

## MeanShift para Densificação de Dados Aplicado à Previsão de Geração de Energia Eólica <sup>\*</sup>

Hellen Dianne P. de Souza <sup>\*</sup> Shigeaki Leite de Lima <sup>\*\*</sup>  
Denisson Queiroz Oliveira <sup>\*\*\*</sup>

<sup>\*</sup> Instituto de Energia Elétrica, Universidade Federal do Maranhão,  
MA, (e-mail: hellen.souza@discente.ufma.br).

<sup>\*\*</sup> Instituto de Energia Elétrica, Universidade Federal do Maranhão,  
MA, (e-mail: shigeaki.lima@ufma.br)

<sup>\*\*\*</sup> Instituto de Energia Elétrica, Universidade Federal do Maranhão,  
MA, (e-mail: dq.oliveira@ufma.br)

---

**Abstract:** The forecast of wind energy generation is used by many energy companies in order to safely estimate the wind resource and in this way guarantee both the company and the consumer the increase in revenue and the supply of energy from a renewable source over a period of time. Some techniques were developed for this purpose and allow estimating the wind resource and the energy generated for a few hours, days or even weeks. In this work, an information theoretic learning technique for the densification of a wind measurement database is presented, generating new data with the same probability density function of the original set until obtaining the mode of the set, using these virtual data as scenarios of the forecast.

**Resumo:** A previsão de geração de energia eólica é usada por muitas empresas de energia com a finalidade de estimar com segurança o recurso eólico, e desta forma garantir tanto para a empresa quanto ao consumidor o aumento da receita e o fornecimento de energia de fonte renovável durante um período de tempo. Muitas técnicas foram desenvolvidas com este objetivo e permitem estimar o recurso eólico e a energia gerada desde algumas horas, dias ou mesmo semanas. Neste trabalho é apresentado uma técnica baseada em teoria da informação para a densificação de uma base de dados de medições de vento, gerando novos dados com a mesma função densidade de probabilidade do conjunto original até obter a moda do conjunto, utilizando estes dados virtuais como cenários de previsão.

*Keywords:* Mean Shift; Forecast; Models; Wind energy; Measurement.

*Palavras-chaves:* Mean Shift; Previsão; Modelos; Energia eólica; Medição.

---

### 1. INTRODUÇÃO

Os avanços tecnológicos e aumento da geração e consumo dos vários setores da sociedade, vem exigindo uma maior demanda de energia elétrica. Em vista de suprir a necessidade energética das indústrias e da população, países tem investido no desenvolvimento de tecnologias que possibilitam maior geração de energia ou que auxiliem de forma a sustentar a demanda por energia elétrica (EPE, 2021).

No Brasil as hidrelétricas são predominante na matriz energética, contudo há uma dependência da sazonalidades das cheias e vazantes dos rios onde estão instaladas, necessitando do apoio das térmicas para suprir a demanda (EPE, 2021). Discussões sobre o aquecimento global e a preservação do meio ambiente tem originado resistências a instalações de novas térmicas e outras fontes baseadas em petróleo, permitindo o protagonismo de outras fontes de energia limpa (ABEEolica, 2020).

Nos últimos 10 anos outras fontes de energia têm se destacado, impulsionadas pelos programas e políticas gover-

<sup>\*</sup> Suporte financeiro para o desenvolvimento do projeto e bolsa via Equatorial Maranhão e CNPQ

namentais como a geração de energia fotovoltaica e eólica. Todavia, o nordeste brasileiro se destaca na geração de energia eólica, com instalações de parques eólicos em quase todos os estados da região, em função de muitos estarem na orla e possuírem potencial eólico adequado (ABEEolica, 2020).

Como estratégia para determinar os locais adequados para a instalação dos parques, investimentos em pesquisa sobre o potencial de geração tem crescido conforme novos parques tem surgido (Maria et al., 2008).

Nesse contexto a realização de pesquisas voltadas a previsão de geração de energia torna-se indispensável para qualquer projeto em potencial. Como ferramentas para prever o potencial energético de uma região são utilizados algoritmos baseados em modelos paramétricos, não paramétricos e híbridos.

Neste trabalho foi desenvolvido um modelo de previsão híbrido, que utiliza o algoritmo *Mean Shift* que é baseado na teoria da informação, com o objetivo de realizar a previsão a partir da densificação de séries temporais resultantes de medições atmosféricas, que podem ser convertidos em novos cenários de geração.

## 2. MODELOS DE PREVISÃO

A previsão do potencial de geração de energia elétrica é um fator importante para elaboração de projetos eólicos, para a gestão, planejamento de parques já instalados, reduzir os riscos e agregar valor a operação (Oliveira et al., 2020) (Li et al., 2020).

Devido a variedade de agentes que podem interferir na velocidade do vento, os modelos de previsão são progressivamente aperfeiçoados, formando portanto técnicas mais precisas e complexas de previsão que podem variar de alguns minutos a previsão de anos (Tabela 1) (Alencar et al., 2018)(Yan et al., 2017). Deste modo, é possível classificar os modelos de previsão em quatro tipos:

Tabela 1. Escala temporal aplicada a modelos de previsão.

Escala	Horizonte
Curtíssimo Prazo	Alguns minutos a 1 hora
Curto Prazo	1 hora a várias horas
Médio Prazo	Várias horas a 1 semana
Longo Prazo	De 1 semana a 1 ano ou mais

A definição da escala temporal a ser utilizada no modelo de previsão, implica em um erro maior ou menor (taxa de assertividade) em função do período que se deseja estimar e conseqüentemente do volume de dados e das informações que deverão estar contidas no modelo (Oliveira et al., 2020).

De forma geral a empresa representante do parque eólico, quando este parque tem contrato de geração outorgado por meio de leilão de energia, tem como principal objetivo o despacho do valor em MW contratado e a previsão tem este viés em dirimir incertezas que possam comprometer a receita. Desta forma, quanto melhor o método de previsão, menor impacto nas receitas (Li et al., 2020)(Alencar et al., 2018).

### 2.1 Métodos paramétricos

Os métodos classificados como paramétricos são baseados em uma abordagem física e fazem uma análise voltada para técnicas de observações meteorológicas a partir de informações enviadas via satélite para centros de tratamento de dados, sendo normalmente utilizados para previsões de 6 a 72 horas a frente (Oliveira et al., 2020).

Dentre as ferramentas mais utilizadas está o *Numerical Weather Prediction (NWP)*, que são algoritmos de previsão do tempo, sendo um dos mais precisos em relação a irradiação solar e porcentagem de cobertura de nuvens (Senatore et al., 2020)(Wu and Wu, 2020).

Os modelos NWP são desenvolvidos com base nos dados obtidos sobre o tempo, temperatura, pressão e rugosidade da superfície. Tais modelos podem ser divididos em macro escala e meso-escala, e para realizarem previsões demandam um alto poder computacional devido a quantidade de informações processadas (Senatore et al., 2020)(Yan et al., 2017).

O mesmo ocorre com o modelo *Global Forecast System (GFS)* capaz de realizar previsões para até 10 dias com uma resolução espacial de 13 km e temporal de uma hora,

e até 16 dias com uma resolução de 34 km e três horas de intervalo de tempo (Wu and Wu, 2020) (França et al., 2020).

### 2.2 Métodos não-paramétricos

Já os métodos não-paramétricos utilizam uma abordagem numérica, estatística, séries temporais, inteligência artificial por exemplo, para obter um modelo de previsão. Esses métodos utilizam os dados histórico das medições para modelar o recurso eólico usando técnicas que variam de acordo com o problema, as mais utilizadas são: (França et al., 2020)

**ARMA** O modelo consiste em duas partes elementares, uma parte autorregressiva (AR) e uma parte de média móvel (MA). Esta técnica é um dos principais modelos de séries temporais e permite uma análise regressiva dos valores passados das variáveis, bem como a modelagem do termo de erro, como uma combinação linear de termos de erros que ocorrem no presente e no passado em vários momentos (Alencar et al., 2018)(França et al., 2020). Pode ser expresso por:

$$x_t = \sum_{i=1}^p \varphi_i x_{t-1} + \sum_{i=1}^q \theta_i \varepsilon_{t-1} \quad (1)$$

Em que  $t$  é um índice representado por um número inteiro e  $x_t$  são números reais. O  $\varphi_i$  são os parâmetros da parte autor-regressiva do modelo,  $\theta_i$  são os parâmetros da parte de média móvel e  $\varepsilon_t$  representa o ruído branco (Alencar et al., 2018)(Li et al., 2020).

**ARIMA** O modelo ARIMA é o desenvolvido para processos aleatórios não integrados, logo este modelo supre a lacuna deixada pelo ARMA que pressupõe a estacionariedade do processo e pode ser representado por (Chang et al., 2016):

$$(1 - \sum_{i=1}^p \phi_i L^i)(1 - L)^d x_t = (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t \quad (2)$$

Em que o processo ARIMA  $(p, d, q)$  expressa a propriedade de fatoração deste polinômio com  $p = p' - d$ , é uma série temporal autorregressiva integrada de médias móveis, em que  $p$  denota os números dos termos autorregressivos,  $d$  o número de vezes que a série deve ser diferenciada antes de tornar-se estacionária e  $q$  o número de termos de média móvel,  $L$  é operador de defasagem, o  $\phi_i$  são os parâmetros da parte auto-regressiva do modelo,  $\theta_i$  são os parâmetros da parte de média móvel e  $\varepsilon_t$  representa o termo de erro (Alencar et al., 2018) e (Chang et al., 2016).

**ARMAX** As técnicas de previsão ARMA e ARIMA não envolvem as variáveis exógenas do processo(Oliveira et al., 2020). Para considerar entradas externas é necessário utilizar o modelo autorregressivo com média móvel e entradas exógenas. Este modelo pode ser apresentado na seguinte forma:

$$x_t = \sum_{i=1}^p \varphi_i x_{t-1} + \sum_{i=1}^q \theta_i \varepsilon_{t-1} + \sum_{i=1}^b \eta_i d_{t-1} \quad (3)$$

Em que  $(p, q, b)$  se refere ao modelo com  $p$  termos autor-regressivos,  $q$  termos de média móvel,  $b$  termos de *input* exógeno e  $\eta_1, \dots, \eta_b$  são os parâmetros  $d_t$ . O  $\varphi_i$  são os parâmetros da parte auto-regressiva do modelo,  $\theta_i$  são os parâmetros da parte de média móvel e  $\varepsilon_t$  representa o termo de erro.

IA As técnicas baseadas no uso da Inteligência Artificial (IA) são amplamente usadas nas modelagens e previsões, dentre as quais as Redes Neurais Artificiais (RNAs) de múltiplas camadas são as mais utilizadas (Figura 1), devido a capacidade de mapear as não linearidades para qualquer nível de previsão. Deste modo, os algoritmos desenvolvidos em RNAs para previsões se tornam mais precisos (Haykin et al., 1999).

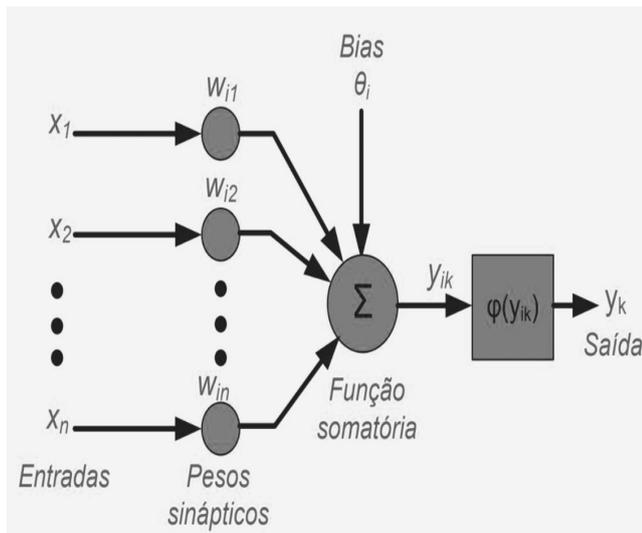


Figura 1. Neurônio Artificial

Por exemplo, um *Multilayer Perceptron (MLP)* consegue prever numa escala de 24 horas, ou seja um dia a frente. Já uma Rede Neural Bayesiana (RNB) consegue realizar uma previsão em escala mundial diariamente, sendo superior em desempenho quando comparado com os modelos de RNAs clássicos e modelos empíricos (Abesamis et al., 2019)(França et al., 2020).

O desenvolvimento de IA para previsões em sistemas eólicos são amplamente utilizadas para melhor atender as especificidades de cada parque. Nesse sentido vários tipos de RNAs podem ser empregados para esse fim, como os baseados em *Backpropagation*, *Multilayer Perceptron* e *Bayesiana* (França et al., 2020).

### 2.3 MÉTODOS HÍBRIDOS

A metodologia híbrida somente a título de informação, se baseia em unir técnicas de imagem de satélites com os de modelagem numérica, tendo portando um melhor desempenho que as demais metodológicas, visto que desta forma consegue extrair informações mais precisas, o que impacta nos resultados da previsão (Alencar et al., 2018)(Oliveira et al., 2020).

### 3. ALGORITMO *MEAN SHIFT*

O Algoritmo *Mean Shift* foi desenvolvido baseado na teoria da informação e entropia de Rényi, sendo comumente utilizado para determinar a moda de conjuntos, a clusterização, a classificação, a segmentação, o *tracking* e recentemente aplicado a densificação (Lima et al., 2013).

Este algoritmo possibilita a captura das estruturas preponderantes de  $n$  conglomerados de dados por meio de uma função de densidade de probabilidade (FDP) com Kernel Gaussiano, sendo um método muito utilizado em análises de processamento de imagens (Rêgo et al., 2016).

A utilização do algoritmo *Mean Shift* no contexto da análise de previsão de sistemas eólicos implica em uma abordagem diferenciada para provisão em séries temporais (Wu et al., 2018). Logo, admitindo uma amostra de dados  $x_0 = (x_1)_{i=1}^n ER^D$ , com base no modelo não-paramétrico de janela de Parzen, a função de densidade de probabilidade é definida por:

$$p(x, \sigma) = 1/n \sum_{i=1}^n G_{\sigma}(x - x_1) \quad (4)$$

Em que  $G_0$  um kernel gaussiano dado por  $G(t) = e^{-1/2}$  com largura de banda  $\sigma > 0$ .

Por meio do algoritmo *Mean Shift* é possível obter a moda de uma conjunto em uma base de dados, em que  $\rho(x) = 0$  (Lima et al., 2013)(Miranda et al., 2012). Para isso é necessário que o ponto estacionário seja dado por:

$$m(x) = \frac{\sum_{i=1}^n G_{\sigma}(x - x_i) * x_i}{\sum_{i=1}^n G_{\sigma}(x - x_i)} \quad (5)$$

Em que a diferença entre  $m(x)$ - $x$ , é denominado de *Mean Shift* (Wu and Yang, 2007)(Rêgo et al., 2016). Os parâmetros de utilização deste algoritmo reduzem a entropia de  $x$ ,  $H(x)$ , à medida em que conserva a distância de Cauchy-Schwartz  $D_{cs}(x, x_0)$ , entre  $x$  e  $x_0$  anexado em algum valor de  $K$ . Desta forma, a função FDP pode ser calculada segundo a entropia de Renyi's, como:

$$H(x) = -\log \int_{-\infty}^{+\infty} p_2(x) dx \quad (6)$$

Substituindo (4) em (6) têm-se:

$$H(x) = -\log V(x) \quad (7)$$

Em que o  $\sigma' = \sqrt{2\sigma}$  e  $V(x)$  é denominado como a informação potencial da FDP de  $p(x)$ (Miranda et al., 2012)(Wu and Yang, 2007). Na sequência a entropia cruzada entre duas FDPs pode ser calculada por:

$$H(x, x_0) = -\log V(x, x_0) \Rightarrow V(x, x_0) = \frac{1}{n^2} \sum_{j=1}^n G_{\sigma'}(x, x_0) \quad (8)$$

Para duas FDPs  $p$  e  $q$ , a distância de Cauchy-Schwartz pode ser expressa por (Rêgo et al., 2016):

$$D(x, x_0) = \log \frac{(\int p^2(x)dx)(\int q^2(x)dx)}{(\int p(x)q(x)dx)^2} \quad (9)$$

Considerando o conceito de que o cosseno de um ângulo entre dois vetores é a medida da distância, têm-se:

$$D_{cs}(x, x_0) = -[H(x) + H(x_0) - 2H(x, x_0)] \quad (10)$$

O deslocamento entre os dois objetivos pode ser otimizado e representado pela fórmula de otimização irrestrita sob o parâmetro  $\lambda$  (Lima et al., 2013)(Wu et al., 2018).

$$J(x) = \min H(x) + \lambda[D_{cs}(x, x_0) - K] \quad (11)$$

Com base na diferença de  $J(x)$  em relação a cada  $x_i \in x$  faculta a transmutação de  $x_0$  em outro conjunto  $x^{t+1}$  na interação  $t + 1$ , fazendo uso da informação inclusa em  $x$  na interação (Wu and Yang, 2007).

O agrupamento das verificações em distintas classes e obtidas pelo deslocamento de cada verificação. Cada deslocamento de um ponto  $x_i^t \in x^t$ , descreve uma FDP de  $x^t$  que contém informações da FDP de  $x_0$  (Miranda et al., 2012).

O conjunto  $x_v = [x^1 \cup x^2 \dots \cup x^t]$  equivalente aos dados virtuais concebidos pelo algoritmo *Mean Shift*, formam um *cluster* denso possuindo propriedades compatíveis com a série original  $x_0$  (Tuzel et al., 2009) (Lima et al., 2013).

#### 4. PREVISÃO EÓLICA

A região denominada Parque Nacional dos Lençóis Maranhenses, composta pelos Municípios de Barreirinhas, Paulinho Neves, Tutóia, Santo Amaro do Maranhão e Humberto de Campos, tem apresentado um potencial latente para geração de energia eólica, tendo despertado interesse de empresas do setor elétrico/eólico.

Em busca de identificar os locais que apresentem o melhores condições para geração de energia, projetos como o EOSOLAR vem sendo desenvolvidos. Este visa avaliar os efeitos micrometeorológicos em diferentes escalas temporais e espaciais, e são voltados ao planejamento e operação de parques eólicos e fotovoltaicos. O EOSOLAR é liderado pelo Instituto de Energia Elétrica da Universidade Federal do Maranhão e tem financiamento de empresas do setor elétrico do Maranhão.

Os estudos realizados pelo EOSOLAR visam mapear e avaliar o potencial eólico (recurso eólico) e solar do estado do Maranhão, fazendo medições na região de Barreirinhas e Paulino Neves em vários pontos de interesse, e para isto utiliza uma das mais avançada tecnologia existente no mercado internacional com foco na prospecção do vento.

O estudo de previsão de geração a partir de base de dados de velocidade e direção de vento associados ao *Mean Shift* foi realizado em três etapas. A primeira etapa envolveu a aquisição e formatação das variáveis de interesse de um banco de dados de ventos da região Norte/Nordeste, especificamente do município de São Luís/MA.

A segunda etapa envolve a seleção do período de análises, a filtragem das variáveis de interesse do arquivo bruto

(velocidade, direção e altura das medições) e a validação destes dados no sentido de verificar se não há erros abruptos na medição.

A terceira e última etapa envolve o desenvolvimento do algoritmo, as simulações a partir dos dados obtidos de velocidade, direção e altura, a plotagem dos resultados para discussão e o cálculo da energia futura em uma escala de curto e médio prazo, o que representa um horizonte de um dia a uma semana.

Uma vez que a previsão é obtida a partir de um conjunto de cenários gerados pela densificação do conjunto original, o qual traça uma trajetória de novos pontos/cenários a cada interação de  $\sigma$  é importante informar que para estimar a densidade a partir das janelas de Parzen é aplicando uma função kernel, que identifica um ponto local com conjunto contínuo derivado da FDP discriminada pelo kernel (Figura 2). Logo, com o uso do kernel gaussiano os pesos são os que estão mais próximos de  $x_i$ , havendo uma suavização da estimativa da densidade.

Ainda na Figura 2 ao estimar a janela de Parzen, se o  $\sigma$  for muito pequeno, não ocorre a interação entres os pontos estabelecidos pelos kernel, sendo, portanto, uma função cheia de modas, diferente da estimação de FDP. Caso o  $\sigma$  seja grande ocorre a perda da informação, já que os detalhes são extintos. Nesta proposta foram utilizados valores do  $\sigma$  que permitem gerar cenários com a estimação da FDP até atingir a moda que é o ponto de parada do algoritmo.

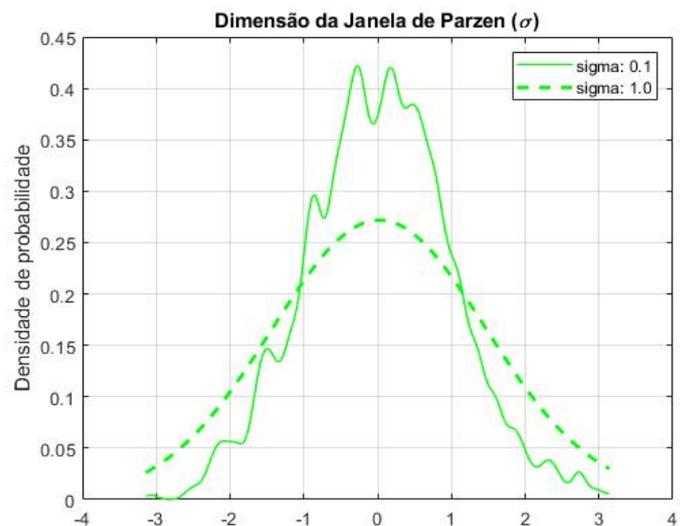


Figura 2. Dimensão da Janela de Parzen.

Na Figura 3 é ilustrado uma amostra de medição dos últimos três meses (outubro a dezembro), realizadas entre 50 e 250 metros de altura (medições a cada 10 metros), com predominância de ventos vindos do nordeste (distribuição entre 10 a 70 graus) e velocidades que variam de 5 m/s a 14 m/s.

Se for desprezado o eixo da altura e ilustrado apenas direção e velocidade (Figura 4) é possível observar que a predominância dos ventos nordeste se mantêm entre 8 m/s e 14 m/s nas direções entre 30 a 50 graus e alguns *outlier* sudeste quase nulos.

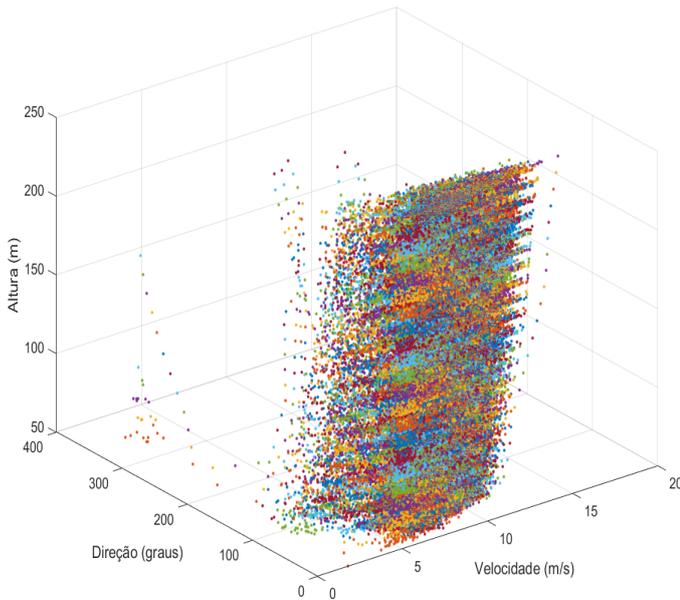


Figura 3. Dados de medição de outubro a dezembro 2021.

Quando se relaciona a altura (Figura 3), estas velocidades se mantêm presentes a partir dos 50 metros até os 250 metros medidos, o que pode indicar que o vento varre toda a circunferência formada pelas pás das turbinas com a mesma velocidade, e isto implicaria em menores esforços na estrutura da turbina e possivelmente menor manutenção.

A partir dos dados ilustrados na Figura 3 foi realizado a simulação com base no *Mean Shift*, com variação do  $\sigma$  em  $n$  interações (quanto mais interações mais cenários são gerados), de forma a gerar o gráfico ilustrado na Figura 5.

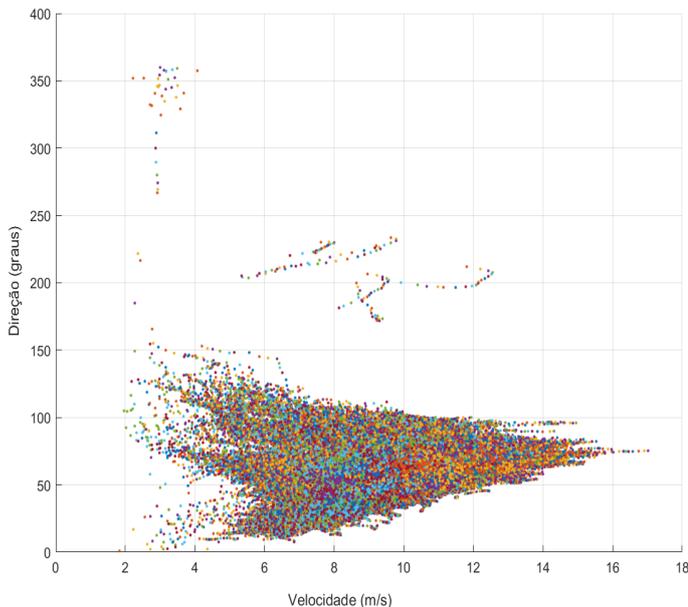


Figura 4. Direção x velocidade.

Na Figura 5 é possível observar os pontos em amarelo que representam o arquivo original (24 horas), após 30 interações, e os cenários gerados que são os pontos em vermelho, fazem uma trajetória gerando novas FDPs até

o centro do conjunto onde se encontra a moda (em verde que representam a última interação do conjunto).

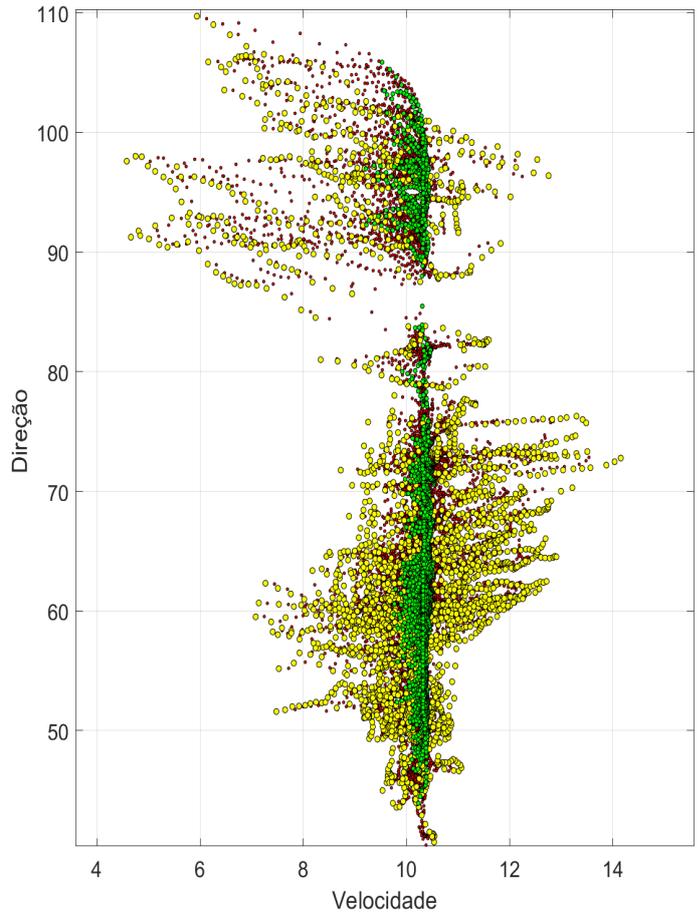


Figura 5. *Mean Shift* após 30 interações.

Para a simulação com o *Mean Shift* foram utilizados alguns meses de medições, considerando as três variáveis principais (velocidade, direção e altura) e a partir de  $n$  cenários gerados foi realizado a validação a curto e médio prazo (1 dia a 1 semana) com o cálculo de erro entre o esperado e o obtido.

A partir dos cenários gerados (30 no total) é calculado a potência eólica e a energia prevista para o dia seguinte e para uma semana a frente. Uma vez que é possível calcular a potência e energia para ventos entre 50 a 250 metros foi observado uma constância de 10 m/s, sendo utilizado este valor como referência, embora seja possível calcular diferentes valores para estas alturas.

$$P = \frac{1}{2} \rho A V^3 \quad (12)$$

Em que  $P$  é a Potência [W],  $\rho$  é a Densidade do ar [kg/m<sup>3</sup>],  $A$  é a Área varrida pelas pás do rotor [m<sup>2</sup>] e  $V$  a Velocidade do vento [m/s].

Desta forma, o algoritmo gera os cenários, retorna os perfis de vento, calcula a potência e energia para cada um dos cenários e compara com o valor real do dia, semana ou mês a frente. Este valor previsto e esperado permite calcular o erro da previsão para cada um dos cenários e definir o pior e melhor caso (Figura 6).

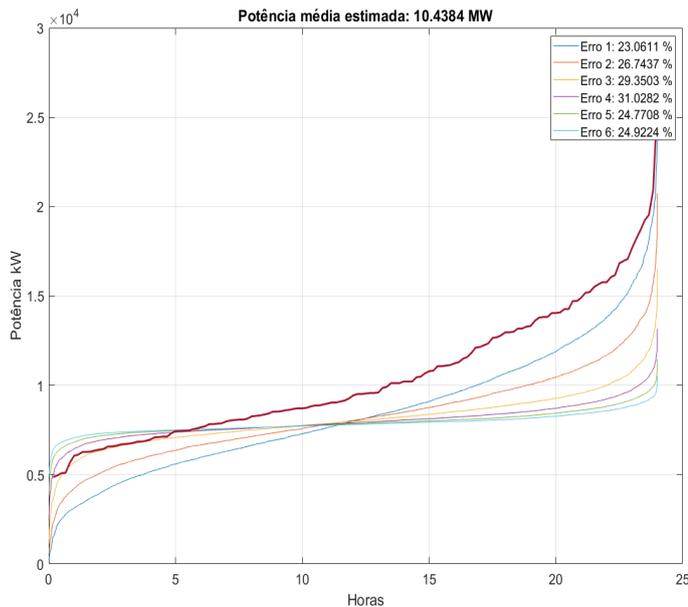


Figura 6. *Mean Shift* para 30 dias a frente.

Ainda na Figura 6, a previsão para 30 dias a frente (1 mês), no qual há seis possíveis cenários com a mesma FDP da base de dados anual. Neste caso a potência média estimada foi de 7.8 MW e os cenários apresentam erros de 24 % a 30 %. Na Figura 7 a projeção é para sete dias a frente (1 semana), no qual o erro obtido foi de 7 % a 15 %.

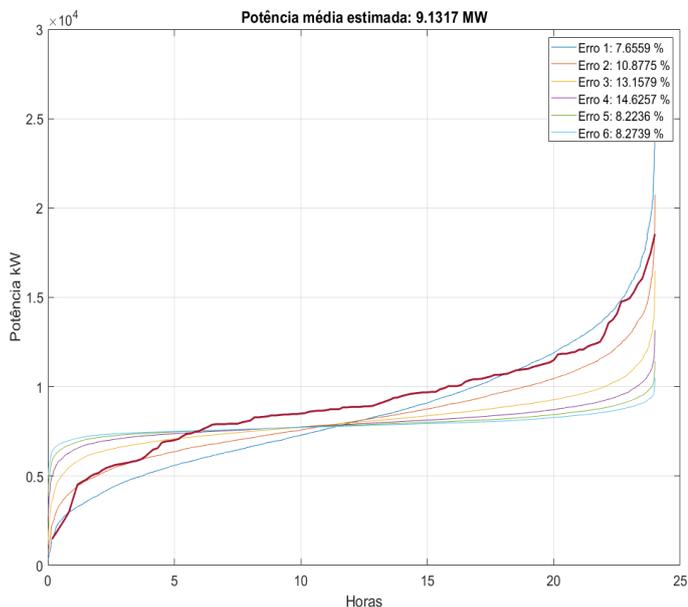


Figura 7. *Mean Shift* para 7 dia a frente.

Já na Figura 8 a previsão foi para 1 dia a frente (24 horas seguinte), no qual os erros variam entre 1 % a 5 %. Estes erros foram calculados tendo como referência a curva em vermelho, que representa o dia seguinte, o sétimo dia ou o trigésimo dia no qual se baseia a previsão. Os dados obtidos sugerem que a melhor taxa de acerto está para o dia seguinte, ou seja, as próximas 24 horas, apresentando um menor acerto quando se deseja obter um horizonte bem mais a frente, caso dos 30 dias.

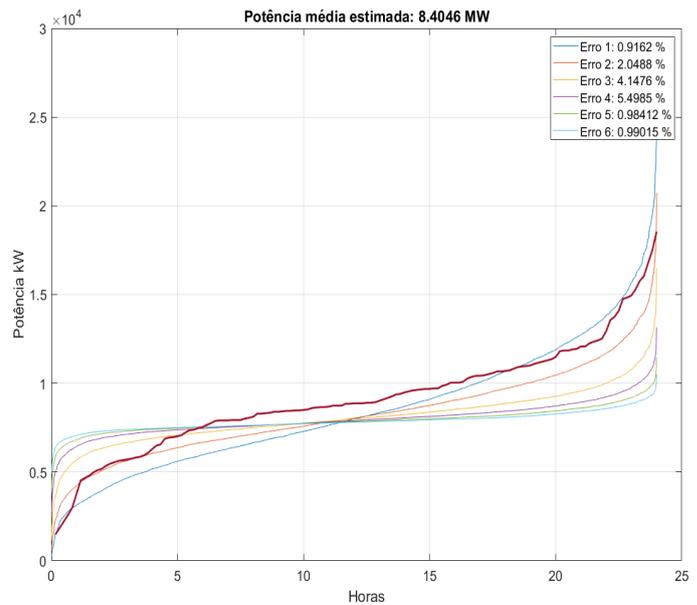


Figura 8. *Mean Shift* para 1 dia a frente.

Para a previsão de geração, no que se refere produção de energia para garantir certa margem de erro, os cenários eólicos projetados baseado na teoria da informação, apresentaram resultados que foram ilustrados em três cenários, ilustrados na Figura 9. O cenário 1 representa a energia gerada 30 dias a frente (em azul) e a energia prevista (em laranja). O erro para o cenário 1 foi de 13 %, para o cenário 2 (7 dias a frente) foi de 6 % e para o cenário 3 (24 horas a frente) foi de 3 %.

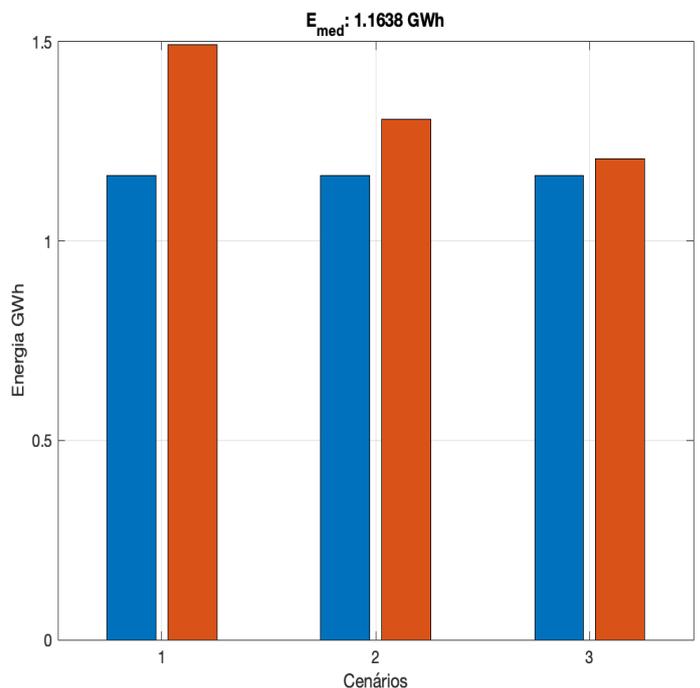


Figura 9. Energia prevista e esperada para 1, 7 e 30 dias a frente.

Logo, para os cenários de energia a previsão é mais eficaz em todos os horizontes permitindo uma gestão assertiva, uma vez que pelo menos a demanda de energia contratada será gerada pela empresa.

## 5. CONCLUSÃO

Como estratégia para determinar os locais mais adequados e com maior potencial eólico, métodos e modelos paramétricos e não paramétricos são empregados para prever o potencial energético de uma região. Este artigo aplicou o algoritmo *Mean Shift* como técnica para densificar uma base de dados de vento, com a finalidade de prever do potencial eólico da região Norte/Nordeste com foco principal no município de São Luís/MA.

Com base na aplicação do algoritmo *Mean Shift* foi possível densificar o conjunto de dados, a partir do original, possibilitando identificar a moda do conjunto. Diante de novos dados gerados, cenários eólicos foram projetados baseados na teoria da informação, apresentando resultados coerentes que viabilizou realizar a previsão do potencial eólico.

## AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - Brasil, a Equatorial Energia, a Gera Maranhão e a CAPES pelo apoio financeiro e cessão dos dados para o estudo.

## REFERÊNCIAS

- ABEEólica (2020). Associação brasileira de energia eólica | boletim anual de geração 2020. Disponível em: <http://abeeolica.org.br/>. Acesso em: 05 de novembro de 2021.
- Abesamis, K., Ang, P., Bisquera, F.I., Catabay, G., Tindogan, P., Ostia, C., and Pacis, M. (2019). Short-term wind power forecasting using structured neural network. In *IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNC-CEM)*, 1–4. IEEE.
- Alencar, D.B.d. et al. (2018). Modelo híbrido baseado em séries temporais e redes neurais para previsão da geração de energia eólica. *Tese. Programa de pós-graduação em Engenharia Elétrica do Instituto de Tecnologia (ITEC) da Universidade Federal do Pará*.
- Chang, G., Lu, H., Hsu, L., and Chen, Y. (2016). A hybrid model for forecasting wind speed and wind power generation. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 1–5. IEEE.
- EPE (2021). Empresa de pesquisa energética | balanço energético nacional: Relatório síntese 2021. Disponível em: <https://www.epe.gov.br/>. Acesso em: 04 de novembro de 2021.
- França, C.G., Faceroli, S.T., Nascimento, P.S.C., Kligerman, A.S., et al. (2020). Análise comparativa de modelos de previsão de geração de energia eólica baseados em machine learning. *Revista de Sistemas e Computação-RSC*, 9(2).
- Haykin, S., Haykin, S., and HAYKIN, S. (1999). *Neural Networks: A Comprehensive Foundation*. International edition. Prentice Hall. URL <https://books.google.com.br/books?id=bX4pAQAAMAAJ>.
- Li, M., Yang, M., Yu, Y., Li, P., Si, Z., and Yang, J. (2020). Wind power probabilistic forecasting based on wind correction using weather research and forecasting model.

- In *2020 IEEE 3rd Student Conference on Electrical Machines and Systems (SCEMS)*, 619–624. IEEE.
- Lima, S.L.d. et al. (2013). Diagnostico de falhas incipientes em transformadores de potência. *Tese. Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Maranhão*.
- Maria, P.H.S.d., Costa, A.A., and Sombra, S.S. (2008). Modelagem numérica em alta resolução para previsão de geração de energia eólica no ceará. *Revista Brasileira de Meteorologia*, 23, 477–489.
- Miranda, V., Castro, A.R.G., and Lima, S. (2012). Diagnosing faults in power transformers with autoassociative neural networks and mean shift. *IEEE Transactions on Power Delivery*, 27(3), 1350–1357.
- Oliveira, F.M.B. et al. (2020). Previsão da velocidade do vento utilizando redes neurais artificiais e modelos autorregressivos. *Dissertação. Programa de Pós-graduação em Sistemas de Energia da Universidade Tecnológica Federal do Paraná*.
- Rêgo, L.P. et al. (2016). Estratégia para predição de consumo de energia elétrica de curto prazo: uma abordagem baseada em densificação com mean shift para tratamento de dias especiais. *Tese. Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Pará*.
- Senatore, A., Fuoco, D., Mendicino, G., Lepore, M., Tozzi, G., and Iorio, P. (2020). A combined modelling system for short-term wind power forecasting based on mesoscale numerical weather prediction. In *2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, 1–5. IEEE.
- Tuzel, O., Porikli, F., and Meer, P. (2009). Kernel methods for weakly supervised mean shift clustering. In *2009 IEEE 12th International Conference on Computer Vision*, 48–55. IEEE.
- Wu, K.L. and Yang, M.S. (2007). Mean shift-based clustering. *Pattern Recognition*, 40(11), 3035–3052.
- Wu, S.H. and Wu, Y.K. (2020). Probabilistic wind power forecasts considering different nwp models. In *2020 International Symposium on Computer, Consumer and Control (IS3C)*, 428–431. IEEE.
- Wu, T., Xia, L., Luo, J., Zhou, X., Hu, X., Ma, J., and Song, X. (2018). Computationally efficient mean-shift parallel segmentation algorithm for high-resolution remote sensing images. *Journal of the Indian Society of Remote Sensing*, 46(11), 1805–1814.
- Yan, J., Zhang, H., Liu, Y., Han, S., Li, L., and Lu, Z. (2017). Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping. *IEEE Transactions on Power Systems*, 33(3), 3276–3284.