejamento e estudos de viabilidade econômica.

Os modelos de previsão de velocidade do vento são, normalmente, categorizados em três classes principais: métodos de previsão físicos, métodos de previsão baseados em estatísticas ou Séries Temporais, do inglês, *Time Series*, TS, e métodos baseados em Inteligência Artificial, do inglês, *Artificial Intelligence*, AI [6]. No primeiro caso, cuja previsão é baseada no clima, a precisão da previsão do vento está diretamente relacionada à topologia do terreno onde as turbinas eólicas estão instaladas. Esses modelos são também chamados de Previsão Numérica do Clima, do inglês, *Numerical Weather Prediction*, NWP, e, a depender do relevo, fatores como a medição da velocidade do vento a determinada altura do solo, a temperatura e pressão do ar ambiente são variáveis que devem ser consideradas.

Os métodos estatísticos dependem, unicamente, das medições anteriores da velocidade do vento para prever os valores futuros. Os métodos baseados em AI vêm, normalmente, sendo empregados para previsão de velocidade do vento de curto e longo prazo [7]. Esses métodos são, inerentemente, não lineares e orientados por dados, sendo capazes de lidar com padrões temporais complexos. Apesar de apresentarem vantagens em relações aos modelos lineares, técnicas como Regressão de Vetor de Suporte, do inglês, *Support Vector Regression*, SVR, Redes Neurais Artificiais, do inglês, *Artificial Neural Networks*, ANN, e Perceptron Multicamada, do inglês, *Multilayer Perceptron*, MLP, quando usadas em séries temporais reais, podem produzir resultados não tão precisos, devido aos problemas de *overfitting* e *underfitting* e, também, a falhas na especificação do modelo.

Levando-se em conta as características dos métodos apresentados, incluindo-se aí suas limitações, busca-se, através da combinação de técnicas, obter um desempenho melhor do que

cada técnica individualmente, reduzindo, dessa forma, o risco de especificações incorretas do modelo. Deve-se, no entanto, considerar que para previsões mais assertivas, a diversidade dos modelos deve ser garantida. As séries temporais e séries residuais são usadas como base para os sistemas híbridos baseados em modelagem residual [8], tendo sido empregados com sucesso para previsões de velocidade do vento de curto, médio e longo prazo [9]. Esses sistemas híbridos superam a limitação dos modelos únicos, lineares e não lineares, quando usados individualmente. Sua implementação envolve, basicamente, uma técnica estatística linear, usada na modelagem das séries temporais, e uma técnica de Inteligência Artificial não linear, usada na previsão da série de resíduos do modelo linear. Obtidas as saídas desses dois modelos, ambas são utilizadas para produzir a previsão da velocidade do vento.

Neste artigo é utilizado um modelo híbrido, apresentado inicialmente em [8], composto por três etapas: modelagem de uma série temporal usando um preditor linear, uma técnica não linear, utilizada na modelagem dos resíduos e, na etapa final, é utilizado um modelo de AI na combinação das previsões das séries temporais e das séries de erros, etapa esta utilizada na maximização da precisão do sistema. O modelo híbrido é utilizado em três conjuntos de dados mensais de velocidade do vento de diferentes estados da região Nordeste do Brasil, apresentando desempenho melhor do que os modelos tradicionais simples e combinados.

Este artigo está organizado da seguinte forma: a Seção II estabelece os métodos de previsão da velocidade do vento; A seção III apresenta o modelo híbrido utilizado; A Seção IV mostra os resultados experimentais alcançados; A Seção V discute os resultados obtidos e a Seção VI apresenta os caminhos para pesquisas futuras.

## 2 MÉTODOS DE PREVISÃO

Os modelos de Previsão da Velocidade do Vento, do inglês, *Wind Speed Forecasting*, WSF, encontram-se, normalmente, relacionados ao mercado de eletricidade, projetos de geração de energia eólica, suporte aos sistemas de geração existentes e, em alguns casos, otimização dos algoritmos de previsão estabelecidos.

### 2.1 HORIZONTES DE PREVISÃO

A adequada classificação da velocidade do vento deve considerar o horizonte de previsão. A estimativa dessa duração, em especial para os operadores do sistema de transmissão, é fundamental, tendo em vista que, parte das receitas decorrentes da venda da energia eólica gerada se deve à confiabilidade dos dados obtidos. A escala de tempo, utilizada pelos métodos de previsão de velocidade, é mostrada abaixo [5]:

- Muito curto prazo (alguns segundos a 30 min);
- Curto prazo (30 minutos a 6 horas);
- Médio prazo (6 horas a 24 horas);
- Longo prazo (24 horas a 72 horas);
- Muito longo prazo (acima de 72 horas).

A classificação, utilizando os limites de tempo definidos acima, não apenas simplifica o estudo, mas, também, ajuda no estabelecimento do método mais preciso de previsão, dependendo da aplicação do modelo.

#### 2.1.1 Previsão de Muito Curto Prazo

O uso adequado dos algoritmos de previsão da velocidade do vento tem, intrinsecamente, dependência da classificação em escala de tempo. Para que sua utilização seja precisa, os métodos devem ser avaliados em locais com diferentes condições climáticas. Dessa forma, devem ser utilizados parâmetros que calculem, além da precisão do método, as condições adequadas de sua utilização. Esses estudos de otimização são feitos, geralmente, para escolher o método de previsão mais preciso.

Os métodos de previsão de muito curto prazo, preveem valores futuros num horizonte de tempo reduzido, indo, usualmente, de alguns segundos até 30 minutos. Métodos de aprendizagem inteligentes, como as Redes Neurais Profundas, do inglês, *Deep Neural Networks*, DNN, são usados para prever velocidades de vento de 10 min até 30 minutos [5]. No domínio de previsão, houve a evolução dos algoritmos de Aprendizado de Máquina, do inglês, *Machine Learning*, ML, com Algoritmos de Previsão de Decomposição, do inglês, *Decomposition Forecasting Algorithms*, DFA, sendo estes os mais usados. O princípio de um algoritmo de decomposição consiste, basicamente, na decomposição da série temporal em variáveis

individuais [10]. A velocidade e a potência do vento são consideradas unidades individuais e, dessa forma, são analisadas para a previsão, sendo combinadas ao final para obter a série resultante. A predição pode ser realizada com Redes Neurais Artificiais, o modelo Autorregressivo Integrado de Médias Móveis, do inglês, *Autoregressive Integrated Moving Average*, ARIMA, ou Máquinas de Vetor de Suporte, do inglês, *Support Vector Machines*, SVM.

### 2.1.2 Previsão de Curto Prazo

Muitos mercados do dia seguinte precisam, até o final do dia, limpar os cenários estabelecidos. Devido a essa característica, os métodos de previsão da velocidade do vento de curto prazo são os mais utilizados. Entre os modelos existentes, a previsão utilizando dois ou mais métodos de aprendizado de máquina, combinados com um modelo de série temporal como, por exemplo, ARMA, ARIMA ou ARIMAX têm se destacado [5].

Entre os modelos estatísticos, o Modelo Autorregressivo de Médias Móveis, do inglês, *Autoregressive Moving Average*, ARMA, é um método de previsão bastante utilizado. Esse modelo, cuja ordem denota a defasagem os valores presentes e passados da variável aleatória, tem como objetivo a previsão de valores futuros, podendo esta ser de curto ou longo prazo. Dessa forma, uma combinação autorregressiva, de ordem  $p$ , e média móvel, de ordem  $q$ , formam um modelo ARMA ( $p, q$ ), cuja ordem denota a defasagem entre os valores presentes e passados da variável.

O ARIMA, também conhecido como metodologia Box-Jenkins [11], de ordem ( $p, d, q$ ), é uma generalização do modelo ARMA, sendo utilizado em situações em que a série temporal não é estacionária. Nestes casos, antes de sua utilização, é necessária uma etapa inicial de diferenciação, que corresponde à parte integrada do modelo, podendo esta ser aplicada uma ou mais vezes à série temporal, de forma a eliminar sua não estacionariedade. As previsões feitas com o ARIMA minimizam o Erro Médio Quadrático, do inglês, *Mean Squared Error*, MSE, do modelo. A identificação do modelo ARIMA mais adequado aos dados é feita com base na função de autocorrelação e autocorrelação parcial estimadas.

O modelo Autorregressivo Integrado de Médias Móveis com Entradas Exógenas, do inglês, *Autoregressive Integrated Moving Average with Exogenous Variable*, ARIMAX, representa uma

generalização do ARIMA, sendo conhecido como metodologia de Box-Tiao [12]. O modelo ARIMAX é considerado multivariado, acrescentando um componente linear em função das observações das variáveis exógenas. O ARIMAX pode ser entendido como a combinação dos modelos Autorregressivo, do inglês, *Autoregressive*, AR, de ordem  $p$ , Integrado, do inglês, *Integrated*, I, de ordem  $d$ , e Média Móvel, do inglês, *Moving Average*, MA, de ordem  $q$ , e Exógeno, do inglês, *Exogenous*, X, de ordem  $r$ , podendo ser representado por ARIMAX ( $p, d, q, r$ ).

### 2.1.3 Previsão de Médio Prazo

Previsões para este horizonte são, normalmente, baseadas no uso de ANNs, modelos físicos de clima ou modelos híbridos, combinando ambos ou alguns desses modelos. O trabalho desenvolvido em [13] utiliza a Transformada Discreta Wavelet, do inglês, *Discrete Wavelet Transform*, DWT, para decompor o sinal, separando os dados em diferentes componentes de frequência. Um algoritmo de previsão aproxima os coeficientes das formas de onda decompostas que são, então, reconstruídas, passando a conter a amostra original, somada à previsão das próximas 24 horas.

Um estudo de caso usando Redes Neurais Diretas, do inglês, *Feedforward Neural Network*, FNN, é apresentado em [14] para previsões de 1 a 24 horas à frente. O número ideal de neurônios, neste estudo, é obtido por meio do método de tentativa e erro. Os resultados fornecem um Erro Absoluto Médio Percentual, do inglês, *Mean Absolute Percentage Error*, MAPE, 12% menor para as FNN, comparando-as com sua persistência. Uma abordagem híbrida para previsões no intervalo de 1 a 48 horas é desenvolvida em [15], sendo a previsão preliminar feita com base em modelos de Previsão Numérica do Clima. A precisão da previsão realizada, feita com base nos erros entre os valores previstos e os valores reais da potência e direção do vento, além de valores de NWP imprecisos, é melhorada por meio da aplicação de regras difusas.

### 2.1.4 Previsão de Longo e Muito Longo Prazo

As previsões de longo prazo encontram sua principal aplicação na análise de comprometimento de equipamentos e na programação da manutenção de sistemas elétricos. Em [16], um modelo híbrido,

combinando a Otimização por Enxame de Partículas, do inglês, *Particle Swarm Optimization*, PSO, e o Sistema de Inferência Difusa Baseado em Rede Adaptativa, do inglês, *Adaptive Neuro-Fuzzy Inference System*, ANFIS, definido como modelo PSO+ANFIS, foi usado para fazer previsões, com 1 dia de antecedência, em intervalos de 15 minutos.

Uma abordagem física para previsão da capacidade de geração de energia eólica, no horizonte de 1 a 10 dias à frente, utilizando Previsões de Clima *Ensemble*, do inglês, *Weather Ensemble Predictions*, WEP, é desenvolvida em [17]. O modelo é calibrado e suavizado para estimar adequadamente a incerteza às condições climáticas. A comparação dos resultados com métodos estatísticos de série temporal, baseada em modelos ARMA, mostram que o WEP apresentou resultados mais precisos ao longo de uma semana.

## 2.2 ESTUDOS RELACIONADOS

A confiabilidade e precisão dos modelos de WSF é essencial para a exploração eficiente da energia eólica. Em particular, a Previsão da Velocidade do Vento de Curto Prazo de Alta Precisão, do inglês, *Short-term Wind Speed Forecasting*, ST-WSF, tem impacto significativo na eficiência desses sistemas.

Em [18], é desenvolvido um modelo híbrido de WSF com base na rede de Memória Longa de Curto Prazo, do inglês, *Long Short-Term Memory*, LSTM, e métodos de decomposição com Otimizador Lobo Cinzento, do inglês, *Grey Wolf Optimizer*, GWO. No estágio de pré-processamento, os dados ausentes são preenchidos pelo método de Média Móvel Ponderada, do inglês, *Weighted Moving Average*, WMA, e os dados da série temporal são suavizados pela filtragem WMA, sendo estes dados usados como entrada do modelo.

Um modelo híbrido adaptativo, baseado em Decomposição de Modo Variacional, do inglês, *Variational Mode Decomposition*, VMD, modelo ARIMA e Rede de Crença Profunda, do inglês, *Deep Belief Network*, DBN, é proposto [19]. Neste trabalho, cuja eficácia foi testada dados de dois parques eólicos da China, a velocidade do vento é decomposta em componentes regulares e irregulares por VMD, sendo aplicado o modelo ARIMA na previsão dos componentes regulares e a DBN, na previsão dos componentes irregulares.

Os autores desenvolvem em [20] um modelo baseado em Aprendizagem Profunda, do inglês, *Deep Learning*, DL, que integra a Transformada

Discreta de Pacote Wavelet, do inglês, *Discrete Wavelet Packet Transform*, DWPT, e a Memória Bidirecional Longa de Curto Prazo, do inglês, *Bidirectional Long Short-Term Memory*, BLSTM. O método captura, com precisão, características temporais profundas e aprende a variabilidade da velocidade do vento nas séries temporais incorporando, por meio da memória bidirecional das redes BLSTM, uma forma eficaz de lidar com as incertezas, investigando, dessa forma, os dados das camadas ocultas anteriores e futuras. Em [21] é feito um estudo comparativo entre técnicas de ST-WSF, utilizando-se ANN. O trabalho mostra que a memória local e as entradas relevantes tornam as Redes Neurais Recorrentes, do inglês, *Recurrent Neural Networks*, RNN, mais adequadas para uso com TS do que as redes de alimentação direta.

Um modelo híbrido, utilizando um Codificador Automático, do inglês, *Autoencoder*, AE, uma rede BLSTM, para predição do intervalo da velocidade do vento é apresentado em [22]. O AE extrai, inicialmente, informações relevantes dos dados de velocidade do vento, gerando os dados que a rede BLSTM usará para gerar os intervalos de predição. O trabalho propõe que, para tarefas de previsão de séries temporais, a extração de recursos utilizando um AE é mais eficaz do que criar DDNs. Em [23], a previsão da velocidade do vento, com base no histórico de geração de energia eólica, é feita pela combinação dos modelos ARMA e SVM, sendo a otimização dos parâmetros da SVM feita pelo uso do PSO. O trabalho apresenta, ainda, um modelo híbrido, baseado no modelo ARMA e PSO-SVM, onde os pesos ideais são obtidos pelo emprego do método de minimização de covariância e PSO. Por fim, em [24], um método de WSF, utilizando DNN, é constituído. O modelo, que utiliza dados em tempo real do parque eólico, é estabelecido usando um Codificador Automático Empilhado, do inglês, *Autoencoder Stacked*, AES, para eliminação de ruído e uma rede LSTM.

## 3 MODELO HÍBRIDO UTILIZADO

Nesse trabalho, a modelagem do sistema híbrido segue a hipótese demonstrada na Equação 1, que foi proposta por Zhang, em 2003.

$$Z_t = L_t + N_t \quad (1)$$

Na equação, tem-se que a série temporal  $Z_t$  é dada pela combinação linear de padrões lineares  $L_t$  e padrões não-lineares  $N_t$ . Como a metodologia

Box-Jenkins garante a modelagem dos padrões lineares, a série de erro é obtida subtraindo da série  $Z_t$  a previsão dos padrões lineares  $\widehat{L}_t$ , sendo calculada por:

$$E_t = Z_t - \widehat{L}_t \quad (2)$$

E essa série de erro (resíduo), por não possuir correlação temporal significativa, é modelada pelas técnicas de ML. A previsão de  $\widehat{N}_t$  é dada por:

$$\widehat{N}_t = f(e, e_{t-1}, \dots, e_{t-n}) + \varepsilon_t \quad (3)$$

Em que  $f$  é uma função gerada pela técnica de ML,  $(e, e_{t-1}, \dots, e_{t-n})$  são os *lags* (retardos) do erro e  $\varepsilon_t$  é um componente aleatório. Por fim, a previsão final é dada pela soma das previsões lineares e não lineares dadas pela Equação 4.

$$\widehat{Z}_t = \widehat{L}_t + \widehat{N}_t \quad (4)$$

O modelo híbrido utilizará a equação 4 como base, onde as previsões lineares e não lineares serão somadas para chegarmos à previsão final.

## 4 DESCRIÇÃO DAS BASES DE DADOS

As três bases de dados possuem, cada uma, um total de 144 amostras. São registros numéricos das médias mensais de velocidade do vento, em metros por segundo, de três capitais do Nordeste: Recife, Fortaleza e Natal. Não há presença de dados não numéricos ou dados ausentes.

Os registros são datados de janeiro de 2003 a dezembro de 2014, entretanto, não há uma coluna índice na base que especifique a data das observações. Essas amostras estão apenas enumeradas em sequência, sendo a média das velocidades de janeiro de 2003 a observação 0 e a média das velocidades de dezembro de 2014 a observação 143. Foi criada, então, utilizando a biblioteca *datetime* do *Python*, uma coluna índice, especificando as datas de cada registro, possibilitando a execução dos modelos. Esses dados também foram utilizados na elaboração de um modelo híbrido de previsão mensal de velocidade dos ventos, tendo sido disponibilizados. [8]

## 5 PROTOCOLO EXPERIMENTAL E RESULTADOS

Para se chegar ao escopo do artigo, foi necessária a utilização de algumas métricas e definições de uso corrente. O horizonte de previsão, o erro entre as velocidades prevista e real, bem como as potências

prevista e real, podem ser mensurados por meio de algumas definições de erro padrão. As expressões matemáticas para o cálculo do Erro Absoluto Médio Percentual, do inglês, *Mean Absolute Percentage Error*, MAPE, Erro Quadrático Médio, do inglês, *Mean Squared Error*, MSE, Erro Absoluto Médio, do inglês, *Mean Absolute Error*, MAE, e o Erro Quadrático Médio, do inglês, *Root Mean Square Error*, RMSE, são mostradas nas Equações (5-8) [7].

$$MAPE = \frac{100}{N} \sum_{t=1}^N \frac{|p_f - p|}{p_f} \quad (5)$$

$$MSE = \frac{1}{N} \sum_{t=1}^N (p_f - p)^2 \quad (6)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |p_f - p| \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (p_f - p)^2}{N}} \quad (8)$$

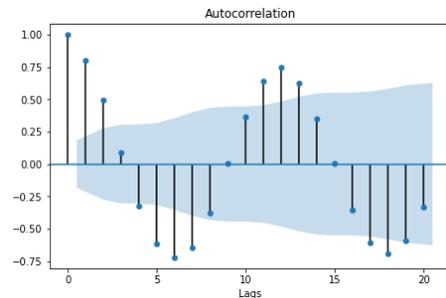
em que  $N$  é o horizonte de previsão, ou o tamanho do conjunto avaliado,  $p_f$  é a energia eólica prevista e  $p$  é o valor real força do vento.

O primeiro passo na execução das atividades experimentais foi aplicar o modelo estatístico linear às três bases, obtendo, assim, a série de resíduos de cada uma delas. O modelo estatístico utilizado foi o ARMA para as séries de Natal e Fortaleza, já que não foi necessário aplicar a diferenciação para torná-las estacionárias. Foi, também, realizado um teste de estacionariedade em cada uma das séries.

A série de Recife, após o teste, foi identificada como não estacionária. Entretanto, ao se aplicar a diferenciação, a série passou a apresentar *lags* com correlação linear significativa em seu resíduo. Foi dada, então, sequência à elaboração do modelo sem que esta série fosse diferenciada, gerando resultados mais satisfatórios nos modelos híbridos.

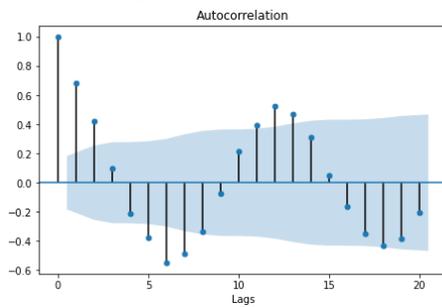
As funções de autocorrelação, do inglês, *Autocorrelation Function*, ACF, das três bases, mostradas nas Figuras 1, 2 e 3, mostram que há uma correlação linear significativa. A área azulada corresponde ao intervalo de confiança.

Figura 1 – ACF Fortaleza.



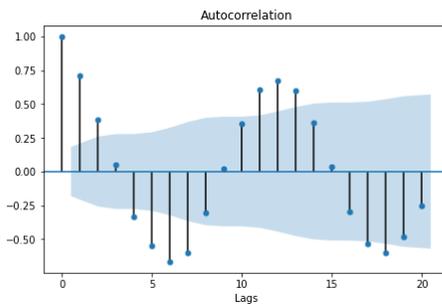
Fonte: Própria.

**Figura 2 – ACF Natal.**



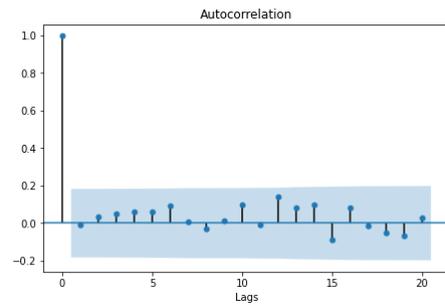
Fonte: Própria.

**Figura 3 – ACF Recife.**



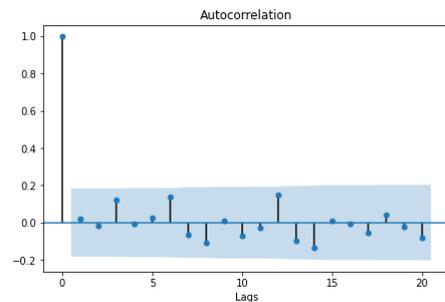
Fonte: Própria.

**Figura 5 – Resíduo da série de Natal.**



Fonte: Própria.

**Figura 6 – Resíduo da série de Fortaleza.**

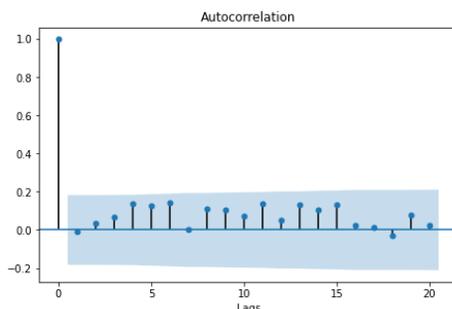


Fonte: Própria.

Um problema na execução do modelo linear se deveu ao critério da invertibilidade das ordens de médias móveis e autorregressivas, dificultando a escolha dos *lags* que poderiam ser candidatos para o modelo. A invertibilidade é um critério matemático que deve ser respeitado, de forma a viabilizar a execução do modelo. A biblioteca utilizada para essa modelagem foi o *statsmodel*, enquanto para a modelagem não linear se utilizou a *scikit-learn*.

Feita a modelagem, através do uso das funções de autocorrelação das séries, a etapa seguinte foi avaliar as séries de resíduos obtidas. A partir dos respectivos correlogramas, foi observado que após a aplicação do modelo estatístico, o resíduo apresentou o comportamento de um ruído Independente e Identicamente Distribuído, do inglês, *Identically and Independently Distributed*, IID, mostrado nas Figuras 4, 5 e 6:

**Figura 4 – Resíduo da série de Recife.**



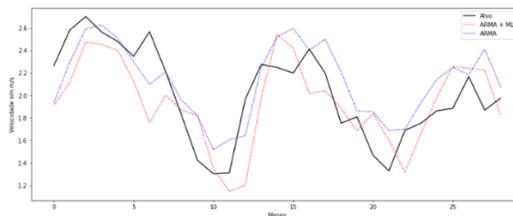
Fonte: Própria.

Nos gráficos, observa-se que os resíduos não possuem correlação temporal significativa, já que os retardos não ultrapassam o intervalo de confiança de 95%, que corresponde à área azulada. Para a modelagem da série de resíduos, foram utilizadas duas técnicas de AI distintas: as redes MLP e LSTM. A *perceptron* multicamada é uma rede neural que consegue driblar bem o *overfitting* e não possui tantos parâmetros, não obstante, sua desvantagem é a inicialização aleatória dos pesos, sendo a LSTM uma rede profunda. Para esse tipo de rede, o ideal é que o conjunto de treinamento possua dados suficientes para sua execução.

Cada uma dessas técnicas de aprendizado de máquina foi combinada com o ARMA, sendo feita a comparação dos resultados. Foram utilizados 86 pontos para treino, 29 para validação e 29 para teste. A seleção dos hiperparâmetros foi feita por um *grid search*, um *loop* onde são testados diversos hiperparâmetros. Os valores que tiveram o melhor resultado, compuseram a arquitetura final da rede, correspondendo à quantidade de neurônios em cada camada, quantidade de camadas ocultas, função de ativação e taxa de aprendizado. As métricas de comparação do desempenho foram o MAPE e o MSE, calculados por meio das expressões matemáticas descritas nas equações (5, 6).

Para as bases, foram aplicados cinco modelos, todos eles sem se levar em consideração variáveis exógenas. O modelo estatístico linear foi o primeiro, seguido pela MLP e LSTM, ambos executados de forma individual. Por fim, foi feita a combinação do modelo estatístico com a modelagem residual feita por uma técnica de ML.

Para as bases de Recife e Natal, o modelo que apresentou menores valores de MAPE e MSE foi a combinação dos modelos ARMA + MLP. Já na base de Fortaleza, o modelo estatístico foi trabalhado de forma individual, tendo apresentado melhor resultado do que os modelos híbridos e de IA trabalhados isoladamente. Fazendo um comparativo das técnicas de IA entre si, os modelos que utilizaram MLP obtiveram melhores resultados que os modelos com LSTM. As Figuras 7, 8 e 9 fazem um comparativo do melhor modelo híbrido com o modelo estatístico linear utilizado individualmente. Somente em Fortaleza, o modelo estatístico obteve melhores resultados. Na Tabela 1, os valores de MAPE e MSE para todos os modelos utilizados são exibidos, destacando-se, em negrito, os melhores valores obtidos.



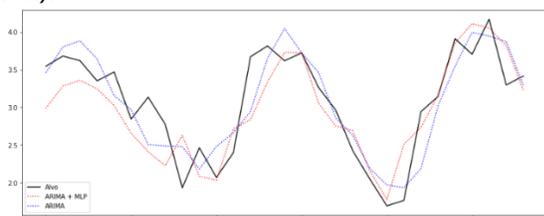
Fonte: Própria.

Tabela 1 – Valores de MAPE e MSE para cada modelo.

BASE	MÉTODO	MAPE	MSE
Recife	ARMA	16,384	0,123
	MLP	16,658	0,142
	LSTM	22,07	0,24
	ARMA + MLP	<b>10,5</b>	<b>0,06</b>
	ARMA + LSTM	15,19	0,11
Fortaleza	ARMA	<b>9,636</b>	<b>0,102</b>
	MLP	11,25	0,171
	LSTM	14,841	0,247
	ARMA + MLP	10,46	0,15
	ARMA + LSTM	11,415	0,181
Natal	ARMA	12,363	0,312
	MLP	12,023	0,394
	LSTM	18,74	1,16
	ARMA + MLP	<b>9,87</b>	<b>0,228</b>
	ARMA + LSTM	13,94	0,555

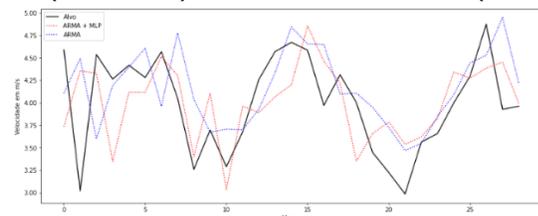
Fonte: Própria.

Figura 7 – Comparativo das previsões de Fortaleza. Modelo híbrido (ARMA+MLP) versus modelo estatístico (ARMA).



Fonte: Própria.

Figura 8 – Comparativo das previsões de Natal. Modelo híbrido (ARMA+MLP) versus modelo estatístico (ARMA).



Fonte: Própria.

Figura 9 – Comparativo das previsões de Recife. Modelo híbrido (ARMA+MLP) versus modelo estatístico (ARMA).

Prever a velocidade dos ventos permite às companhias organizar melhor a distribuição de energia num horário específico, de modo que seja mais proveitoso para a rede. Em outras palavras, é possível prever a quantidade de energia a ser entregue num determinado horizonte de previsão.

Combinar modelos estatísticos lineares de previsão com modelos de AM para a previsão de resíduos gera resultados mais satisfatórios, no que diz respeito a assertividade da previsão.

## 6 CONCLUSÃO

A velocidade do vento é uma das principais fontes de energia limpa, por ser inesgotável e estando disponível de forma abundante. A correta previsão de sua velocidade pode beneficiar sobremaneira os operadores das plantas de geração de energia eólica, porém, trata-se de uma tarefa cada vez mais desafiadora.

O uso de técnicas de IA permitem, ainda, através da análise dos dados de sensores, a criação de planos de manutenção preditiva. O planejamento atividades como, por exemplo, a troca de componentes dos aerogeradores, reduz o *downtime*

das manutenções, diminuindo ainda as perdas na geração de energia.

Esse artigo utiliza um modelo híbrido para previsão de velocidade do vento, o que foi feito em três etapas: modelagem estatística linear, utilizando o modelo ARMA, modelagem do resíduo, utilizando MLP e LSTM e, por fim, a combinação do modelo estatístico com o modelo de ML. Além do modelo híbrido, as técnicas disponíveis foram utilizadas, individualmente, a fim de comparar os seus resultados.

Com exceção da série de velocidade dos ventos de Fortaleza, que obteve melhor resultado com o modelo estatístico, as previsões das séries de Recife e Natal foram mais satisfatórias ao combinar o modelo estatístico com o modelo de aprendizado de máquina, evidenciando a viabilidade de se combinar esses modelos, de forma a obter os melhores resultados na previsão da velocidade do vento.

## 7 TRABALHOS FUTUROS

Entre as técnicas de ML mais promissoras para a previsão de séries temporais, a Regressão por Vetor de Suporte não foi utilizada nesse artigo. No que diz respeito à previsão de séries temporais, a SVR obtém excelentes resultados na modelagem da série de resíduos, obtida pelo modelo estatístico linear. Neste artigo não foram, também, consideradas variáveis exógenas. Ao considerar essas variáveis, os resultados podem se tornar melhores, aumentando a acurácia do modelo.

Como sugestão de trabalho futuros, fica a elaboração de um modelo híbrido, que combine o modelo estatístico com outras técnicas de IA [25, 26]. Também fazer uso de abordagens de seleção dinâmica ou *ensembles* [26, 27], passando a considerar variáveis exógenas [26].

## REFERÊNCIAS

- [1] DHIMAN, H. S.; DEB, D. **Decision and Control in Hybrid Wind Farms**. 2019. Disponível em: [https://www.researchgate.net/publication/335208382\\_Decision\\_and\\_Control\\_in\\_Hybrid\\_Wind\\_Farms](https://www.researchgate.net/publication/335208382_Decision_and_Control_in_Hybrid_Wind_Farms).
- [2] DHIMAN, H. S.; DEB, D. **A Review of Wind Speed and Wind Power Forecasting Techniques**. 2020. Disponível em: [https://www.researchgate.net/publication/344150296\\_A\\_Review\\_of\\_Wind\\_Speed\\_and\\_Wind\\_Power\\_Forecasting\\_Techniques](https://www.researchgate.net/publication/344150296_A_Review_of_Wind_Speed_and_Wind_Power_Forecasting_Techniques).
- [3] WU, Y.-K.; HONG, J.-H. **A literature review of wind forecasting technology in the world**. IEEE Power Tech. Conference 1-5, Lausanne, Switzerland, July 2007. Disponível em: <https://ieeexplore.ieee.org/document/4538368>.
- [4] FABRI, A.; GOMEZ, T. **Assessment of the Cost Associated With Wind Generation Prediction Errors in a Liberalized Electricity Market**. IEEE Transactions Power Systems, v. 20, n.3, p. 1440-1146, 2005. Disponível em: <https://ieeexplore.ieee.org/iel5/59/32048/01490597.pdf>.
- [5] DHIMAN, H. S.; DEB, D. **A Review of Wind Speed and Wind Power Forecasting Techniques**. 2020. Disponível em: [https://www.researchgate.net/publication/344150296\\_A\\_Review\\_of\\_Wind\\_Speed\\_and\\_Wind\\_Power\\_Forecasting\\_Techniques](https://www.researchgate.net/publication/344150296_A_Review_of_Wind_Speed_and_Wind_Power_Forecasting_Techniques).
- [6] LIMA JUNIOR, A. R. *et al.* **An Experimental Study of Fitness Function and Time Series Forecasting Using Artificial Neural Networks**. Proceedings of the 12th Annual Conference Companion on Genetic and Evolutionary Computation. p. 2015-2018. Conference, Oregon, USA, July 7-11, 2010. Disponível em: [https://www.researchgate.net/publication/220740909\\_An\\_experimental\\_study\\_of\\_fitness\\_function\\_and\\_time\\_series\\_forecasting\\_using\\_artificial\\_neural\\_networks](https://www.researchgate.net/publication/220740909_An_experimental_study_of_fitness_function_and_time_series_forecasting_using_artificial_neural_networks).
- [7] SAEED, A.; LI, C. *et al.* **Hybrid Bidirectional LSTM Model for Short-Term Wind Speed Interval Prediction**. IEEE Access, v. 8, p. 182283-182294, 2020. Disponível em: <https://ieeexplore.ieee.org/document/9209989>.
- [8] DE MATTOS NETO, P. S. G.; OLIVEIRA, J., F., L. *et al.* **A Hybrid Nonlinear Combination System for Monthly Wind Speed Forecasting**. IEEE Access, v. 8, p. 191365-191377, 2020. Disponível em: <https://ieeexplore.ieee.org/document/9229129>.
- [9] JAO, J. **A Hybrid Forecasting Method for Wind Speed**. MATEC Web of Conferences, Conference 232, January 2018. Disponível em: [https://www.researchgate.net/publication/329037963\\_A\\_Hybrid\\_Forecasting\\_Method\\_for\\_Wind\\_Speed](https://www.researchgate.net/publication/329037963_A_Hybrid_Forecasting_Method_for_Wind_Speed).
- [10] QIAN, Z.; PEI, Y., *et al.* **A Review and Discussion of Decomposition-Based Hybrid Models for Wind Energy Forecasting Applications**. Applied Energy, v. 235, p. 939-953, 2019. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0306261918316489>

Disponível em:  
<https://ieeexplore.ieee.org/document/5224014>.

- [11] BOX, G. E. P.; JENKINS G. M., *et al.* **Time Series Analysis: Forecasting and Control**. 5. ed. New Jersey: Wiley, 2015. Disponível em: <https://books.google.com.br/books?hl=pt-BR&id=rNt5CgAAQBAJ&pg=PR7&dq=Box+and+Jenkins:+Time+Series+Analysis,+Forecastin+g+and+Control>.
- [12] BOX, G. E. P.; TIAO G. C. **Intervention Analysis with Applications to Economic and Environmental Problems**. *Journal of the American Statistical Association*, v. 70, n. 349, p. 70-79, 1975. Disponível em: [https://rmgsc.cr.usgs.gov/outgoing/threshold\\_articles/Box\\_Tiao1975.pdf](https://rmgsc.cr.usgs.gov/outgoing/threshold_articles/Box_Tiao1975.pdf).
- [13] BERREZZEK, F.; KHELIL, K.; BOUADJILA, T. **Efficient Wind Speed Forecasting Using Discrete Wavelet Transform and Artificial Neural Networks**. *Revue d'Intelligence Artificielle*, v. 33, n. 6, p. 447-452, 2019. Disponível em: <http://www.univ-soukahras.dz/en/lab/leer/pub/2467>.
- [14] UPADHYAY, K. G.; CHOUDHARY, A. K.; TRIPATHI, M. M. **Short-Term Wind Speed Forecasting Using Feed-forward Back-Propagation Neural Network**. *International Journal of Engineering, Science and Technology*, v. 3, n. 5, p. 107-112, 2011. Disponível em: [https://www.researchgate.net/publication/235660032\\_Short-term\\_Wind\\_Speed\\_Forecasting\\_using\\_Geed-forward\\_Back-propagation\\_Neural\\_Network](https://www.researchgate.net/publication/235660032_Short-term_Wind_Speed_Forecasting_using_Geed-forward_Back-propagation_Neural_Network).
- [15] SOMAN, S., S.; ZAREIPOUR, H. *et al.* **A Review Of Wind Power and Wind Speed Forecasting Methods With Different Time Horizons**. *North American Power Symposium, Conference 26-28, Texas, USA, September 2010*. Disponível em: <https://ieeexplore.ieee.org/document/5619586>.
- [16] POUSINHO, H., M., I.; CATALÃO, J., P., S.; MENDE, V., M., F. **Wind power short-term prediction by a hybrid PSO-ANFIS approach**. *15th IEEE Mediterranean Electrotechnical Conference. Conference 26-28, Valletta, Malta, April 2010*. Disponível em: <https://ieeexplore.ieee.org/document/5475923>.
- [17] TAYLOR, J. W.; Mcsharry, P. E.; BUIZZA, R. **Wind Power Density Forecasting Using Ensemble Predictions And Time Series Models**. *IEEE Transactions on Energy Conversion*, v. 24, n. 3, p. 775-782, 2009. Disponível em: <https://ieeexplore.ieee.org/document/5224014>.
- [18] ALTAN, A.; KARASU, S.; ZIO, E. **A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer**. *Applied Soft Computing*, v. 100, 2021. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1568494620309352>.
- [19] ZHANG, T.; WEI, Y.; TAN, Z. **An adaptive hybrid model for short term wind speed forecasting**. *Energy*, v. 190, 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0360544219312642>.
- [20] KURDIKERI, R., B.; RAJU, A., B. **Hybrid Deep Learning-Based Model for Wind Speed Forecasting Based on DWPT and Bidirectional LSTM Network**. *IEEE Acces*, v. 8, p. 229219-229232, 2020. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9306756>.
- [21] KURDIKERI, R., B.; RAJU, A., B. **Comparative Study of Short-Term Wind Speed Forecasting Techniques Using Artificial Neural Networks**. *IEEE International Conference on Current Trends towards Converging Technologies. Conference 1-3, Coimbatore, India, March 2018*. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8550849>.
- [22] SAEED, A.; LI, C. *et al.* **Hybrid Bidirectional LSTM Model for Short-Term Wind Speed Interval Prediction**. *IEEE Acces*, v. 8, p. 182283-182294, 2020. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9209989>.
- [23] WANG, Y.; WANG, W. **Clustered Hybrid Wind Power Prediction Model Based on ARMA, PSO-SVM, and Clustering Methods**. *IEEE Acces*, v. 8, p. 17071-17079, 2020. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8964381>.
- [24] LIU, X.; ZHANG, H., *et al.* **Wind Speed Forecasting Using Deep Neural Network With Feature Selection**. *Neurocomputing*, v. 397, p. 393-403, 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0925231220304148>.
- [25] DE OLIVEIRA, J.; F.; L., *et al.* **A hybrid optimized error correction system for time**

**series forecasting.** Applied Soft Computing, v. 87, p. 105970, 2020.

- [26] DE MATTOS NETO, Paulo SG *et al.* **An adaptive hybrid system using deep learning for wind speed forecasting.** Information Sciences, v. 581, p. 495-514, 2021.
- [27] DE OLIVEIRA, João FL; SILVA, Eraylson G.; DE MATTOS NETO, Paulo SG. **A hybrid system based on dynamic selection for time series forecasting.** IEEE Transactions on Neural Networks and Learning Systems, 2021.