

Diogo Friggo

Estimativa da produção de energia de um parque eólico por meio de modelo estocástico

Brasil

2019

Diogo Friggo

Estimativa da produção de energia de um parque eólico por meio de modelo estocástico

O presente trabalho propõe um modelo que represente a estrutura de séries temporais de velocidade de vento de modo a fazer previsões sobre seu comportamento futuro tendo como objetivo estimar a energia que seria gerada por uma turbina eólica.

Universidade Federal do Rio Grande do Sul – UFRGS

Trabalho de Diplomação em Engenharia Física

Orientador: Carlo Requião

Brasil

2019

Diogo Friggo

Estimativa da produção de energia de um parque eólico por meio de modelo estocástico/ Diogo Friggo. – Brasil, 2019-

58 p. : il. (algumas color.) ; 30 cm.

Orientador: Carlo Requião

Graduação – Universidade Federal do Rio Grande do Sul – UFRGS
Trabalho de Diplomação em Engenharia Física, 2019.

1. energia eólica. 2. processo estocástico. 3. cálculo de Ito. I. Universidade Federal do Rio Grande do Sul. II. Engenharia Física.

Diogo Friggo

Estimativa da produção de energia de um parque eólico por meio de modelo estocástico

O presente trabalho propõe um modelo que represente a estrutura de séries temporais de velocidade de vento de modo a fazer previsões sobre seu comportamento futuro tendo como objetivo estimar a energia que seria gerada por uma turbina eólica.

Brasil
2019

Resumo

Todos os agentes envolvidos no planejamento de um parque eólico precisam determinar, com um certo nível de certeza, quanta energia este poderá gerar. Investidores precisam avaliar o risco associado ao financiamento da construção do parque em relação a um retorno futuro. Produtores precisam garantir um determinado nível de produção de energia com base em acordos mensais com compradores. Operadores de subestações, por outro lado, precisam ter uma estimativa da produção em resolução horária de modo a evitar perdas elétricas e para que possam atender adequadamente a demanda por energia a qual sabe-se que varia ao longo do dia de maneira previsível e conhecida a partir de dados históricos de consumo por uma dada região.

Atualmente existem metodologias bem estabelecidas na indústria para a estimativa do recurso eólico de longo prazo e da geração de energia de longo prazo, geralmente para um horizonte de 20 anos. Elas se baseiam no fato de que, a longo prazo, a distribuição de probabilidade da velocidade do vento aproxima-se de uma distribuição de Weibull ([MENČÍK, 2016](#)) a qual é então convertida em uma distribuição de probabilidade de energia por meio de uma curva de potência de uma determinada turbina. Essa distribuição de densidade de probabilidade é a mais difundida na análise de dados de vento ([P., 2003](#)), tendo sido utilizada para estimar tanto o recurso eólico ([AN., 2003](#)) quanto a produção de energia ([M HOOSHMAND R-A; A-H, 2015](#)).

Tais métodos, no entanto, não se aplicam a estimativas de curto prazo, pois, sob essa resolução temporal, o caráter estocástico do vento predomina não sendo possível reduzir o seu comportamento a uma distribuição de probabilidade estática. É necessário recorrer a outra abordagem para tal estimativa. O presente trabalho propõe a modelagem estocástica do processo físico que rege o comportamento do vento por meio de diferentes métodos, conhecidos e aplicados com sucesso na modelagem de outros processos estocásticos, em particular na área financeira. Tal modelagem permite obter uma estimativa da geração de energia em tanto em resolução mensal quanto horária.

Palavras-chave: energia eólica, processo estocástico, modelos autoregressivos

Listas de ilustrações

| | |
|---|----|
| Figura 1 – Comparação entre a composição da matriz energética brasileira no ano de 2017 (esquerda) e no ano de 2018 (direita) | 10 |
| Figura 2 – Ranking de capacidade eólica instalada acumulada | 11 |
| Figura 3 – Evolução da capacidade instalada de parques eólicos no Brasil | 11 |
| Figura 4 – Curva de potência de uma turbina genérica. Fonte: Wikimedia Commons | 13 |
| Figura 5 – À esquerda velocidade do vento em m/s para uma região no sul do Ceará e à direita o preço de fechamento das ações da google para 120 dias de 2013. Fonte: autoria própria, dados (ECMWF, 2017). | 14 |
| Figura 6 – Chapada | 16 |
| Figura 7 – Variação da elevação na direção preferencial de escoamento do vento. . | 17 |
| Figura 8 – Alguns nós da série ERA 5 ao sul do Ceará, Brasil. Google earth V 7.3.2.5776. (14 de Dezembro, 2015). Rio Grande do Sul, Brasil. $7^{\circ} 14' 11,54''$ S, $40^{\circ} 04' 12,31''$ W, Eye alt 1808,45 km. | 20 |
| Figura 9 – Velocidade do vento registrada por satélite na região de interesse nos anos de 2017 e 2018 | 21 |
| Figura 10 – Histograma de velocidades do nó noroeste da série de dados modelo. . | 22 |
| Figura 11 – Gráfico de Cullen e Frey para os dados do nó noroeste da série de dados modelo. | 23 |
| Figura 12 – Rosa dos ventos | 24 |
| Figura 13 – Chapada | 24 |
| Figura 14 – Chapada | 25 |
| Figura 15 – Chapada | 33 |
| Figura 16 – Chapada | 35 |
| Figura 17 – Evolução dos parâmetros p e q de um modelo ARIMA ao longo do tempo. | 36 |
| Figura 18 – Evolução dos parâmetros p e q de um modelo ARIMA ao longo do tempo. | 37 |
| Figura 19 – Localização da região de medição por satélite dos dados de vento discutidos no texto. Fonte: SIO, NOAA, U.S. Navy, NGA, GEBCO. © 2018 Google | 39 |
| Figura 20 – Rosa dos ventos do recurso eólico da cidade de Sento Sé no norte da Bahia - característico do Nordeste brasileiro, a partir de dados medidos para o ano de 2017. Fonte: autoria própria, dados (ECMWF, 2017). . | 40 |
| Figura 21 – Rosas dos ventos mensais do recurso eólico da cidade de Sento Sé no norte da Bahia a partir de dados medidos para o ano de 2017. Fonte: autoria própria, dados (ECMWF, 2017). | 41 |

| | |
|--|----|
| Figura 22 – Velocidade do vento para cada mês do ano de 2017 da cidade de Sento Sé no norte da Bahia. Curva em preto: dados medidos. Curva em azul: curva ajustada aos dados medidos. Fonte: autoria própria, dados (ECMWF, 2017). | 42 |
| Figura 23 – Velocidade do vento para cada hora do dia e para cada mês do ano de 2017 para a cidade de Sento Sé no norte da Bahia. Fonte: autoria própria, dados (ECMWF, 2017). | 43 |
| Figura 24 – Valor de abertura das ações da Google na NASDAQ. Curva em preto: dados medidos. Curva em azul: curva ajustada aos dados medidos. Fonte: autoria própria, dados (FINANCE, 2018). | 46 |
| Figura 25 – A distribuição de velocidades pode ser aproximada por uma distribuição de Weibull. Dados do quadrante central dos dados ERA5 (ECMWF, 2017) para a cidade de Sento Sé no norte da Bahia. Fonte: autoria própria, dados (ECMWF, 2017). | 48 |
| Figura 26 – Distribuição de velocidades de todos os quadrantes dos dados ERA5 (ECMWF, 2017) para a cidade de Sento Sé no norte da Bahia. Fonte: autoria própria, dados (ECMWF, 2017). | 49 |
| Figura 27 – Gráfico de Gantt do planejamento das etapas de desenvolvimento do presente trabalho. Fonte: autoria própria. | 56 |

Lista de abreviaturas e siglas

| | |
|--------|---|
| ECMWF | European Center for Medium-Range Weather Forecast |
| NASDAQ | National Association of Securities Dealers Automated Quotations |
| ARMA | Autoregressive Moving-Average |

Sumário

| | |
|--|-----------|
| Introdução | 10 |
| I ESTUDO DE CASO | 15 |
| 1 CARACTERIZAÇÃO DA REGIÃO | 16 |
| 2 A SÉRIE TEMPORAL MODELO | 19 |
| 3 CARACTERIZAÇÃO DOS DADOS | 21 |
| 3.1 Distribuição de probabilidades | 22 |
| 3.2 Sazonalidade | 23 |
| 4 TIME SERIES ANALYSIS | 26 |
| 5 REGIME ESTACIONÁRIO | 27 |
| 6 JANELA DE DADOS | 28 |
| 7 MENSURANDO A QUALIDADE DA PREVISÃO | 29 |
| 8 MODELO AUTOREGRESSIVO (AR) | 30 |
| 9 FORECASTING: PRINCIPLES AND PRACTICE | 31 |
| 9.1 Horizonte de previsão | 31 |
| 9.2 Métodos | 31 |
| 9.3 Análise visual | 32 |
| 9.4 Medidas de qualidade de previsões | 33 |
| 9.4.1 Erro médio absoluto (MAE) | 34 |
| 9.4.2 Raíz do quadrado da média do erro (RMSE) | 34 |
| 9.4.3 Erro absoluto médio escalonado (MASE) | 34 |
| 9.4.4 Validação cruzada | 34 |
| 9.5 ARIMA | 35 |
| 9.6 Box-Jenkins | 35 |
| 9.7 Modelo Autoregressivo Variável | 35 |
| 9.8 GARCH | 37 |

| | | |
|------------|--|-----------|
| II | ANÁLISE DE DADOS DE VENTO | 38 |
| 10 | CARACTERÍSTICAS DE LONGO PRAZO | 39 |
| 11 | CARACTERÍSTICAS DE CURTO PRAZO | 42 |
| III | EMBASAMENTO TEÓRICO | 44 |
| 12 | EMBASAMENTO MATEMÁTICO | 45 |
| 13 | EMBASAMENTO FÍSICO | 47 |
| 13.1 | Restrições físicas | 47 |
| 13.2 | Abordagem de longo prazo | 48 |
| 13.3 | Desvantagens | 50 |
| 13.4 | Ressalvas | 50 |
| IV | REVISÃO BIBLIOGRÁFICA | 52 |
| V | PLANEJAMENTO | 54 |
| | REFERÊNCIAS | 57 |

Introdução

Ao contrário de fontes de energia tradicionais tais como petróleo ou carvão, o vento é uma fonte de energia limpa e renovável. A energia eólica é, em princípio, inesgotável, pois é função apenas da energia cinética das moléculas presentes na atmosfera, as quais, em última instância, estão em movimento, principalmente, devido a radiação solar e atração gravitacional entre corpos massivos que rege a dinâmica de movimento da Terra e consequentemente sua atmosfera. Ser inesgotável não significa que seja infinita. Existe um limite de quanta energia se pode extrair da atmosfera, no entanto, por ser renovável, essa energia é constante e naturalmente reabastecida.

Diante da finitude das principais fontes de energia em uso no mundo, a geração de energia eólica, assim como de energia solar, tem crescido continuamente ao redor do mundo. Na imagem abaixo é possível constatar a tendência de diminuição do uso de fontes de energia não renováveis e aumento das renováveis, principalmente eólica e solar. Em 2017, na imagem a esquerda, a energia eólica era responsável por 8,1% da matriz energética do país. Em 2018 esse número subiu para 9% enquanto que quase todas as outras fontes recuaram.

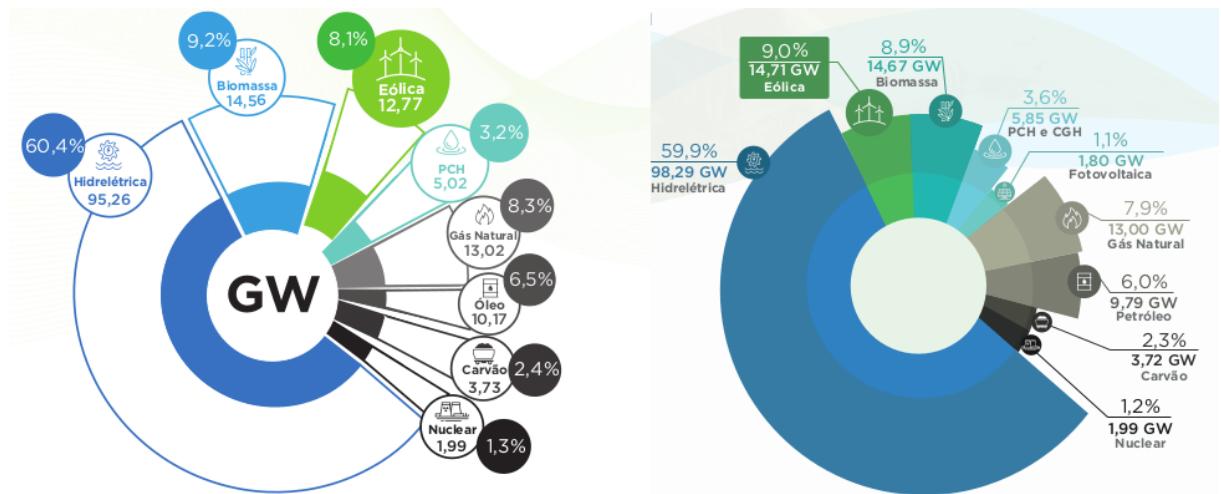


Figura 1 – Comparação entre a composição da matriz energética brasileira no ano de 2017 (esquerda) e no ano de 2018 (direita)

O Brasil conta com uma capacidade instalada de energia eólica de 12,76 GW em seus mais de 500 parques eólicos e figura mundialmente como 8º maior produtor, superando países desenvolvidos como o Canadá.

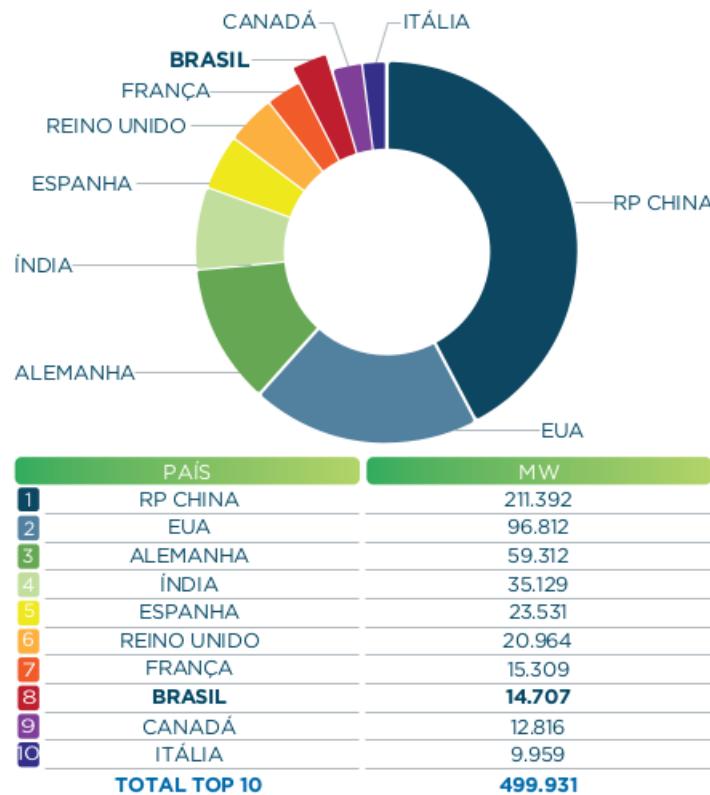


Figura 2 – Ranking de capacidade eólica instalada acumulada

O crescimento da captação de energia eólica no país vem crescendo constatadamente, adicionando em média 2GW de capacidade instalada (máxima capacidade de produção) a cada ano desde 2013. Hoje essa fonte de energia é responsável, em média, por 11% do abastecimento do país e 60% do Nordeste.

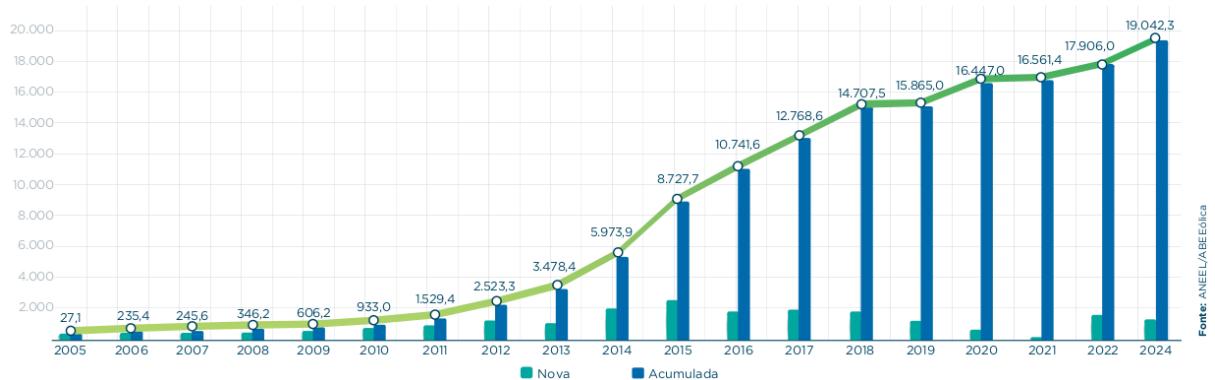


Figura 3 – Evolução da capacidade instalada de parques eólicos no Brasil

A principal fonte de energia elétrica do país vêm de hidrelétricas. Em períodos de estiagem, essa fonte é complementada por termoelétricas, as quais usam combustíveis fósseis para gerar energia. O seu alto custo tem impacto direto no consumidor final que

paga mais pela mesma energia.

A variabilidade do recurso eólico representa um dos maiores desafios ao seu aproveitamento. A previsão de que a maré será alta ou baixa é aceita como verdade sem mais questionamentos, enquanto que a previsão de que, devido ao movimento de massas de ar, provavelmente irá chover é encarada com forte suspeita, pois, ao contrário de outras fontes naturais de energia, tais como a energia hídrica ou solar, a energia eólica apresenta grande variabilidade (ACKERMANN, 2012), sendo, portanto, de maior dificuldade a sua previsão. Dessa forma, é de grande interesse que se entenda o seu comportamento e se desenvolva métodos para sua previsão.

Os principais fatores que influem na geração de energia eólica são a densidade do ar e a velocidade do vento. A potência (em watts) gerada por uma turbina é dada pela equação (CEPEL, 2001) a qual é fornecida por fabricantes de turbinas eólicas:

$$P = \frac{1}{2} \rho \frac{\pi D^2}{4} \nu^3 C_p \eta \quad (1)$$

P = potência elétrica na altura do cubo rotor [W]

ρ = densidade do ar $\left[\frac{kg}{m^3} \right]$

D = diâmetro do rotor [m]

ν = velocidade do vento $\left[\frac{m}{s} \right]$

C_p = coeficiente aerodinâmico de potência do rotor [W]

η = eficiência do conjunto gerador/transmissão

O fato de a potência variar com o cubo da velocidade do vento evidencia a importância da determinação precisa dessa grandeza.

A conversão entre velocidade e potência é feita por uma curva de potência fornecida pelo fabricante da turbina utilizada. Ela representa a potência que seria gerada para cada velocidade incidente perpendicularmente às pás da turbina. Uma vez feita a previsão da velocidade do vento utilizar-se-a, neste trabalho, uma curva de potência genérica como a exibida abaixo para obter a previsão em energia ao invés de velocidade do vento. Observa-se que baixas velocidades de vento resultam em baixa produção de energia e que a velocidades muito altas a produção é constante. Para velocidades muito intensas a turbina é temporariamente desligada para evitar danos estruturais.

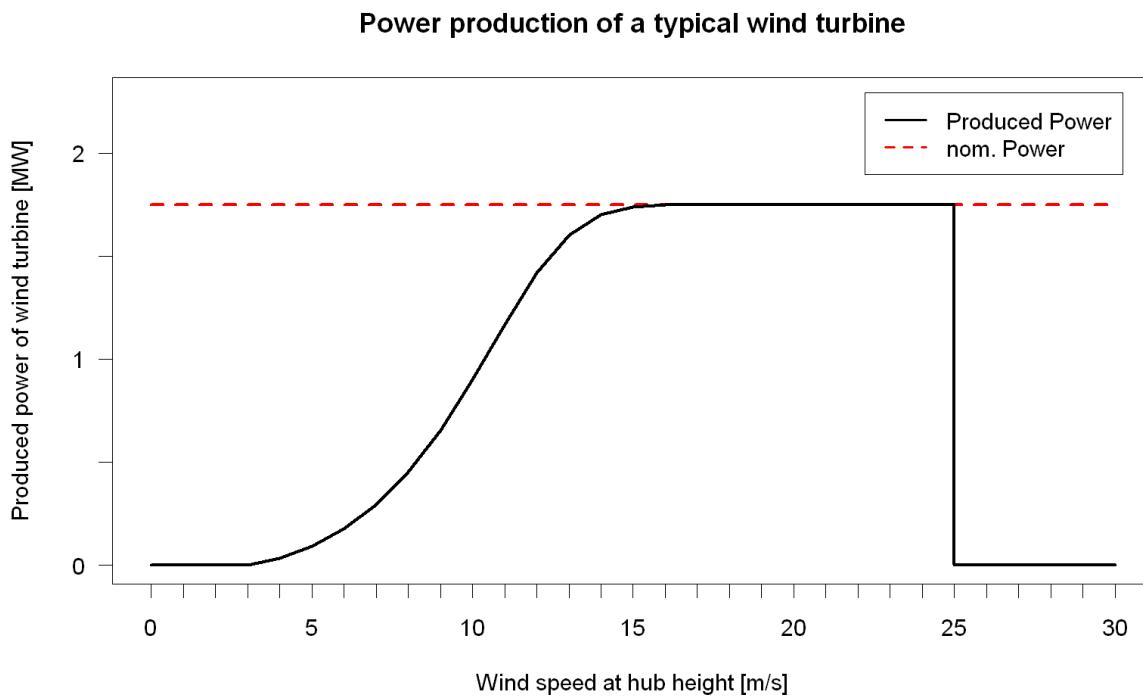


Figura 4 – Curva de potência de uma turbina genérica. Fonte: [Wikimedia Commons](#)

Na equação 1, exceto pela velocidade do vento, os demais parâmetros são controlados ou apresentam pouca variabilidade como a densidade do ar, por exemplo, que é relativamente constante para um dado macroclima. A velocidade do vento, por outro lado, tem caráter fortemente estocástico, ou seja, seu comportamento está associado a um processo aleatório. A sua previsão demanda o uso de modelos matemáticos que levem em conta tal comportamento.

Fenômenos estocásticos são comuns na natureza: o movimento errático que uma partícula macroscópica sofre ao ser imersa num fluído (composto por partículas microscópicas) conhecido como movimento Browniano; o decaimento radiativo de átomos em que não se sabe em qual momento dado átomo emitirá radiação (conhece-se apenas uma taxa característica de emissão).

Exemplos associados a atividade humana também são comuns sendo a evolução temporal do valor de ativos econômicos o exemplo mais marcante e para cujo entendimento se desenvolveu um arcabouço matemático sofisticado. Esse conhecimento pode ser utilizado para estudar quaisquer outros fenômenos estocásticos. A imagem traz dois exemplos de séries estocásticas: uma série de dados de vento (esquerda) e uma série de dados financeira (direita). Os métodos desenvolvidos para estudar séries financeiras podem ser usados com sucesso para séries de velocidade do vento.

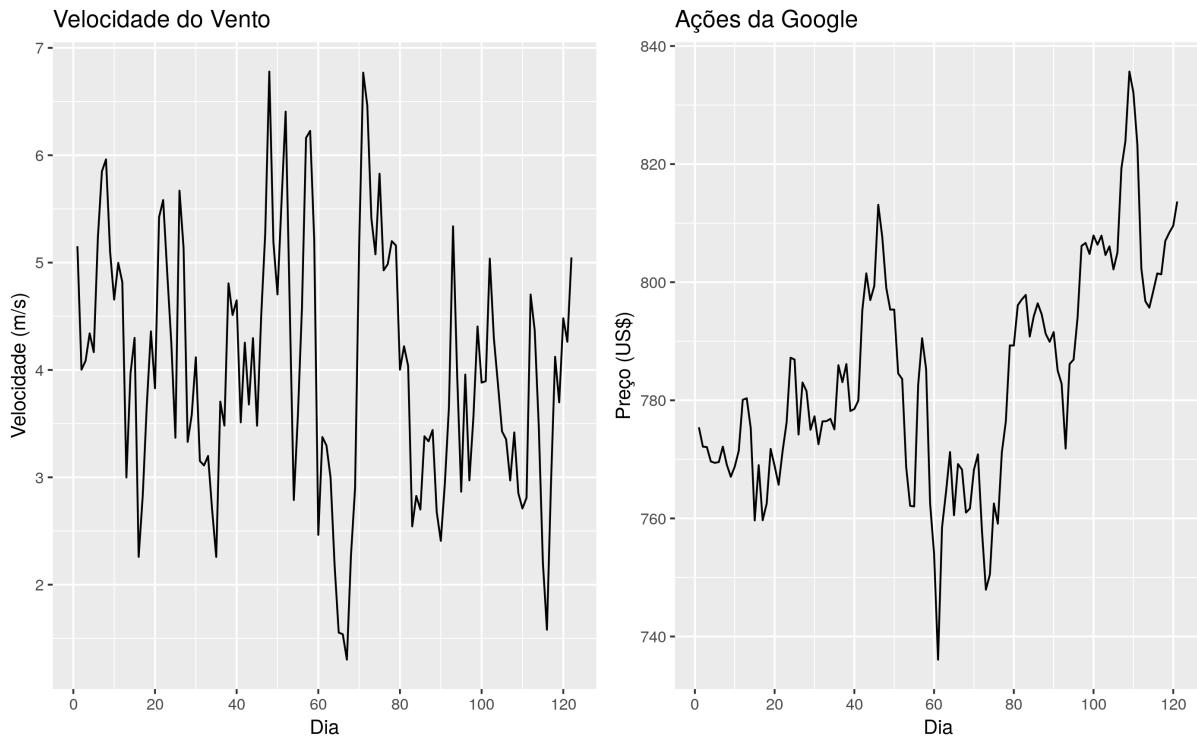


Figura 5 – À esquerda velocidade do vento em m/s para uma região no sul do Ceará e à direita o preço de fechamento das ações da google para 120 dias de 2013. Fonte: autoria própria, dados ([ECMWF, 2017](#)).

Yule e Walker representaram uma grandeza estocástica como resultado de choques aleatórios descritos por um filtro linear aplicado em ruído branco:

$$v_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

Essa ideia dá origem a modelos autoregressivos que descrevem a estrutura de séries temporais levando em conta termos anteriores da própria série. Box e Jenkins descrevem esse método como "deixar que os dados falem por eles próprios", ou seja, nenhuma variável exógena é empregada, apenas a própria série. Esse método tem sido empregado com sucesso para prever o comportamento futuro de séries temporais em diversas áreas. Diversas variações do modelo original foram desenvolvidas ao longo dos anos. O conjunto de todas essas variações são denominadas modelos ARIMA, um acrônimo do inglês *autogressive integrated moving average*, que significa modelo autogressivo (AR) integrado (I) de média móvel (MA).

Parte I

Estudo de caso

1 Caracterização da região

De modo a desenvolver as ideias desse trabalho escolheu-se arbitrariamente uma região localizada no estado do Ceará. Optou-se pela região nordeste por ser a região onde há mais investimentos em energia eólica e que responde pela maioria dos parques em operação no país. Os métodos desenvolvidos são, no entanto, aplicáveis a qualquer região, pois têm como entrada apenas uma série temporal.

A região de interesse encontra-se numa chapada elevada, um local com topografia muito favorável a instalação de um parque eólico sabe-se da dinâmica de fluídos que o escoamento é acelerado quando exposto a um diferencial de elevação. Na imagem abaixo pode-se perceber o quanto elevada a chapada é em relação ao seu redor.

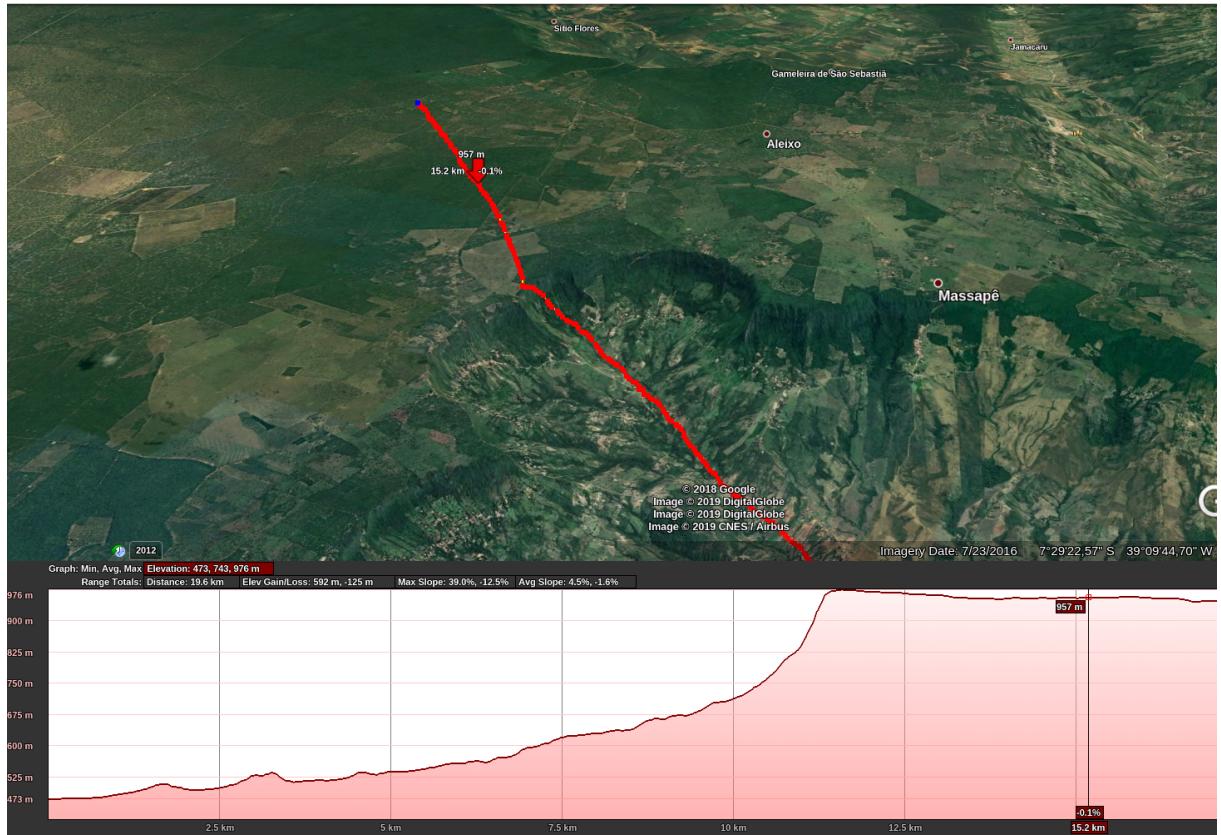


Figura 6 – Chapada

Detalhes da topografia da região pode ser visualizados na imagem abaixo a qual exibe em detalhes a elevação por meio de linhas de contorno além dos declives e a rosa dos ventos. Uma região plana facilita as simulações computacionais necessárias à extração do recurso eólico nas posições de torres de medição às posições de cada turbina.

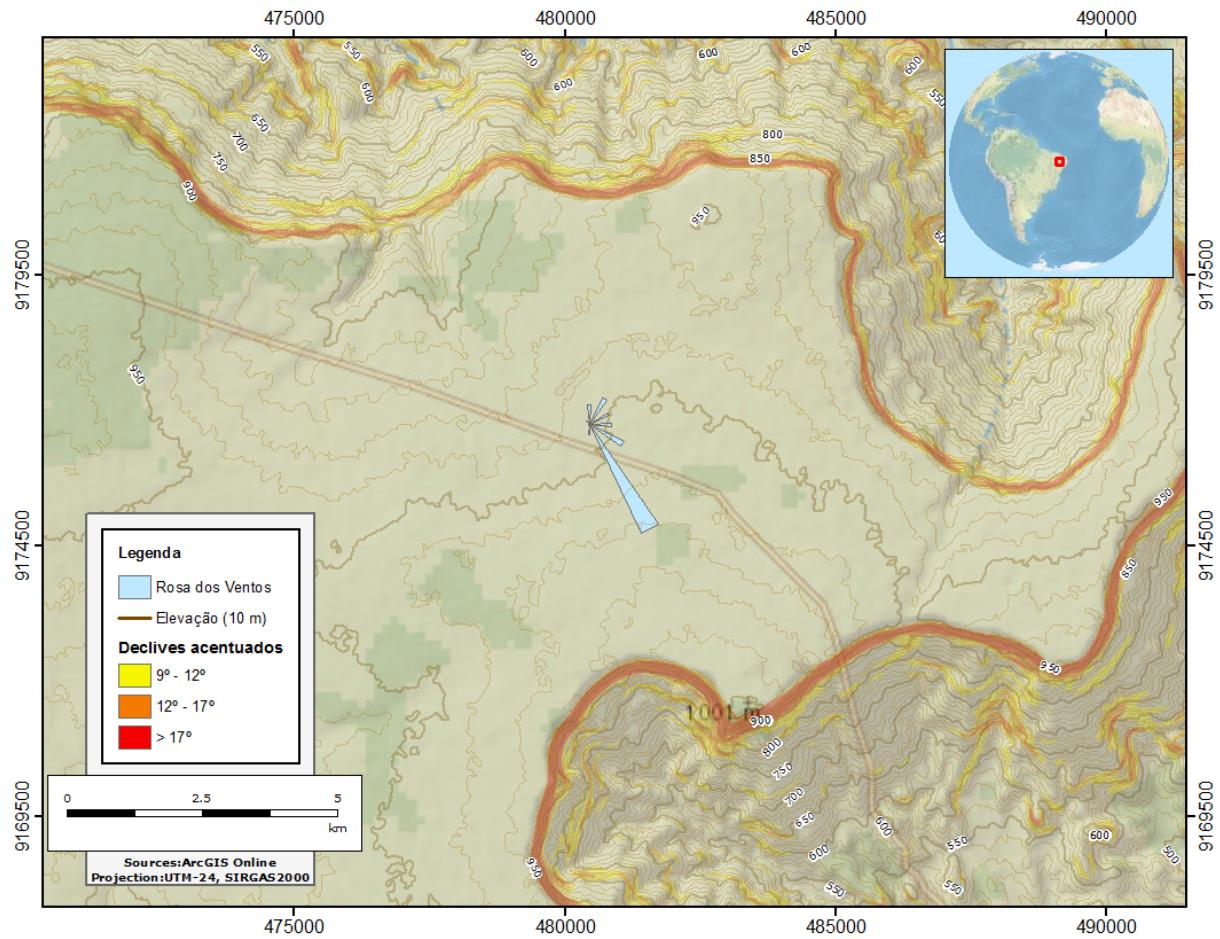


Figura 7 – Variação da elevação na direção preferencial de escoamento do vento.

Na camada da atmosfera próxima ao solo, a velocidade do vento aumenta com a altura conforme uma lei de potência empírica:

$$\frac{v}{v_r} = \left(\frac{h}{h_r} \right)^\alpha \quad (1.1)$$

v = velocidade do vento à altura h

v_r = velocidade do vento à altura de referência h_r

α = coeficiente de cisalhamento, parâmetro estimado a partir dos dados

Em suma, a partir da imagem acima, é possível determinar as melhores posições para se colocar turbinas de um parque eólico:

- Na chapada, onde a velocidade do vento é maior devido a sua elevação de acordo com a lei de potência

- Nas zonas de baixa declividade para facilitar as simulações computacionais
- Nas bordas devido ao diferencial de elevação que causa aceleração do vento
- Orientadas a aproximadamente 150° em relação do norte verdadeiro de modo a captar o vendo oriundo de sua direção preferencial

2 A série temporal modelo

Para exemplificar a elaboração do modelo proposto, utilizou-se uma série de dados disponibilizada publicamente pela organização européia ECWMF. Essa série conta com uma vasta gama de grandezas físicas medidas por satélite oriundas de diversas fontes. Esses dados são agregados e tratados de modo a melhorar sua qualidade. Existem diversas outras séries de dados de satélite com o mesmo propósito, tal como a série ERA-Interim (antecessora da ERA 5) ou as séries MERRA e MERRA 2 produzidas pela NASA. No entanto, observa-se que a qualidade da fonte ERA 5 é muito superior. Isso é empiricamente constatado devido a excelente correlação com dados de torres de medição instaladas no solo. Essa melhoria deve-se tanto a melhorias técnicas na assimilação e agregação de dados quanto a maior resolução espacial da série ERA 5: enquanto MERRA e MERRA 2 apresentam uma resolução espacial de 50 km em latitude e longitude, ERA 5 apresenta 30 km.

A grandeza de interesse para esse trabalho é a velocidade do vento a 100 m de altura em relação ao solo. Essa é a altura disponibilizada pela fonte ERA 5. Outras alturas estão disponíveis mas são obtidas por interpolação e portanto são de qualidade inferior. Essa é uma altura adequada pois a altura do cubo rotor de aerogeradores comerciais varia tipicamente entre 80 m e 130 m.

A série ERA 5, assim como as outras mencionadas, possuem dados para todo o globo, dividindo-o em uma rede de nós com resolução de 30 km x 30 km em latitude e longitude respectivamente. A imagem abaixo exemplifica como esses nós são dispostos na região de interesse.

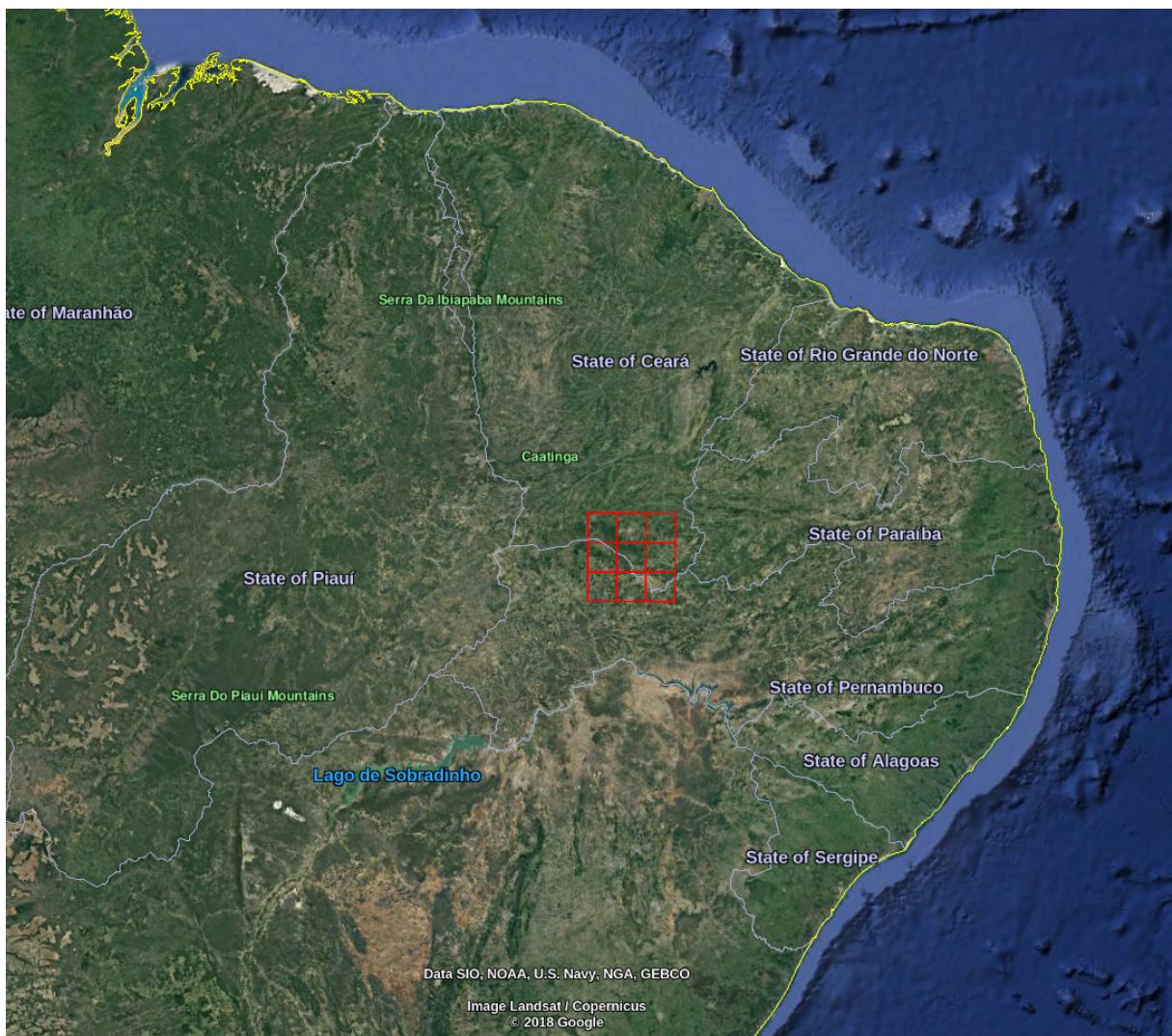


Figura 8 – Alguns nós da série ERA 5 ao sul do Ceará, Brasil.

Google earth V 7.3.2.5776. (14 de Dezembro, 2015). Rio Grande do Sul, Brasil.
 $7^{\circ} 14' 11,54''$ S, $40^{\circ} 04' 12,31''$ W, Eye alt 1808,45 km.

Cada nó cobre uma vasta região. A velocidade reportada para cada nó é uma média das velocidades da respectiva região. A campanha de medição necessária que antecede a construção de um parque eólico exige a instalação de torres de medição no local. Desse modo, a resolução dessas medições é muito maior do que a oferecida pela série ERA 5. No entanto, os dados de medições de torres não são disponibilizados publicamente. Com base na excelente correlação entre as séries ERA 5 e séries medidas por torres de projetos, acredita-se que o procedimento exposto neste trabalho possa ser aplicado com sucesso em dados medidos por torres em solo ou na nacelle dos próprios aerogeradores.

Os dados são fornecidos em base horária e compreendem o período de Janeiro de 2000 a Janeiro de 2019.

3 Caracterização dos dados

O começo de qualquer análise de séries temporais se dá pelo gráfico dos valores que assume ao longo do tempo. Por meio desse gráfico é possível identificar qualitativamente tendências, sazonalidade, ciclicidade, valores atípicos, caráter estacionário e comportamento da variância:

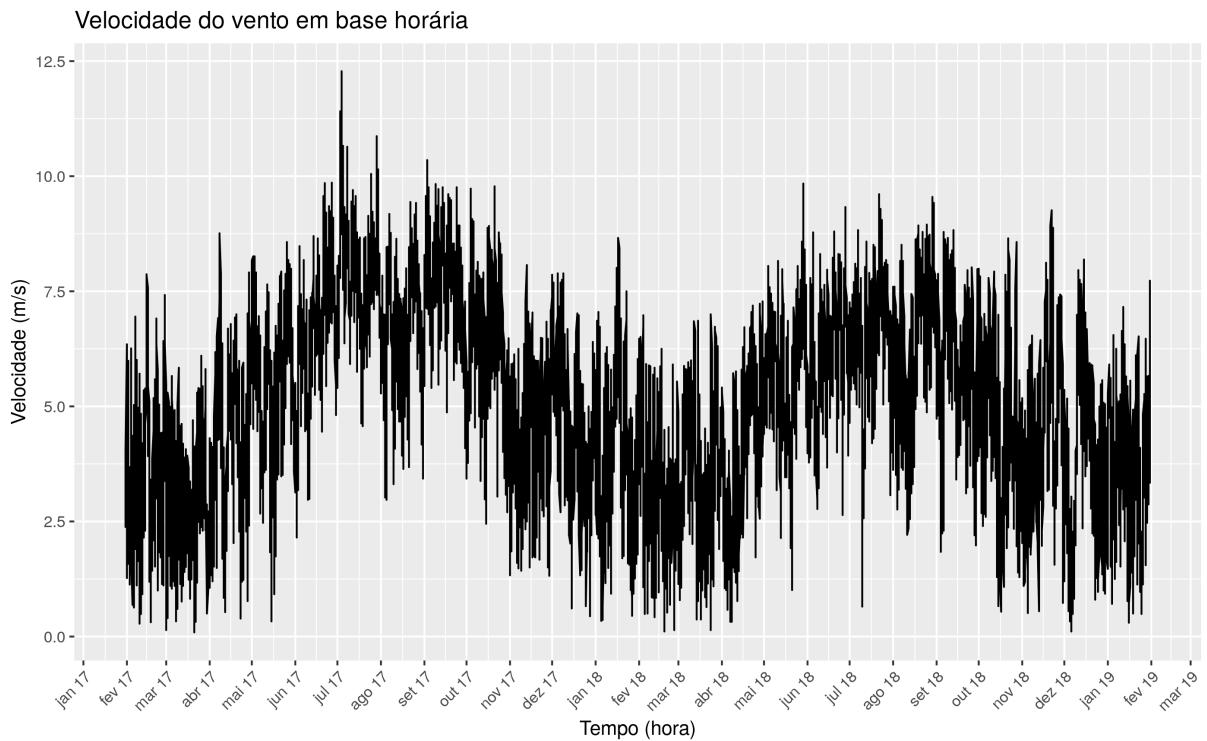


Figura 9 – Velocidade do vento registrada por satélite na região de interesse nos anos de 2017 e 2018

Apenas pela análise visual é possível perceber que a velocidade tem máximos os meses de inverno e mínimos nos de verão, que há sazonalidade anual. A série não aparenta ter derivada, oscila ao longo de um valor médio. Por apresentar sazonalidade a série não é estacionária. Como os métodos apresentados neste trabalho tomam como premissa que a série de entrada é estacionária, é necessário transformá-la previamente em uma série estacionária e ser capaz de fazer o caminho reverso após a sua previsão. Como discutido em REF isso pode ser feito por uma transformação de Box-Cox. Nessa escala não é possível afirmar algo conclusivo sobre a variância da série.

3.1 Distribuição de probabilidades

O histograma de velocidades exibido abaixo indica que a série não segue uma distribuição normal e que uma distribuição de cauda longa à direita seria mais adequada.

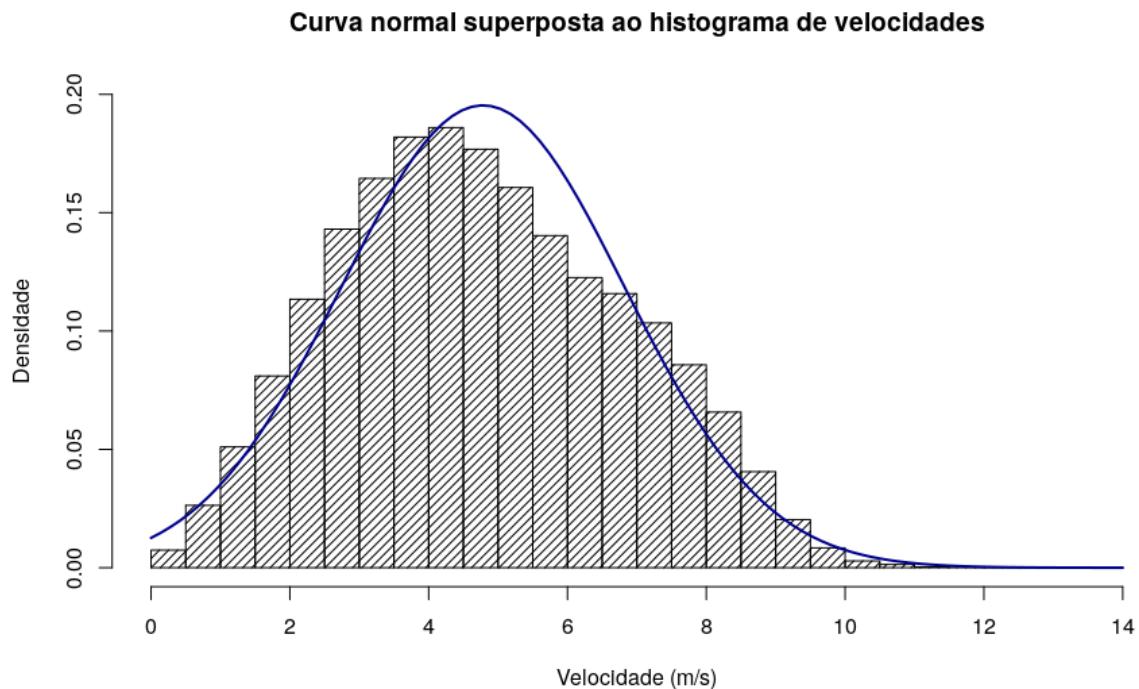


Figura 10 – Histograma de velocidades do nó noroeste da série de dados modelo.

Por meio do gráfico de Cullen e Frey é possível afirmar que a distribuição de probabilidades que melhor descreve os dados se aproxima de uma distribuição beta:

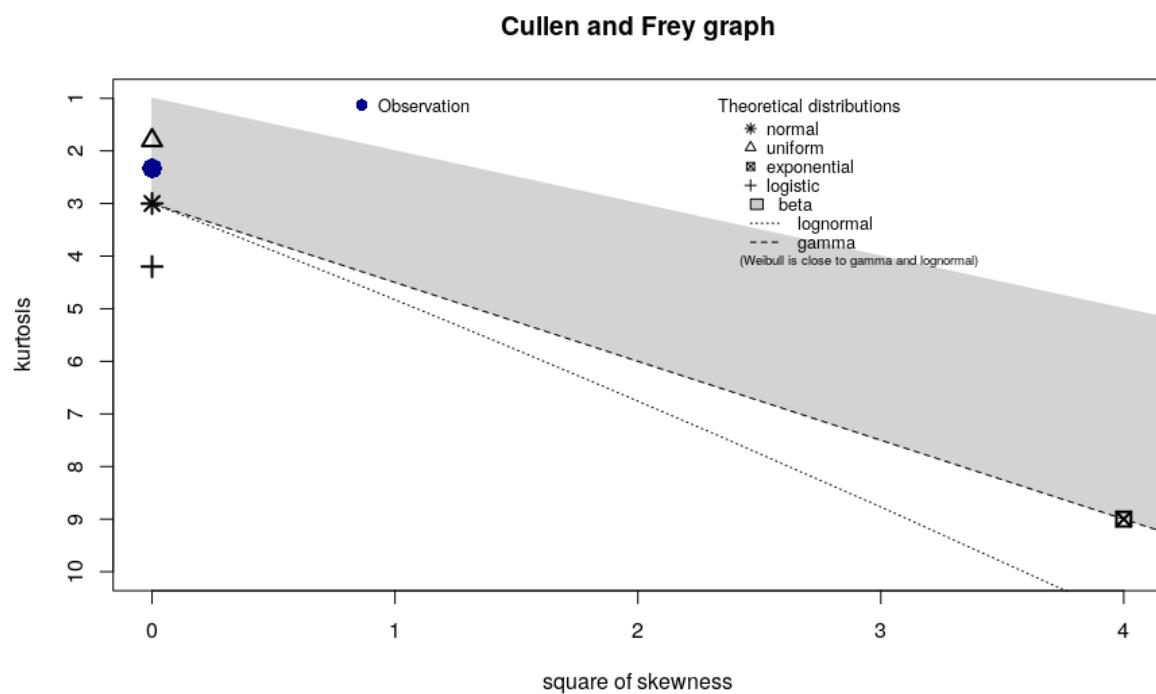


Figura 11 – Gráfico de Cullen e Frey para os dados do nó noroeste da série de dados modelo.

3.2 Sazonalidade

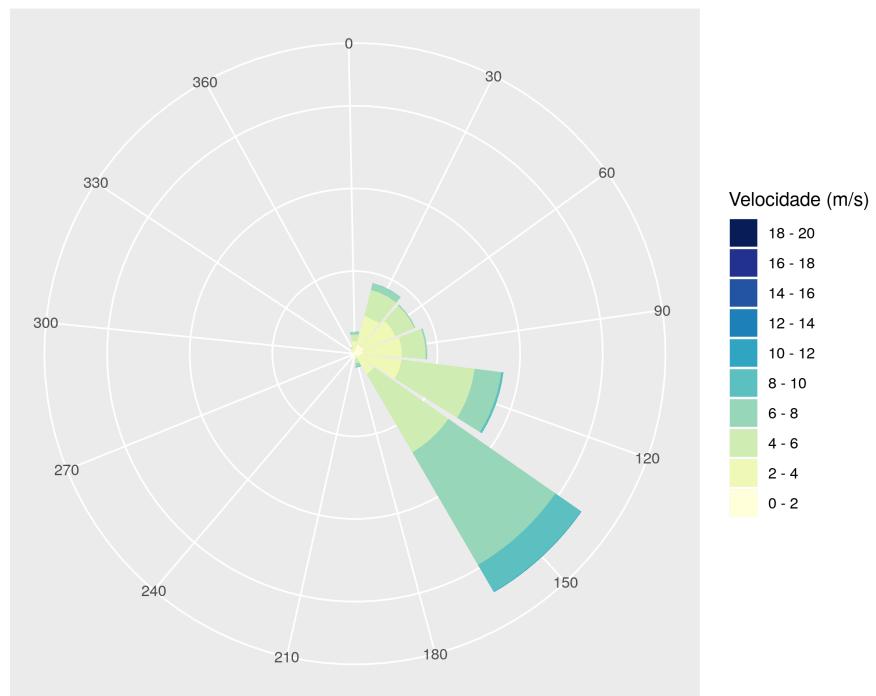


Figura 12 – Rosa dos ventos

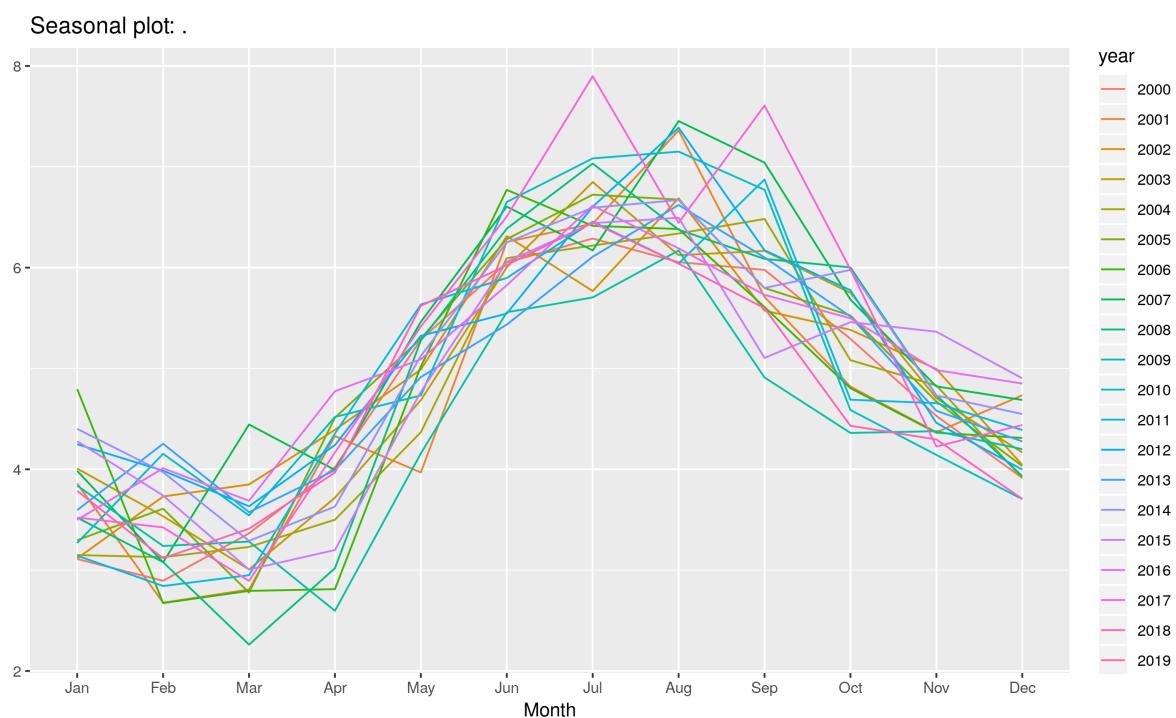


Figura 13 – Chapada

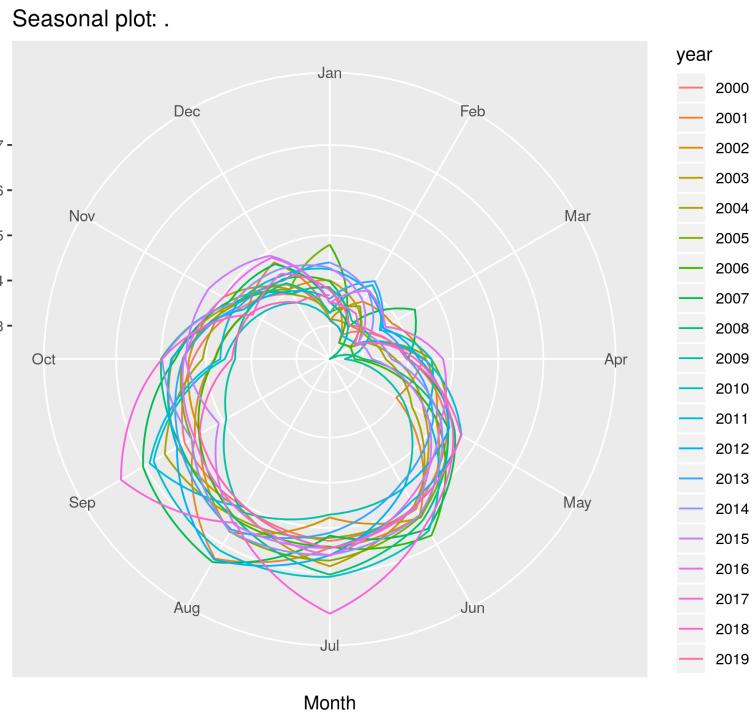


Figura 14 – Chapada

4 Time Series Analysis

Este trabalho faz uso de um modelo linear estocástico que tem como hipótese que a série temporal sob análise é gerada por uma combinação de choques aleatórios. Essa hipótese foi inicialmente proposta por Yule e Walker e extendida por diversos outros autores. Na literatura esse modelo é conhecido como modelo autoregressivo com média móvel (ARMA, do inglês autoregressive moving-average)

Um processo estocástico pode ser visto como o resultado de um filtro linear aplicado em ruído branco a_t . rodapé: Ruído branco é definido como um sinal aleatório o qual tem a mesma intensidade para qualquer frequência [fonte].

$$\begin{aligned} z_t &= \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ z_t &= \mu + a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} \end{aligned}$$

Ruído branco tem média zero e variância constante.

5 Regime estacionário

Um processo estocástico é dito estritamente estacionário se o deslocamento da origem temporal não altera suas propriedades, ou seja, a distribuição conjunta de probabilidades calculada para o sequência de m medições z_1, z_2, \dots, z_m tomadas nos tempos t_1, t_2, \dots, t_m é a mesma àquela calculada em $t_{1+k}, t_{2+k}, \dots, t_{m+k}$ para as m medições $z_{1+k}, z_{2+k}, \dots, z_{m+k}$. $k \in \mathbb{Z}$ pode assumir tanto valores positivos quanto negativos, isto é, o deslocamento temporal pode ser tanto positivo quanto negativo. Dessa forma, um processo estacionário é caracterizado por uma distribuição de probabilidades que não varia no tempo. O modelo autoregressivo com média móvel utilizado neste trabalho toma essa propriedade como hipótese para o seu desenvolvimento, portanto, se faz fundamental garantir que a série de entrada satisfaça tal hipótese.

6 Janela de dados

7 Mensurando a qualidade da previsão

8 Modelo Autoregressivo (AR)

9 Forecasting: Principles and Practice

Qualquer quantidade que seja observada cronologicamente é dita uma série temporal. Exemplos incluem: o preço diário das ações da Google, a quantidade mensal de chuva no Rio Grande do Sul, a produção anual de vinhos na serra gaúcha. Neste trabalho a série de interesse é de velocidade média do vento a 100 m de altura do solo no sul do estado do Ceará. Os dados são coletados por satélite e são disponibilizados em base horária no fuso horário GMT.

9.1 Horizonte de previsão

O horizonte de previsão é variável. Ele depende do propósito para o qual a previsão será utilizada. Alguns parques eólicos no país fazem acordos mensais sobre a energia que será entregue a rede. Nesse caso é de grande utilidade uma previsão de quanta energia será produzida no mês seguinte. A previsão minutal ou horária é, nesse caso, pouco relevante. Ao operador de subestação de distribuição de energia interessa saber quando ocorreram máximos e mínimos de produção de energia em base horária de modo que o sistema possa compensar as faltas e excessos sem perdas. Para ele a previsão minutal ou horária é essencial. Neste trabalho abordou-se tanto a previsão em escala horária quanto mensal.

Quanto mais próximo do último dado medido for a previsão mais certeza se tem do seu valor. A previsão da velocidade do vento 12 meses no futuro é muito menos confiável do que a previsão para o mês seguinte. Por esse motivo, previsões são acompanhados de um intervalo de confiança. Um intervalo de confiança de 95%, por exemplo, indica, com 95% de confiança, o intervalo de valores que velocidade medida poderá assumir.

9.2 Métodos

Existem diversos métodos para previsão de séries temporais. Ao longo do desenvolvimento deste trabalho os seguintes métodos foram considerados:

- Suavização exponencial: os valores passados contribuem para o valor atual com um peso que decai exponencialmente quanto mais distante estão do momento da previsão.
- Box-Jenkins SARIMA: um modelo robusto que faz uso de valores passados (autoregressão, AR) e erros passados (média móvel, MA), inclui integração (I) para tornar a série estacionária e estabilizar variância e conta com um processo iterativo para estimar parâmetros (Box-Jenkins).

- Rede neural de memória curta de longo prazo (LTSM): esse modelo consegue determinar com precisão o peso que deve ser dado a valores muito distantes do momento da previsão.
- Máquinas com vetor de suporte: os dados passam por uma transformação não-linear que permite a sua categorização em um espaço de dimensão elevada e posterior transformação reversa para reportar a previsão.

O método escolhido foi o Box-Jenkins SARIMA. Cada modelo assume uma série de hipóteses sobre o comportamento do que se deseja prever e possui um conjunto de parâmetros que precisam ser estimados para tal propósito. Para o modelo escolhido uma das hipóteses é de que os dados possuem autocorrelação, isto é, o valor atual depende dos n valores que o antecedem: se a velocidade foi alta um instante atrás é provável que ela seja alta agora, se no momento anterior o regime foi turbulento, é provável que no momento seguinte ele permaneça turbulento. Essa hipótese corresponde ao caráter autoregressivo (AR) do modelo. Outra hipótese do modelo é de que há autocorrelação não apenas em valores passados mas também nos erros passados. Essa hipótese corresponde ao caráter de média móvel (MA) do modelo. A interpretação dessa hipótese não é tão óvia mas seu uso se justifica pelo fato de que um processo AR com infinitos termos pode ser descrito por um processo MA com finitos termos e vice-versa. No contexto de séries temporais o princípio da parsimônia dita que dentre os modelos que caracterizem uma série, se escolha aquele com menos parâmetros. Dessa forma a combinação de termos AR e MA permite que a estrutura da série seja capturada com poucos parâmetros.

9.3 Análise visual

A escolha do modelo para prever uma série temporal se dá pelas características dessa série tais como padrões, valores atípicos, tendências, sazonalidade e ciclicidade. Dessa forma o primeiro passo é visualizar a série tanto em sua forma bruta quanto sob outras formas que ressaltem outras características que não sejam visíveis na sua forma bruta. A transformação para o espaço de frequência, por exemplo, permite identificar se há periodicidade na série. Para facilitar a visualização a imagem abaixo apresenta apenas os 2 últimos anos de dados em base horária:

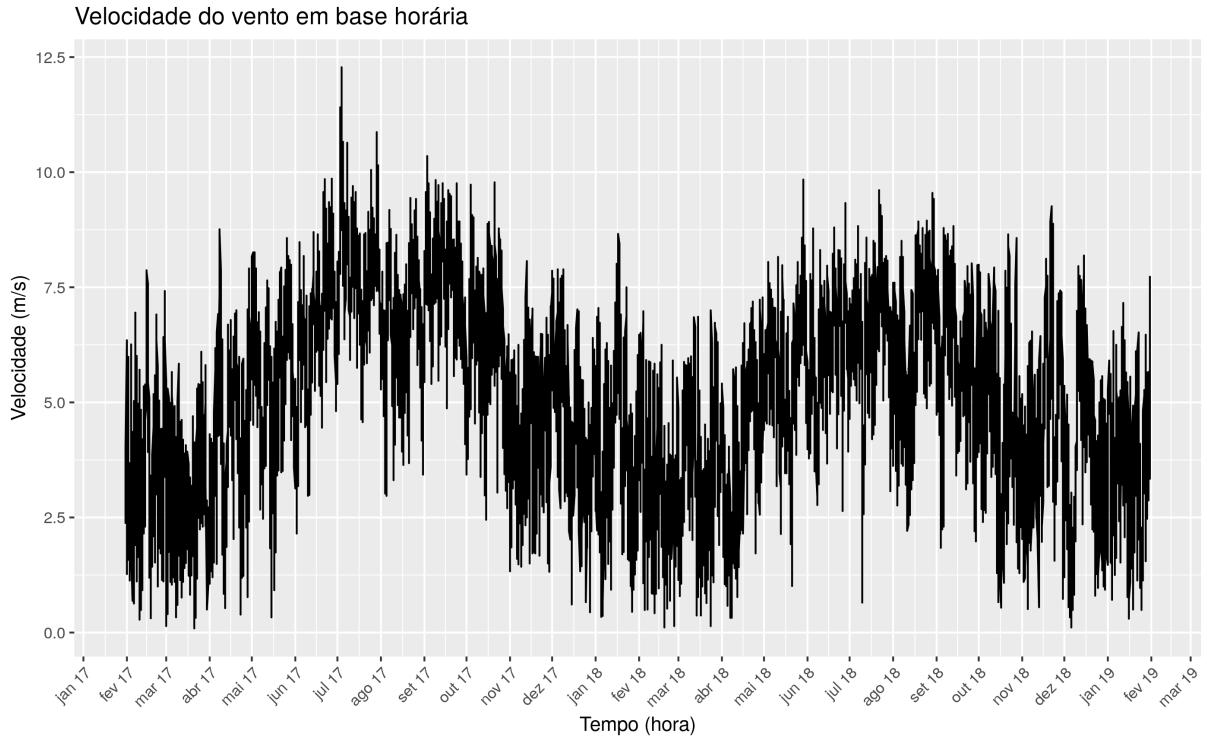


Figura 15 – Chapada

9.4 Medidas de qualidade de previsões

Para prever e quantificar a qualidade da previsão de uma série temporal é necessário dividi-la em um conjunto de treinamento e outro de teste. A divisão pode ser, por exemplo, 80% dos dados destinados ao conjunto de treinamento e os 20% subsequentes ao conjunto de teste. Empiricamente é desejável que o conjunto de teste seja igual ao maior ao tamanho do horizonte de previsão que se deseja obter. O conjunto de treinamento é utilizado para computar os parâmetros do modelo enquanto que o conjunto de teste é utilizado para quantificar o quanto bem o modelo é capaz de prever os dados subsequentes. Existem diversas maneiras de mensurar o quanto bem o modelo é capaz de prever os dados subsequentes. Deve-se evitar usar os resíduos como medida pois eles são resultado da estimativa dos parâmetros do modelo aos dados de treinamento. Como qualquer série pode ser aproximada por um modelo com infinitos parâmetros, os resíduos não dão informação sobre o poder preditivo do modelo para valores futuros, desconhecidos. A

O erro de previsão é dado pela diferença entre o valor medido, y_{t+h} , e o valor previsto pelo modelo, \hat{y}_{t+h} :

$$e_{t+h} = y_{t+h} - \hat{y}_{t+h}$$

Existem muitas maneiras de quantificar o erro de previsão. As seções seguintes

abordam as características de cada uma e identificam as que serão usadas para quantificar os resultados deste trabalho.

9.4.1 Erro médio absoluto (MAE)

O erro absoluto médio é dado pela média do módulo do erro de previsão. Essa medida penaliza o resultado tanto por erros negativos quanto positivos. A minimização dessa medida de erro resulta em estimativas que se aproximam da mediana dos dados.

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |e_t|$$

9.4.2 Raíz do quadrado da média do erro (RMSE)

A raíz do quadrado da média do erro é dada por:

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T e_t^2}$$

A minimização dessa medida de erro resulta em estimativas que se aproximam da média dos dados.

9.4.3 Erro absoluto médio escalonado (MASE)

Ao contrário das medidas anteriores, essa possui a vantagem de que é independente da escala dos dados medidos:

$$\text{MAPE} = \frac{\frac{1}{T} \sum_{t=1}^T |e_t|}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}$$

9.4.4 Validação cruzada

A validação cruzada consiste em usar uma janela de treinamento para calcular o erro de previsão para a próxima medida, expandir a janela em uma medida e calcular o novo erro e assim por diante sob um conjunto de teste. A medida de erro é dada pela média dos erros individuais. Qualquer uma das medidas acima pode ser utilizada no cálculo da validação cruzada. Essa é a maneira mais robusta de quantificar o erro de previsão pois atua efetivamente em dados que não foram utilizados para construir o modelo. A idéia da medida pode ser ilustrada da seguinte forma:

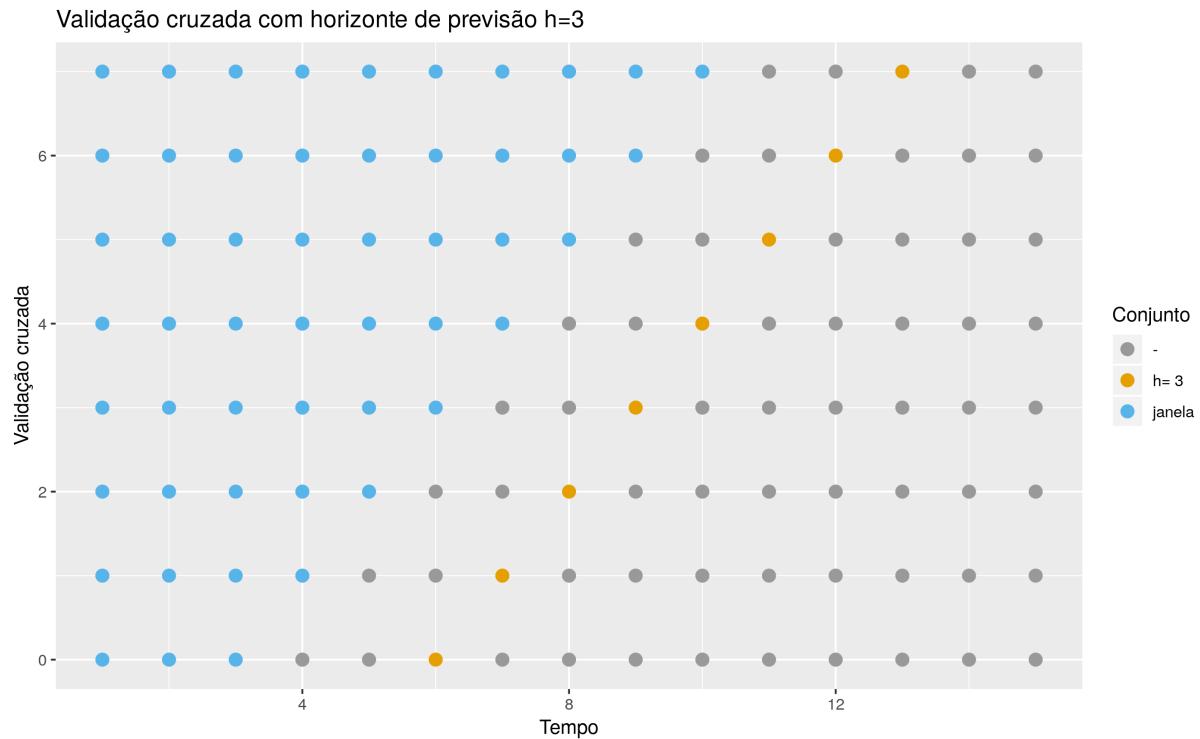


Figura 16 – Chapada

Onde as medidas de treinamento estão em azul e as de teste em amarelo. Na linha de baixo, mensura-se o erro de previsão que resulta ao tentar prever três horas a frente da hora atual. A linha acima inclui a medida seguinte ao conjunto de treinamento e tenta prever, novamente, três horas a frente. Isso é feito várias vezes. Os erros resultantes são calculados, por exemplo, por MASE e a sua média é reportada como valor final.

9.5 ARIMA

9.6 Box-Jenkins

9.7 Modelo Autoregressivo Variável

Ao longo do estudo observou-se que para diferentes períodos da série de dados, diferentes modelos ARIMA apresentavam melhor potencial preditivo não sendo possível elencar um único modelo que capturasse a estrutura da série de forma universal. Duas soluções para esse problema foram exploradas. Uma delas é optar por outro modelo. A outra, abordada nessa seção, é permitir que os parâmetros ARIMA fossem adaptativos, sendo recalculados, a medida que mais dados são coletados. Essa é uma estratégia interessante por vários motivos. É possível estimar os parâmetros de um modelo ARIMA rapidamente

de forma que a previsão esteja disponível ininterruptamente mesmo em escalas temporais da ordem de minutos. Qualquer mudança estrutural na grandeza física, tal como devido ao aumento da temperatura média do planeta, é automaticamente capturada pelo modelo.

A imagem abaixo exibe a evolução temporal dos graus de liberdade do modelo em um espaço de fase sendo possível entender como os parâmetros variam ao longo do tempo e consequentemente como a estrutura dos dados (limitada pela representação por um modelo ARIMA) varia. O tempo é representado pela escala em cor. O parâmetro de diferenciação, d , do modelo assume durante quase todo o período o valor 1 e portanto exibe-se apenas os parâmetros p e q .

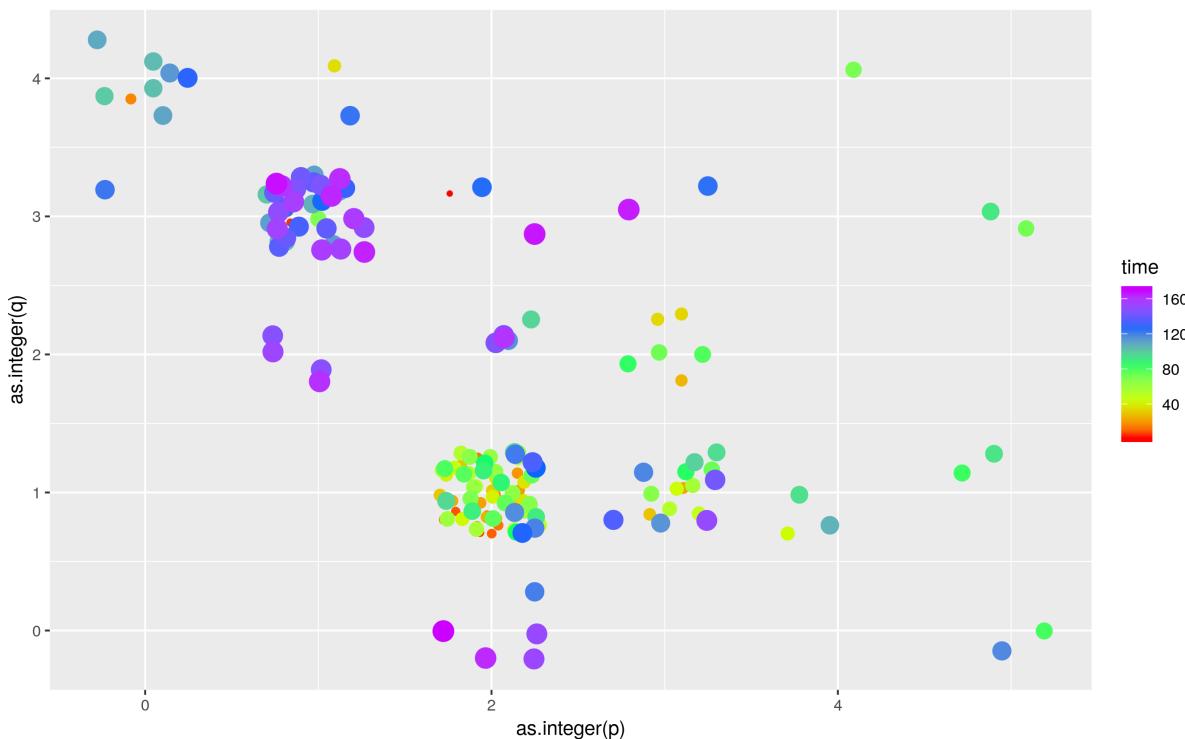


Figura 17 – Evolução dos parâmetros p e q de um modelo ARIMA ao longo do tempo.

Modelos ARIMA são definidos pelos parâmetros p,d,q . O parâmetro p representa o número de termos anteriores da série que influenciam o valor atual. Os pesos com que cada termo anterior contribui para a estimativa do valor seguinte são parâmetros a serem estimados por regressão linear. O parâmetro q tem um papel análogo ao do parâmetro q . Ele refere-se à inclusão dos q erros anteriores na estimativa do valor seguinte.

Os parâmetros assumem apenas valores inteiros. Para facilitar a visualização da evolução temporal adicionou-se um ruído aleatório a cada ponto de modo a evitar a sua sobreposição. Fica claro que no início do período de 160 horas o principal modelo ARIMA é o ARIMA(2,1,1) enquanto que ao final do período o modelo preferencial é o ARIMA(1,1,3). Também observa-se que diversos outros modelos são assumidos com menor frequência.

O resultado é superior ao uso de um único modelo ARIMA para toda a série:

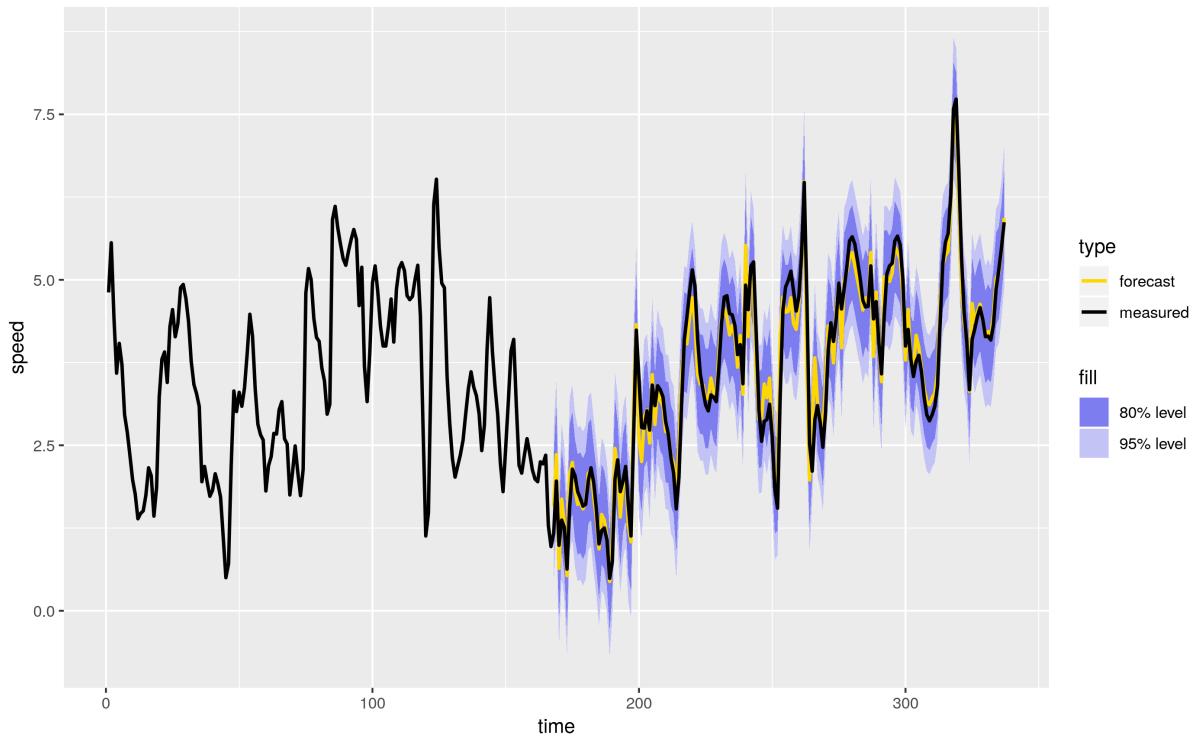


Figura 18 – Evolução dos parâmetros p e q de um modelo ARIMA ao longo do tempo.

9.8 GARCH

Parte II

Análise de dados de vento

10 Características de longo prazo

É interessante exemplificar as discussões subsequentes com dados reais de medição. O Centro Europeu de Previsões Metereológicas de Médio Prazo (ECMWF, sigla em inglês) disponibiliza publicamente dados metereológicos de todo o globo medidos por satélite, incluindo dados de velocidade e direção do vento compilados no banco de dados ERA5 ([ECMWF, 2017](#)). Os dados de velocidade do vento são referentes a uma altura de 80 metros e possuem resolução de 30 quilômetros. As unidades de todas as grandezas referenciadas neste trabalho seguem o padrão SI. Em particular velocidade é dada em metros por segundo e o tempo em segundos. A partir de uma posição do globo dada em latitude e longitude como entrada obtém-se da ECMWF uma série temporal com 18 variáveis: 9 de velocidade do vento e 9 de direção referentes aos nove quadrantes que circundam a posição informada: quadrante central, norte, sul, sudeste, etc.

A figura abaixo mostra esses quadrantes para uma região no norte baiano, na cidade de Sento Sé para a posição 9.8436 S, 41.1912 W.

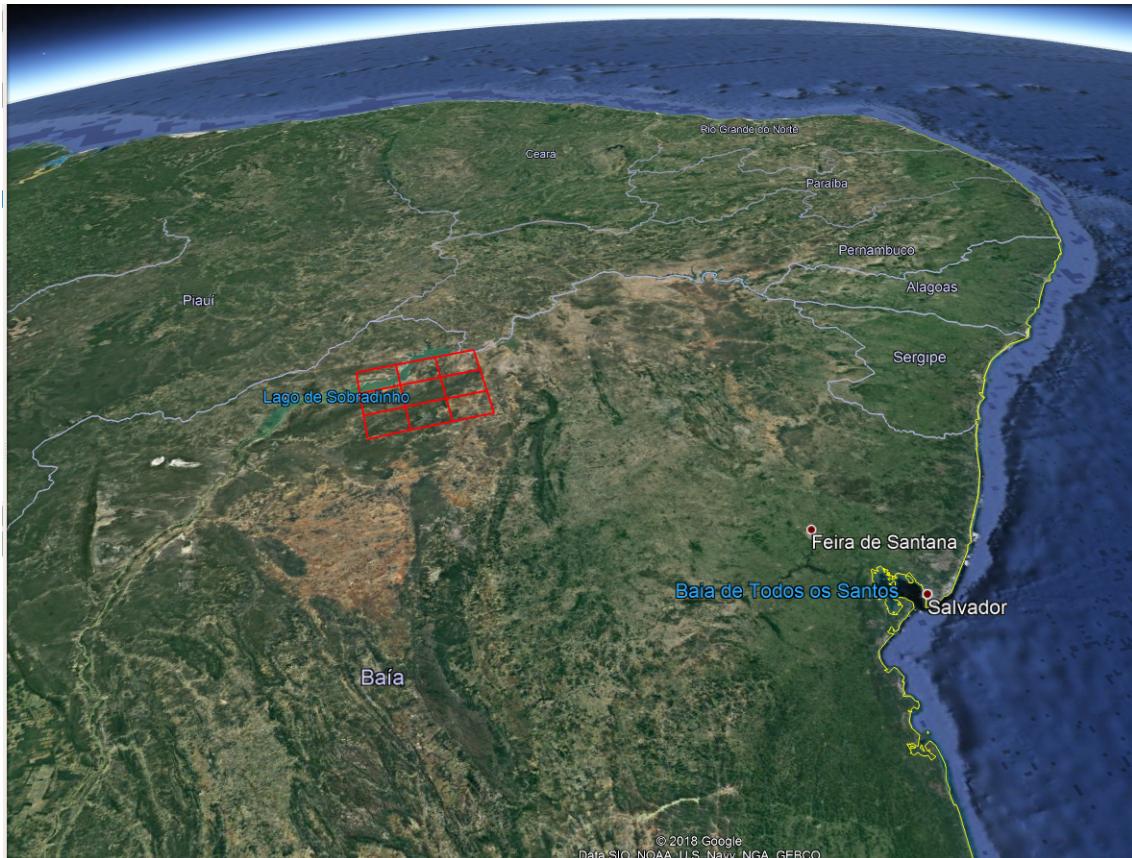


Figura 19 – Localização da região de medição por satélite dos dados de vento discutidos no texto. Fonte: SIO, NOAA, U.S. Navy, NGA, GEBCO. © 2018 Google

A partir de análises de diversas localidades do Nordeste brasileiro sabe-se que o recurso eólico dessa região é representativo de todo o Nordeste do país, isto é, a velocidade média do vento é alta, predominantemente entre 6 e 8 m/s, fortemente unidirecional a jusante de sudeste a 120° em relação ao norte geográfico, com uma componente significativamente menor, mas relevante, a cerca de 150° como pode ser observado na rosa dos ventos gerada a partir dos dados medidos para o ano de 2017:

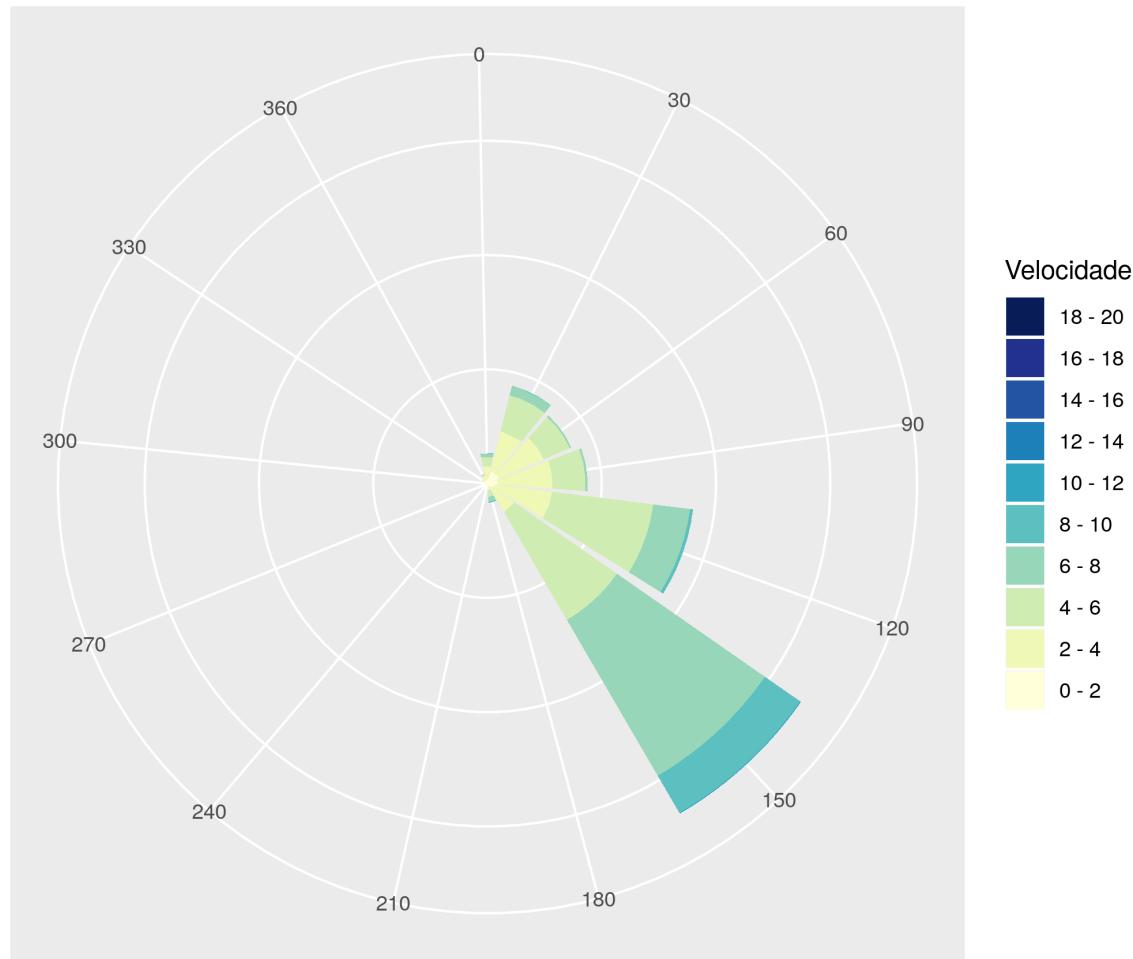


Figura 20 – Rosa dos ventos do recurso eólico da cidade de Sento Sé no norte da Bahia - característico do Nordeste brasileiro, a partir de dados medidos para o ano de 2017. Fonte: autoria própria, dados ([ECMWF, 2017](#)).

Analizando as rosas dos ventos mensais se confirma o comportamento anual e também é possível observar que nos meses de junho a outubro a velocidade do vento é muito alta enquanto que nos demais meses, de Novembro a Maio, ela é muito menor. Com essa informação, é possível, por exemplo, nos meses de baixa velocidade e consequente baixa geração de energia, complementar o suprimento energético do local com outras fontes de energia renovável tal como a energia solar que apresenta o comportamento inverso à energia eólica na região, isto é, nos meses de primavera e verão a elevada altitude solar permite extrair a maior proporção de energia anual ao contrário dos meses de outono e inverno.



Figura 21 – Rosas dos ventos mensais do recurso eólico da cidade de Sento Sé no norte da Bahia a partir de dados medidos para o ano de 2017. Fonte: autoria própria, dados ([ECMWF, 2017](#)).

11 Características de curto prazo

Sabe-se que em períodos curtos, o vento apresenta um comportamento que aparenta ser puramente estocástico, associado ao comportamento turbulento característico da dinâmica de fluídos pouco viscosos tal como o ar. No entanto, observando períodos maiores de tempo, tais como dias, meses e anos observa-se que existem tendências coexistentes com o caráter aleatório, errático do vento. É tarefa da ciência de dados evidenciar a presença desses padrões a partir dos dados medidos, pois, apenas exibindo graficamente a evolução temporal da velocidade do vento leva a crer que não exista qualquer tendência ou padrão nestes dados:

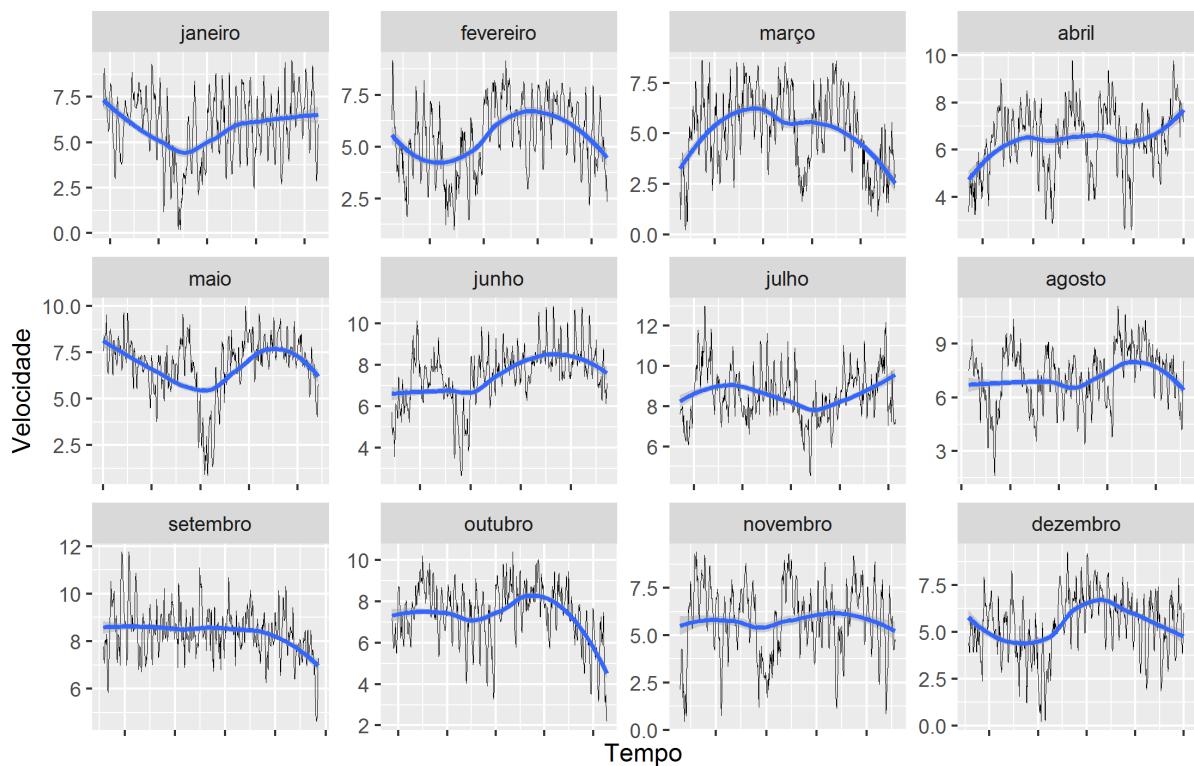


Figura 22 – Velocidade do vento para cada mês do ano de 2017 da cidade de Sento Sé no norte da Bahia. Curva em preto: dados medidos. Curva em azul: curva ajustada aos dados medidos. Fonte: autoria própria, dados ([ECMWF, 2017](#)).

A figura abaixo apresenta a variação horária da velocidade do vento para cada mês do ano para a cidade de Sento Sé. Observa-se que o vento no local apresenta um ciclo diário característico e que está presente em todos os meses do ano. De acordo com ([K SILVA C; S., 2013](#)) o ciclo diário do vento é um fator mais relevante do que a velocidade média do vento para definir a localização de um parque eólico, devido a relação entre a geração horária de energia e a demanda correspondente. Embora a velocidade do vento

seja bem mais alta em Julho do que em Janeiro, em ambos os meses tem-se um mínimo de velocidade durante o meio da tarde, uma taxa de variação positiva a medida que a noite se aproxima e uma estabilização durante a madrugada e primeiras horas da manhã. O mesmo vale para os demais meses. Pode-se explicar esse fenômeno observado que durante o dia o aquecimento da atmosfera pelo Sol torna o vento turbulento sendo mais difícil o seu aproveitamento para geração de energia, pois sua velocidade direcional é baixa. À noite, por outro lado, a atmosfera se estratifica, o vento fica pouco turbulento e a sua velocidade direcional aumenta. Essa explicação embora válida para este local pode não ser para outro, pois existem diversos fatores, além do aquecimento local da atmosfera, que definem o comportamento do vento. Em regiões desérticas, por exemplo, observa-se uma alta estabilidade na atmosfera devido à baixa umidade. A mesma explicação pode ser estendida para descrever o comportamento mensal: durante meses de inverno e outono a velocidade do vento é alta, enquanto que em meses de verão e primavera ela é baixa.

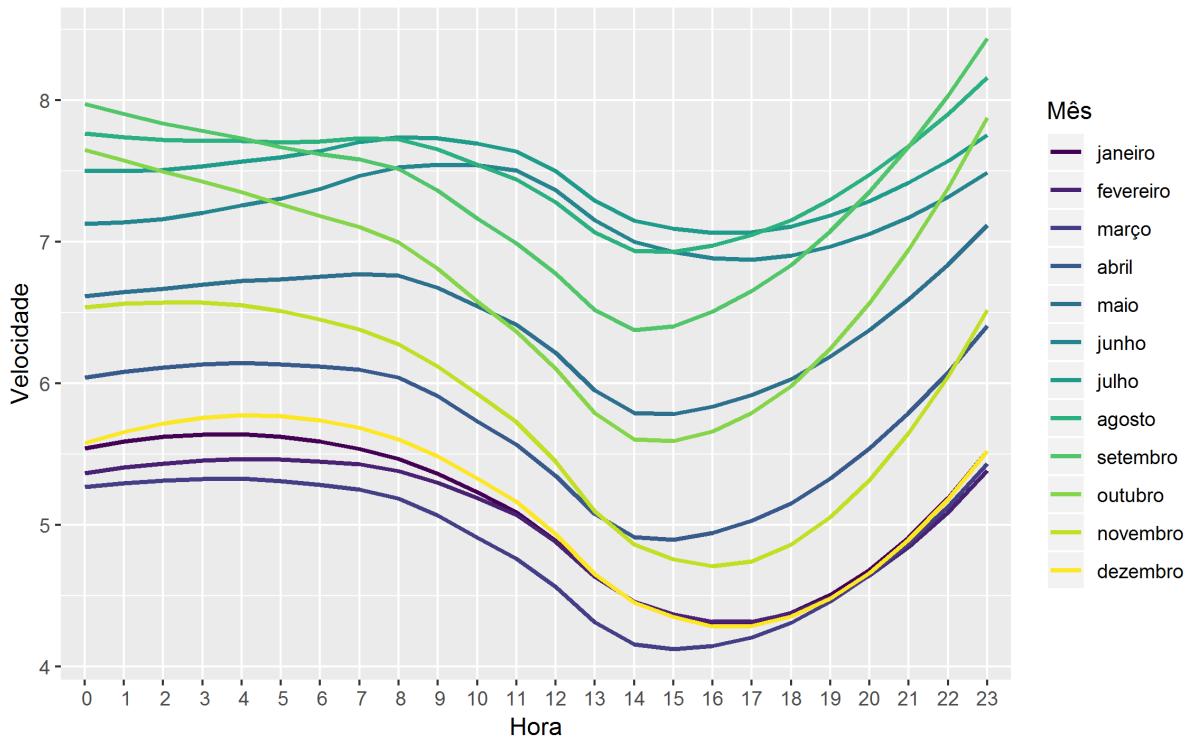


Figura 23 – Velocidade do vento para cada hora do dia e para cada mês do ano de 2017 para a cidade de Sento Sé no norte da Bahia. Fonte: autoria própria, dados ([ECMWF, 2017](#)).

Parte III

Embasamento teórico

12 Embasamento matemático

A separação entre o que é fundamentalmente aleatório e o que exibe uma tendência em alguma janela temporal é de extrema relevância para a concepção de um modelo matemático que descreva tal fenômeno. Essas contribuições podem ter origens físicas diferentes e a formulação das equações diferenciais estocásticas envolve a distinção entre um termo associado à deriva e outro à difusão o qual não contribui para o valor esperado da variável em questão, isto é, na ausência de deriva, o valor esperado da variável estocástica seria zero, em outras palavras, ele distribui a variância igualmente para mais e para menos.

Um processo estocástico contínuo no tempo, $W(t)$, definido para $t \geq 0$ com $W(0) = 0$ tal que o incremento $W(t) - W(s)$ é dado por uma distribuição gaussiana com média 0 e variância $t - s$ para qualquer $0 \leq s < t$ sendo tais incrementos, restritos aos que não se sobrepõem no tempo, independentes, recebe o nome de processo de Wiener ([WIENER, 1923](#)). O movimento browniano é um exemplo particular de processo de Wiener.

O processo S_t descrito pela equação estocástica abaixo, por exemplo ([R, 2018](#)), possui o termo de deriva $\mu S_t dt$ e o termo de difusão dado por $\sigma S_t dW_t$ o qual está associado a um processo de Wiener representado por W_t :

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t \quad (12.1)$$

Um processo estocástico cujo valor esperado futuro é igual ao valor presente é classificado como um *martingale* ([E., 2004](#)). Essa é uma característica importante da análise, pois diversos teoremas a tem como hipótese. É claro que a teoria encontraria pouca aplicação prática se não contemplasse processos que envolvem deriva. Para tanto, o teorema de Girsanov ([E., 2004](#)) garante que tal processo, sujeito a algumas hipóteses pouco restritivas, pode tornar-se um *martingale* por meio de uma mudança de escala - uma função multiplicativa apropriada a qual é conhecida como derivada de Radon-Nikodym ([E., 2004](#)).

A evolução histórica do preço das ações da Google na NASDAQ bolsa, por exemplo, apesar de ter caráter estocástico, apresenta claramente uma tendência de subida (uma deriva, ou *drift* do inglês):



Figura 24 – Valor de abertura das ações da Google na NASDAQ. Curva em preto: dados medidos. Curva em azul: curva ajustada aos dados medidos. Fonte: autoria própria, dados ([FINANCE, 2018](#)).

13 Embasamento físico

13.1 Restrições físicas

Como fora exposto, a velocidade do vento apresenta várias tendências de longo prazo, comportamentos periódicos - sazonais e até horários e padrões característicos de orografia, clima e latitude. No entanto, essa é uma grandeza fundamentalmente estocástica, isto é, está associada a uma variável aleatória e portanto, a sua previsão está fadada a uma incerteza intrínseca. Antes da formulação da mecânica quântica acreditava-se que se conhecendo perfeitamente o Hamiltoniano de um sistema, isto é, sabendo a posição e momento de todas as partículas que o compõe, seria possível prever com exatidão a sua evolução temporal, ao menos em tese (dado que tal feito é inviável para qualquer sistema moderadamente complexo). No entanto, o princípio da incerteza de Heisenberg pôs fim a essa crença, visto que enuncia que não somente é impossível saber simultaneamente a posição e momento exatos de uma partícula como não faz sentido falar da simultaneidade das duas grandezas; tampouco é útil a informação sobre velocidade e momento de todas as partículas de um sistema, pois as grandezas de interesse são de natureza estatística, como a temperatura por exemplo.

A mecânica estatística desenvolveu-se em paralelo ignorando a busca pela determinação exata da evolução temporal de um sistema. Ao invés disso, concentrou-se em determinar valores médios de quantidades de interesse emergentes do comportamento caótico fundamentada na hipótese de que todas as possíveis configurações do sistema ocorrem com igual probabilidade. Apesar do grande sucesso em diversas áreas da física foi incapaz de explicar o comportamento turbulento de fluidos, foi incapaz de explicar como surge ordem a partir do caos. Relativamente pouco progresso foi feito até então (PEITGEN, 2004). É por esse motivo que frequentemente a previsão do tempo falha, até mesmo em uma escala de horas independente de quantas estações metereológicas sejam utilizadas, já que de acordo com a teoria do caos existe um horizonte de previsão (STEWART, 1989) a partir do qual não é possível obter estimativas confiáveis. A teoria do caos mostrou que a ignorância quanto a natureza da complexidade era muito maior. Tome como exemplo um sistema simples com uma variável discreta x_t e uma regra de evolução temporal não-linear, $x_{t+1} = x_t(1 + x_t)$ (NOSRATI; SHAFIEE, 2018). Conhece-se tudo sobre esse sistema. Ele é perfeitamente determinístico. Ainda assim existe um certo instante de tempo a partir do qual não é possível predizer o seu estado futuro. Esse instante chama-se de transição ao caos (LORENZ, 1996).

13.2 Abordagem de longo prazo

Fenômenos fortemente não-lineares como a turbulência, não são, no entanto, completamente imprevisíveis. Acredita-se que dependendo do sistema e sujeito às restrições expostas acima a tarefa seja factível, porém não existe uma única maneira de abordar tal problema. Este trabalho propõe a investigação de qual modelo produz melhores resultados de estimativa de produção de energia, começando por um modelo de evolução temporal de velocidade do vento que faça uso dos dados de medição para encontrar a distribuição de probabilidade associada à essa grandeza.

É prática comum na indústria a determinação da distribuição de probabilidade de energia a partir de simulações de Monte Carlo ([KALOS; WHITLOCK, 1986](#)) tendo como hipótese que a velocidade do vento é descrita por uma distribuição de Weibull a qual já foi verificada experimentalmente a partir de muitas análises de dados de vento e tem sólido embasamento teórico conforme o teorema de Fisher–Tippett–Gnedenko, o qual prova que o máximo de um conjunto de variáveis aleatórias independentes, após renormalização, converge apenas para uma de três classes de distribuições: a distribuição de Gumbel, a distribuição de Fréchet ou a distribuição de Weibull ([BASRAK, 2011](#)).

Para a cidade de Sento Sé, por exemplo, a distribuição de velocidades do quadrante central é exibida abaixo:

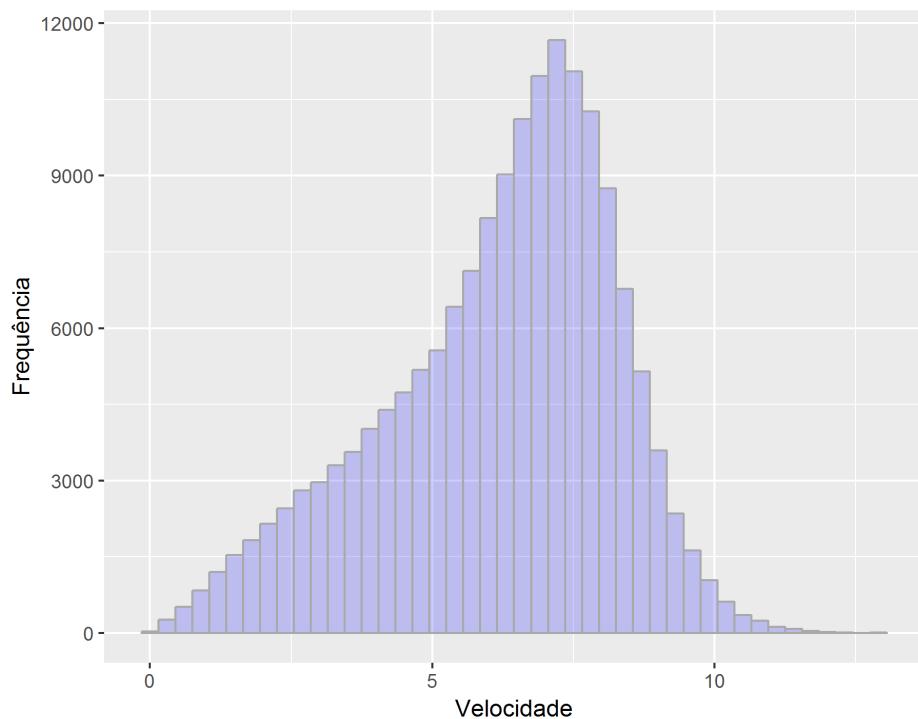


Figura 25 – A distribuição de velocidades pode ser aproximada por uma distribuição de Weibull. Dados do quadrante central dos dados ERA5 ([ECMWF, 2017](#)) para a cidade de Sento Sé no norte da Bahia. Fonte: autoria própria, dados ([ECMWF, 2017](#)).

Essa distribuição claramente não é gaussiana porque possui uma cauda unilateral. Tal distribuição é representada satisfatoriamente por uma distribuição de Weibull por meio de dois parâmetros: o parâmetro de forma k e o parâmetro de escala λ . Seus primeiros momentos são definidos por (MENČÍK, 2016):

$$\langle v \rangle = \lambda \Gamma \left(1 + \frac{1}{k} \right) \quad (13.1)$$

$$\sigma^2 = \lambda^2 \left\{ \Gamma \left(1 + \frac{2}{k} \right) - \left[\Gamma \left(1 + \frac{1}{k} \right) \right]^2 \right\} \quad (13.2)$$

Avaliando a distribuição de densidade de probabilidade de velocidade para todos os quadrantes dos dados de vento da cidade de Sento Sé, confirma-se que todos seguem uma distribuição do tipo weibull. Sabendo que a direção preferencial do vento é sul-sudeste observa-se, a partir das curvas para norte, nordeste e noroeste, que à medida que flui no local a velocidade média aumenta.

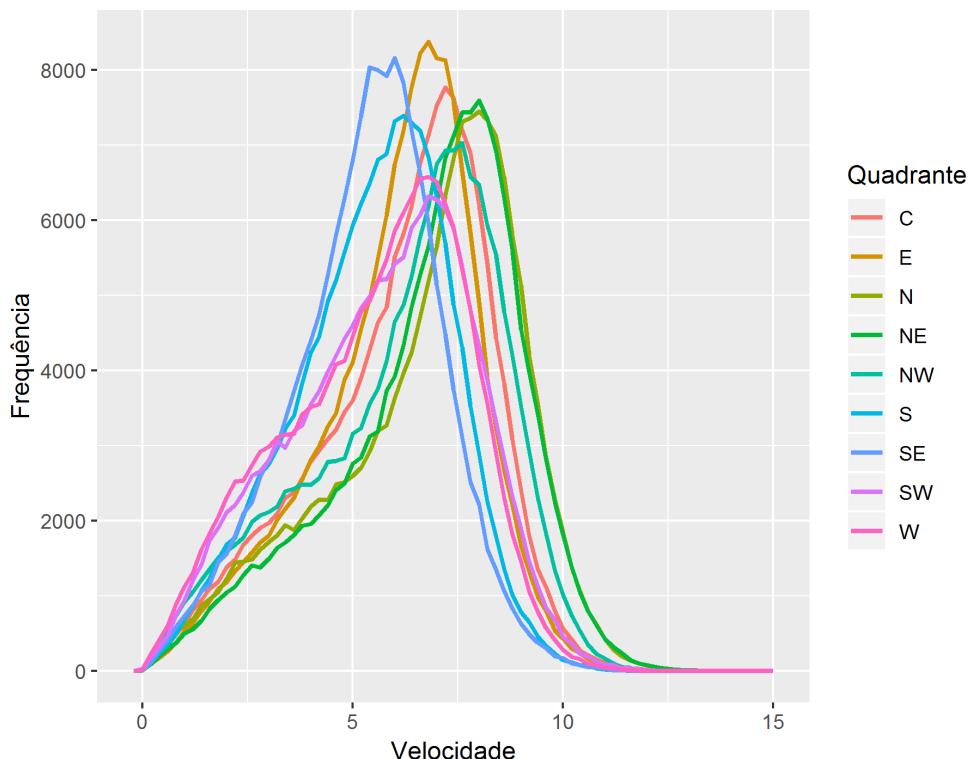


Figura 26 – Distribuição de velocidades de todos os quadrantes dos dados ERA5 (ECMWF, 2017) para a cidade de Sento Sé no norte da Bahia. Fonte: autoria própria, dados (ECMWF, 2017).

Os parâmetros λ e k são utilizados para caracterizar o recurso eólico em uma dada posição (latitude e longitude) e altura. Para obter o recurso eólico em outra posição ou altura (posição e altura de uma turbina por exemplo) é feito um escalonamento da distribuição:

muda-se a sua forma para obter um valor esperado de velocidade correspondente à nova posição.

13.3 Desvantagens

Essa abordagem embora adequada para estimativas de longo prazo e muito utilizada na indústria possui várias desvantagens:

- A hipótese da distribuição de Weibull é completamente inválida a curto prazo, visto que é uma distribuição constante e não pode, portanto, representar as flutuações incessantes características do comportamento turbulento do vento ([LOUKATOUA SYDNEY HOWELLB; DUCKC, 2011](#)). Uma distribuição de probabilidade que descreva adequadamente a velocidade do vento deve ser dependente do tempo. A evolução temporal horária ou até mesmo minuto a minuto é uma informação essencial para operadores de subestações de distribuição de energia elétrica, pois permite adequar oferta de energia a picos e baixas de demanda os quais são conhecidos
- A simulação é lenta. O Método de Monte Carlo, como aplicado pela indústria, é um método de força bruta, ou seja, lança mão de poder computacional para atingir seu objetivo. E deve ser realizado toda a vez que se deseje obter o mesmo resultado ao passo que um modelo físico permite que o esforço computacional seja empregado apenas uma vez e, além disso, exige menos poder computacional. É impensável refazer tais cálculos a cada hora.
- O método de Monte Carlo, como aplicado nessa problema, é apenas um recurso computacional, não descreve o comportamento físico do sistema e portanto não pode ser empregado em análises sofisticadas tais como mensurar a dependência temporal e espacial entre dois parques eólicos, ou seja, mensurar o efeito de esteira ([CEPEL, 2001](#)) em resolução horária, podendo, por exemplo, como resultado conceber uma estratégia de gerenciamente setorial conjunto entre os dois parques, ou seja, orientar as turbinas dos dois parques de maneira mutuamente benéfica. Outro exemplo, seria prever o melhor intervalo de tempo para realizar a manutenção de uma turbina levando em conta o tempo necessário para realizar a manutenção as quais são, via de regra, custosas e demoradas.

13.4 Ressalvas

Existem muitas fontes de incerteza ao longo de todo o processo de estimativa da produção de energia. Incerteza na obtenção dos dados de medição, na conversão entre velocidade e energia, na extração horizontal (do local de medição à posição das

turbinas), da extração vertical (da altura da medição à altura das turbinas) entre muitas outras. A consideração dessas fontes de incerteza fogem ao escopo do presente trabalho mas devem ser todas consideradas em qualquer trabalho comercial. Pode-se dizer que este trabalho estuda o caso particular em que a medição do vento é feita a alguns diâmetros de rotor de distância da turbina, exatamente na altura do cubo rotor.

Parte IV

Revisão bibliográfica

A área de estudo deste trabalho possui vasta literatura. Modelos autoregressivos de média móvel (ARMA) foram utilizados em (E, 2011) para prever tanto a velocidade quanto a direção do vento.

Uma revisão bibliográfica dos vários métodos de previsão de dados de vento é feita em (WY., 2014). Este trabalho conclui que esses métodos possuem bom desempenho em diferentes resoluções temporais e que o erro de previsão aumenta com o tempo desejado de previsão.

Modelos para as flutuações da velocidade do vento foram explorados em (R., 2012) com base na solução analítica da equação de Fokker–Planck–Kolmogorov e várias classes de distribuições de probabilidade são propostas como solução. Neste trabalho, também foi feita uma análise teórica do efeito que tufões de vento causam na geração de energia por turbinas.

Em (R.; FG, 2012) são feitas simulações estocásticas de caminho aleatório multi-fractal, isto é, um espectro contínuo de expoentes se faz necessário para caracterizar a geometria do objeto de estudo. Usando uma curva de potência experimental os autores produzem uma série temporal de produção de energia.

As flutuações na velocidade do vento são modelados em (PJ; RB., 2001) por meio de um processo de Ornstein-Uhlenbeck-Rayleigh bidimensional. Este trabalho observa que o modelo markoviano nele proposto não reproduz a variabilidade de curto-prazo exibida pelo vento. De acordo com (EB MORALES JM; H., 2016) o estudo de equações diferenciais estocásticas para a previsão da velocidade do vento é vasto na literatura, mas o mesmo não é verdade para a estimativa da produção de energia. Apesar da observação, o trabalho não foge à regra e também aborda apenas modelos estocásticos para a velocidade do vento. Somente em (JK ZUGNO M, 2016) que modelos estocásticos voltados à produção de energia são explorados, com o objetivo de obter uma previsão de até 2 dias.

Parte V

Planejamento

O principal objetivo deste trabalho é conceber uma maneira mais rápida, eficiente e que produza melhores resultados para estimar a produção de energia a partir de dados de vento. Existem muitos métodos disponíveis para tentar cumprir este objetivo, como exposto na revisão bibliográfica, entre eles:

- Modelos de autoregressão (ARMA);
- Redes neurais;
- Equações diferenciais estocásticas para velocidade do vento, produção de energia ou ambos;
- Teoria combinatória com peso variável;
- Caminho aleatório multifractal;
- Movimento Browniano geométrico;
- Combinações ou variações entre esses modelos

É parte deste trabalho avaliar qual o método produz a melhor estimativa para a produção de energia. A validação do modelo será feita pela comparação com dados medidos de parques eólicos em operação e os resultados obtidos com o modelo.

A execução desse trabalho consiste em várias etapas. As principais delas são elencadas abaixo:

- Estudar a teoria matemática do cálculo estocástico
- Investigar se ideias da teoria do caos podem ser utilizadas na formulação do modelo
- Investigar o que já foi feito na literatura na área de previsão de geração de energia eólica
- Avaliar os diferentes métodos disponíveis para modelar dados estocásticos
- Formular o modelo com base no método mais promissor
- Obter permissão para utilizar os dados medidos
- Estudar sobre solução numérica de equações diferenciais estocásticas
- Resolver o problema pelo método tradicional
- Comparar os resultados entre métodos
- Comparar os resultados com dados medidos de energia

- Redigir o texto final
- Apresentar à banca

No gráfico de Gantt abaixo identificadas as macro-etapas do trabalho e são estabelecidos durações e prazos para a conclusão dessas etapas

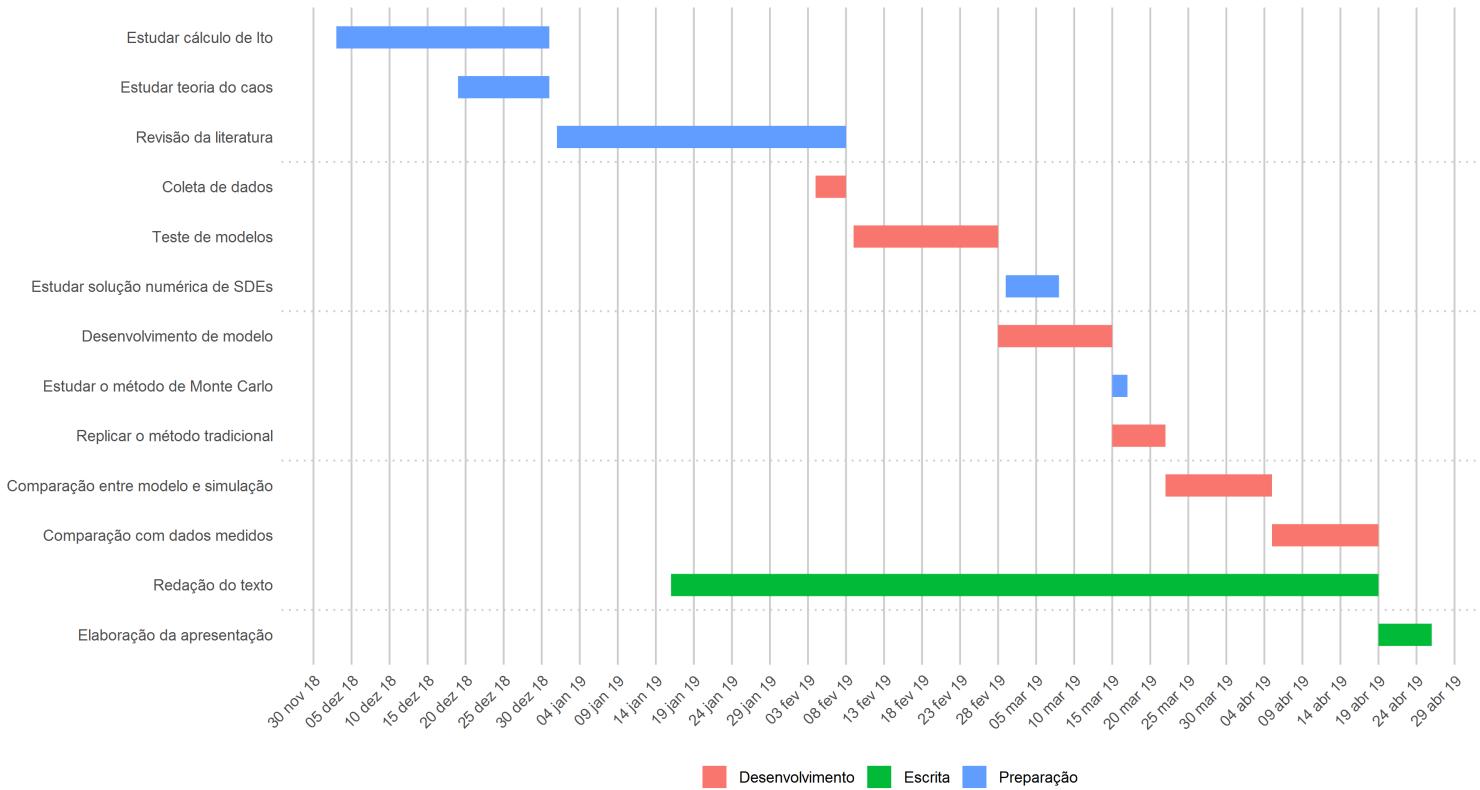


Figura 27 – Gráfico de Gantt do planejamento das etapas de desenvolvimento do presente trabalho. Fonte: autoria própria.

Referências

- ACKERMANN, T. *Wind Power in Power Systems*. [S.l.]: John Wiley & Sons, Ltd, 2012. Citado na página 12.
- AN., C. Energy output estimation for small-scale wind power generators using weibull-representative wind data. *Journal of Wind Engineering and Industrial Aerodynamics*, 2003. Citado na página 4.
- BASRAK, B. Fisher-tippett theorem. In: _____. *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 525–526. ISBN 978-3-642-04898-2. Disponível em: <https://doi.org/10.1007/978-3-642-04898-2_254>. Citado na página 48.
- CENTRO DE PESQUISAS DE ENERGIA ELÉTRICA, CEPEL. *Atlas do Potencial Eólico Brasileiro*. Brasília, 2001. Disponível em: <http://www.cresesb.cepel.br/publicacoes/download/atlas_eolico/atlas%20do%20potencial%20eolico%20brasileiro.pdf>. Citado 2 vezes nas páginas 12 e 50.
- E, S. J. E. Arma based approaches for forecasting the tuple of wind speed and direction. *Journal of Applied Energy*, 2011. Citado na página 53.
- E., S. S. *Stochastic Calculus for Finance II: Continuous-Time Models*. [S.l.]: Springer Finance - Pittsburg, 2004. Citado na página 45.
- EB MORALES JM, M. J. I.; H., M. Short-term probabilistic forecasting of wind speed using stochastic differential equations. *International Journal of Forecasting*, 2016. Citado na página 53.
- ECMWF. *ERA5 Reanalysis Database*. 2017. Disponível em: <<https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era5>>. Acesso em: 03 nov 2018. Citado 10 vezes nas páginas 5, 6, 14, 39, 40, 41, 42, 43, 48 e 49.
- FINANCE, G. *Google Finanças*. 2018. Disponível em: <<https://www.google.com/search?q=NASDAQ:GOOG&tbo=fin>>. Acesso em: 11 nov 2018. Citado 2 vezes nas páginas 6 e 46.
- JK ZUGNO M, M. H. M. Probabilistic forecasts of wind power generation by stochastic differential equation models. *International Journal of Forecasting*, 2016. Citado na página 53.
- K SILVA C, F. P. S.; S., C. Wind power design in isolated energy systems: impacts of daily wind pattern. *Applied Energy*, n. 1, 2013. Citado na página 42.
- KALOS, M.; WHITLOCK, P. *Monte Carlo Methods*. New York, NY, USA: John Wiley and Sons, A Wiley-Interscience Publication, 1986. Citado na página 48.
- LORENZ, E. *Essence of Chaos*. [S.l.]: University of Washington Press - Nova Iorque, 1996. Citado na página 47.

- LOUKATOUA SYDNEY HOWELLB, P. J. A.; DUCKC, P. Stochastic wind speed modelling for estimation of expected wind power output. *Journal of Applied Energy*, 2011. Citado na página 50.
- M HOOSHMAND R-A, K. A. P.; A-H, Z. Joint operation of wind farm, photovoltaic, pump-storage and energy storage devices in energy and reserve markets. *International Journal of Electrical Power & Energy Systems*, 2015. Citado na página 4.
- MENČÍK, J. Weibull distribution. *Springer, Berlin, Heidelberg*, v. 1, n. 1, p. 71–97, 2016. Citado 2 vezes nas páginas 4 e 49.
- NOSRATI, K.; SHAFIEE, M. Fractional-order singular logistic map: Stability, bifurcation and chaos analysis. *Chaos, Solitons & Fractals*, v. 115, p. 224 – 238, 2018. ISSN 0960-0779. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0960077918308932>>. Citado na página 47.
- P, V. S. R. A review of wind speed probability distributions used in wind energy analysis: case studies in the canary islands. *Renewable Sustainable Energy Review*, 2003. Citado na página 4.
- PEITGEN, J. e. S. *Chaos and Fractals: New Frontiers of Science 2nd*. [S.l.]: Springer Finance - Bremen, 2004. Citado na página 47.
- PJ, E.; RB., H. Level-crossing statistics of the horizontal wind speed in the planetary surface boundary layer. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2001. Citado na página 53.
- R., C. Pdf models and synthetic model for the wind speed fluctuations based on the resolution of langevin equation. *Journal of Applied Energy*, 2012. Citado na página 53.
- R., C.; FG, S. Modeling of atmospheric wind speed sequence using a lognormal continuous stochastic equation. *Journal of Wind Engineering and Industrial Aerodynamics*, 2012. Citado na página 53.
- R, M. F. Z.-M. Construction of sde-based wind speed models with exponentially decaying autocorrelation. *Journal of Applied Energy*, 2018. Citado na página 45.
- STEWART, I. *Does God Play Dice: The New Mathematics of Chaos*. [S.l.]: Blackwell Publishing - Londres, 1989. Citado na página 47.
- WIENER, N. *Differential Space*. [S.l.]: Wiley, 1923. Citado na página 45.
- WY., C. A literature review of wind forecasting methods. *Journal of Power Energy Engineering*, 2014. Citado na página 53.