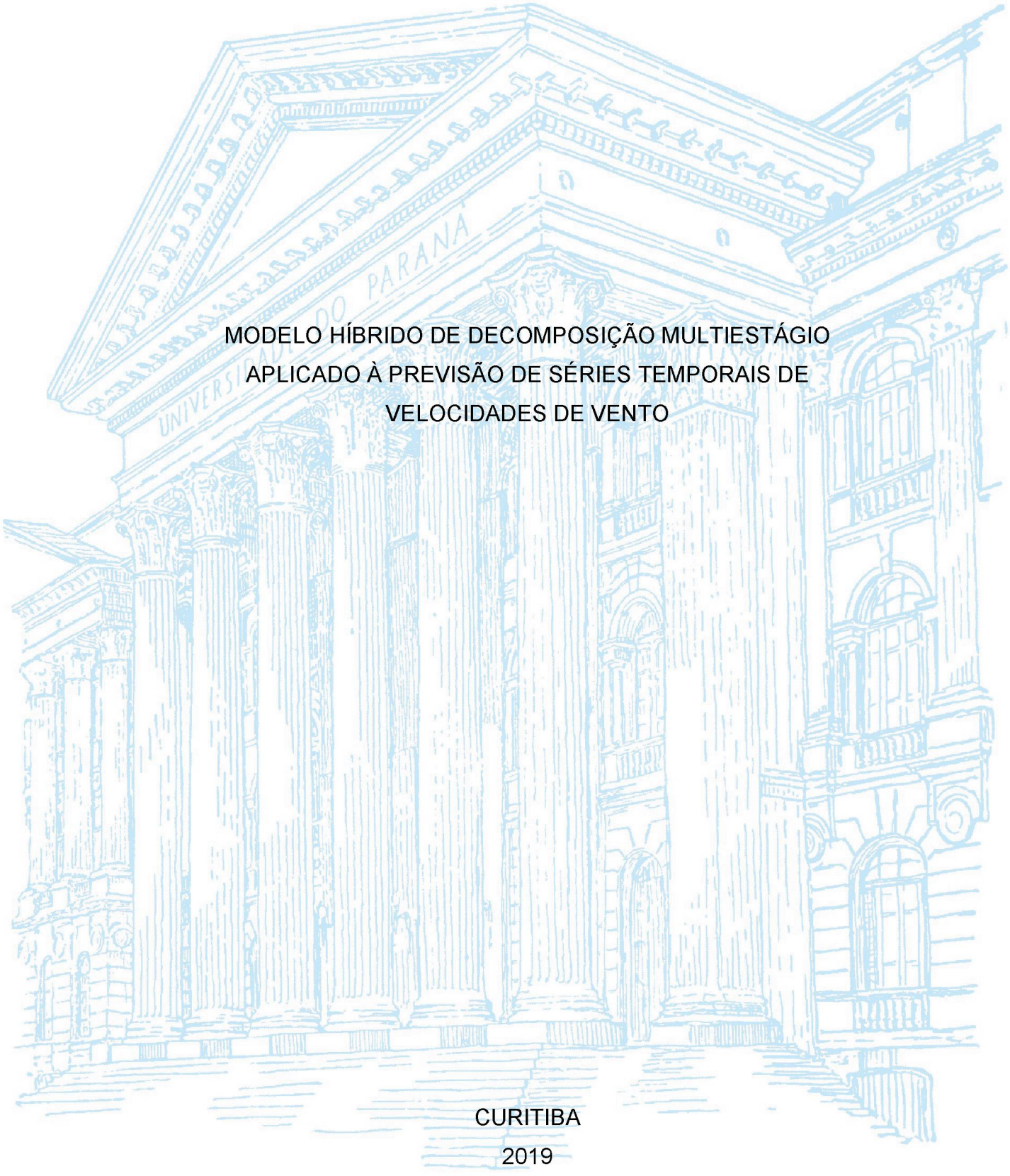
UNIVERSIDADE FEDERAL DO PARANÁ

SINVALDO RODRIGUES MORENO



MODELO HÍBRIDO DE DECOMPOSIÇÃO MULTIESTÁGIO APLICADO À PREVISÃO DE SÉRIES TEMPORAIS DE VELOCIDADES DE VENTO

CURITIBA

2019

SINVALDO RODRIGUES MORENO

MODELO HÍBRIDO DE DECOMPOSIÇÃO MULTIESTÁGIO APLICADO À PREVISÃO DE SÉRIES TEMPORAIS DE VELOCIDADES DE VENTO

Tese apresentada ao curso de Pós-Graduação em Engenharia Elétrica, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Dr. Leandro dos Santos Coelho

CURITIBA

2019

Catalogação na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

M843m Moreno, Sinvaldo Rodrigues
Modelo híbrido de decomposição multiestágio aplicado à previsão de séries temporais de velocidades de vento / Sinvaldo Rodrigues Moreno. – Curitiba, 2019.

Tese - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, 2019.

Orientador: Leandro dos Santos Coelho.

1. Energia eólica. 2. Análise de séries temporais. 3. Redes neurais (Computação). I. Universidade Federal do Paraná. II. Coelho, Leandro dos Santos. III. Título.

CDD: 621.312136

Bibliotecária: Vanusa Maciel CRB- 9/1928

MINISTÉRIO DA EDUCAÇÃO
SETOR DE TECNOLOGIA
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA ELÉTRICA - 40001016043P4

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA ELÉTRICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **SINVALDO RODRIGUES MORENO** intitulada: **Modelo Híbrido de Decomposição Multi-Estágio aplicado à Previsão de Séries Temporais de Velocidades de Vento**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 01 de Julho de 2019.

LEANDRO DOS SANTOS COELHO
Presidente da Banca Examinadora

RAPHAEL AUGUSTO DE S BENEDITO
Avaliador Externo (UTFPR)

GIDEON VILLAR LEANDRO
Avaliador Interno (UFPR)

VIVIANA COCCO MARIANI
Avaliador Externo (PUC/PR)

JULIO CÉSAR NIEVOLA
Avaliador Externo (PUC/PR)

DEDICATÓRIA

À minha esposa Quézia, pelo amor, carinho e principalmente pela compreensão quando optei por me dedicar a este trabalho.

Ao meu filho Felipe (Lipe) de apenas dois anos e meio, pelas pausas inesperadas na pesquisa, solicitadas através de um pedido de colo, ou de vários pedidos em uma única frase ainda mal pronunciada: *Papai qué passeá! Papai, papai oi, ..., papai brincá! Papai pintá, papai, ..., papai, ..., papai piscina, paia.*

AGRADECIMENTOS

À Deus, por ter me proporcionado saúde e perseverança para trilhar este caminho.

Ao Prof. Dr. Leandro, pela confiança, paciência e tempo dedicado à orientação deste trabalho.

Ao Prof. Dr. Vilson Roiz, por acreditar no meu projeto e recomendar ao PPGEE-UFPR a minha aceitação como aluno.

Ao Prof. Dr. Raphael Augusto (UTFPR) por recomendar-me ao PPGEE-UFPR, e por compor a banca.

Ao Dr. Julio César Nievóla e Dr. Gideon Villar pelas contribuições e pelo incentivo já na etapa da Qualificação.

A Profa. Dra. Viviana Cocco Mariani pela parceria na publicação de trabalhos de alto nível e por compor a banca.

Ao DELT e PPGEE-UFPR pela oportunidade disponibilizada a mim, favorecendo à concretização deste sonho.

EPÍGRAFE

Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes.

Marthin Luther King

RESUMO

O rápido desenvolvimento da tecnologia de geração de energia eólica, bem como o aumento da inserção da geração de energia eólica na matriz energética nacional, requer pesquisa e desenvolvimento em várias áreas dessa tecnologia. Como a energia eólica é dependente das variações do clima, caracterizando-a como variável e intermitente no longo prazo, modelos de previsão com grau de precisão para geração de energia eólica são reconhecidos como uma importante contribuição para a integração confiável da energia eólica em larga escala. Atualmente as incertezas na previsão de velocidades de vento tem impactado o modelo de operação do sistema elétrico em vários lugares do mundo, resultando em contingências não planejadas pelo operador do sistema elétrico para evitar a falta de abastecimento energético, devido aos erros de previsão de produção eólica. Esta tese propõe uma abordagem para estabilizar os erros de previsão de velocidades de vento. A solução proposta considera a hibridização de duas técnicas de decomposição que atuam em domínios distintos, ou seja, a técnica de análise espectral singular (*singular spectrum analysis*) que atua na decomposição de séries temporais no domínio do tempo, e a decomposição de modo variacional (*variational mode decomposition*), que é aplicada especificamente no domínio da frequência. A hibridização é considerada para construir um modelo de decomposição multiestágio que será acoplado ao modelo de previsão de séries temporais. Outros trabalhos que exploram a decomposição como estágio inicial do modelo de previsão, consideram normalmente, apenas um dos domínios de decomposição. O modelo híbrido proposto, dá ênfase na extração das características relevantes da série temporal que só podem ser identificadas em sua estrutura no domínio da frequência, tais como componentes harmônicas ou em oposição de fase. Já as características relevantes no domínio do tempo são identificadas no segundo estágio da decomposição, onde o principal objetivo é reduzir a componente residual a níveis de variância limitados a unidade, de forma que se assemelhem ao ruído branco gaussiano, sendo esta uma condição requerida nos modelos clássicos de decomposição, mas não necessária no modelo proposto. A hibridização dessas duas técnicas busca mitigar um problema presente na técnica de análise espectral singular, relacionado a dificuldade em obter separabilidade aproximada das componentes de um sinal no domínio do tempo. Os resultados obtidos, através da decomposição de várias séries temporais de velocidades de vento, mostram que a abordagem proposta consegue extrair componentes descorrelacionadas, ou com um grau mínimo de correlação, separando de forma satisfatória as componentes de tendência, sazonalidade e componentes harmônicas. Como contribuição do modelo de decomposição proposto nesta tese, pode-se destacar a estabilização dos erros de previsão, independente da técnica de previsão adotada, pois cada componente obtida tem um comportamento específico, reduzindo desta forma a complexidade do modelo previsor. O modelo híbrido foi acoplado a duas técnicas de previsão para avaliar a hipótese considerada na pesquisa. A primeira técnica pertence a classe de modelos paramétricos, e são os modelos auto regressivos e de média móvel integrada (ARIMA), a segunda técnica pertencente aos modelos não paramétricos e ao grupo de inteligência computacional, sendo adotado uma rede neural com

unidades de memória de longo e curto prazo (LSTM). Estas técnicas foram aplicadas para à obtenção de previsões para diferentes séries temporais de 1 a 12 passos-a-frente e os resultados, quando comparados aos obtidos sem a etapa de decomposição, favorecem ao modelo sugerido, mesmo utilizando séries temporais com pequeno número de amostras.

Palavras-chave: Previsão de velocidades de vento. Energia eólica. Séries temporais. Análise espectral singular. Decomposição de modo variacional.

ABSTRACT

Rapid growth in wind power industry, as well as increase of wind generation penetration on national grid, requires continuous researching in specific areas related to wind energy forecasting. As wind power is weather dependent, variable and intermittent over various time-scales, thus accurate forecasting of wind power is recognized as a major contribution for reliable large-scale wind power integration. Currently, uncertainties in the prediction of wind speeds have impacted the oepration model of the electrical system in several places of the world, resulting into unplanned contingencies conducted by the electrical system operator, aiming to avoid the lack of energy supply, mostly due to errors of wind power's prediction, shuch as overestimation. In this thesis is propposed an approach aiming to stabilize the erros into wind speeds prediction's models. The proposed apporach considers the hybridization of two decomposition techniques that operate in distinct domains. The first, singular spectrum analysis technique, that acts in time domain to perform the decomposition of time series, and the second, variational mode decomposition, which is applied specifically in the frequency domain. The hybridization scheme is considered to construct a multistage decomposition model that will be coupled as first stage into time series prediction model. Other works that explore decomposition as an early stage of the prediction model, usually consider only one specific and limited domain to apply the decomposition step. The proposed hybrid model emphasizes the extraction of relevant time series characteristics or structure, that can only be identified in frequency domain, such as harmonic or opposite phase components. However, the relevant characteristics which remains into time series structure, will be identified in the second decompostion step, where the main objective is reduce the variance of residual component to levels limited to unity, resembling the white Gaussian noise, which one being a sufficient but not required condiction in the proposed model. The proposed hybridization of these two techniques seeks to mitigate a problem present in the singular spectral analysis technique, related to the difficulty in obtaining approximate separability of the components of a signal in the time domain. The results obtained as decomposition of several time series of wind velocities, can show that the proposed approach is able to extract uncorrelated components, or with a minimum degree of correlation, satisfactorily separating the components of trend, seasonality and harmonic components. As a contribution of the decomposition model proposed in this thesis, it is possible to highlight the stabilization of prediction errors, independent of the prediction technique adopted, since each component obtained has a specific behavior, reducing in this way, the complexity of the predictive model. The hybrid model was coupled to two forecasting techniques to evaluate the hypothesis considered in this research. The first technique belongs to the class of parametric models, the well-known auto-regressive and integrated moving average models (ARIMA), the second technique belonging to the non-parametric models and the group of computational intelligence, was a neural network with short and long memory units term (LSTM). These techniques were applied to obtain 1 to 12 steps-ahead forecasting intervals for different time series, and the results, when compared to those obtained without the decomposition

step, highlight the characteristics of the proposed model, even when applied to time series with small number of samples.

Key-words: Wind forecasting. Wind energy. Time series. Singular spectrum analysis. Variational mode decomposition.

LISTA DE FIGURAS

Figura 1 Evolução da capacidade instalada de energia eólica 2

Figura 2 Matriz energética brasileira 2

Figura 3 Espectro da série de velocidades de vento amostradas a cada 10min. .. 11

Figura 4 Espectro de 5 componentes da série de velocidades de vento amostradas a cada 10min. 12

Figura 5 Etapas na identificação de sistemas 15

Figura 6 Etapas do modelo proposto nesta tese 17

Figura 7 Estágios da SSA 23

Figura 8 Série chegada de passageiros em aeroportos nos USA 30

Figura 9 Decomposição da série com $L = 20$ 30

Figura 10 Contribuição de cada componente na explicação da variância da série . 31

Figura 11 Primeira componente reconstruída 32

Figura 12 Segunda componente reconstruída 33

Figura 13 Terceira componente reconstruída 33

Figura 14 Quarta componente reconstruída 34

Figura 15 Quinta componente reconstruída 34

Figura 16 Correlação ponderada das componentes obtidas por meio do SSA 35

Figura 17 Correlação cruzada das componentes obtidas por meio do SSA 36

Figura 18 Influência de α no filtro de Wiener 39

Figura 19 Sinal C - composto por conteúdo nas frequências 18 e 50Hz 42

Figura 20 Sinais misturados que originaram o Sinal C 42

Figura 21 Espectro do sinal C 43

Figura 22 Espectro do sinal C e modos estimados pelo algoritmo VMD 43

Figura 23 Modo 1 de $18Hz$ reconstruído e comparado ao sinal A 44

Figura 24 Modo 2 de $50Hz$ reconstruído e comparado ao sinal B 44
Figura 25 Coerência entre o modo 1 de $18Hz$ e o sinal C 45
Figura 26 Coerência entre o modo 2 de $50Hz$ e o sinal C 46
Figura 27 Estrutura do modelo proposto 53
Figura 28 Centro de frequências para cinco modos convergência após 90 iterações 54
Figura 29 Estrutura do modelo proposto sem a etapa de previsão 57
Figura 30 Série de velocidades de vento amostrada a cada 10min - 01 de janeiro 2019 60
Figura 31 Série de velocidades de vento amostrada a cada 10min - 02 de janeiro 2019 61
Figura 32 Série de velocidades de vento amostrada a cada 10min - 03 de janeiro 2019 61
Figura 33 Série de velocidades de vento amostrada a cada 10min - 04 de janeiro 2019 62
Figura 34 Série de velocidades de vento amostrada a cada 10min - 05 de janeiro 2019 62
Figura 35 Série de velocidades de vento amostrada a cada 1 dia - trimestre composto pelos meses de novembro 2018 a janeiro 2019 63
Figura 36 Função de autocorrelação parcial - série de velocidades de vento com 140 amostras 63
Figura 37 Decomposição VMD-SSA da série temporal velocidades de vento (10min) de 01/jan/2019 67
Figura 38 Correlação entre componentes obtidas pela decomposição VMD-SSA para série vel. vento(10min) de 01/jan/2019 67
Figura 39 Decomposição VMD-SSA da série temporal velocidades de vento (10min) de 01/jan/2019 80
Figura 40 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/jan/2019 80
Figura 41 Decomposição VMD-SSA para série temporal velocidades de vento (10min)

de 02/jan/2019 81

Figura 42 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/jan/2019 81

Figura 43 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/jan/2019 82

Figura 44 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/jan/2019 82

Figura 45 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/jan/2019 83

Figura 46 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/jan/2019 83

Figura 47 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/jan/2019 84

Figura 48 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/jan/2019 84

Figura 49 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/jan/2019 85

Figura 50 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/jan/2019 85

Figura 51 Decomposição VMD-SSA da série temporal velocidades de vento (10min) de 20/jan/2019 86

Figura 52 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/jan/2019 86

Figura 53 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/jan/2019 87

Figura 54 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/jan/2019 87

Figura 55 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/dez/2018 88

Figura 56 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/dez/2018 88

Figura 57 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/dez/2018 89
Figura 58 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/dez/2018 89
Figura 59 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/dez/2018 90
Figura 60 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/dez/2018 90
Figura 61 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/dez/2018 91
Figura 62 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/dez/2018 91
Figura 63 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/dez/2018 92
Figura 64 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/dez/2018 92
Figura 65 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/dez/2018 93
Figura 66 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/dez/2018 93
Figura 67 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/dez/2018 94
Figura 68 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/dez/2018 94
Figura 69 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/dez/2018 95
Figura 70 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/dez/2018 95
Figura 71 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/nov/2018 96

Figura 72 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/nov/2018 96

Figura 73 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/nov/2018 97

Figura 74 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/nov/2018 97

Figura 75 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/nov/2018 98

Figura 76 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/nov/2018 98

Figura 77 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/nov/2018 99

Figura 78 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/nov/2018 99

Figura 79 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/nov/2018 100

Figura 80 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/nov/2018 100

Figura 81 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/nov/2018 101

Figura 82 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/nov/2018 101

Figura 83 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/nov/2018 102

Figura 84 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/nov/2018 102

Figura 85 Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/nov/2018 103

Figura 86 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/nov/2018 103

Figura 87 Decomposição VMD-SSA da série temporal velocidades de vento (média diária) trimestre 01/nov/2018 à 31/jan/2019 104

Figura 88 Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (média diária) trimestre 01/nov/2018 à 31/jan/2019 104

Figura 89 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 01/jan/2019 através do método de Box & Jenkins. 110

Figura 90 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 01 e 02/Jan/19. 110

Figura 91 Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 01 e 02/Jan/19. 111

Figura 92 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 01/jan/2019 através da LSTM 112

Figura 93 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 01/jan/19. 112

Figura 94 Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 01/jan/19. 113

Figura 95 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 02/Jan/2019 através do método de Box & Jenkins 116

Figura 96 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 02 e 03/jan/19. 117

Figura 97 Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 02 e 03/jan/19. 117

Figura 98 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 02/jan/2019 através da LSTM 118

Figura 99 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 02/jan/19. 118

Figura 100 Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 02/jan/19. 119

Figura 101 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 03/jan/2019 através do método de Box & Jenkins 122

Figura 102 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 03 e 04/jan/19. 123

Figura 103 Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 03 e 04/jan/19. 123

Figura 104 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 03/jan/2019 através da LSTM 124

Figura 105 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 03/jan/19. 125

Figura 106 Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 03/jan/19. 125

Figura 107 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 04/jan/2019 através do método de Box & Jenkins 128

Figura 108 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 04 e 05/jan/19. 128

Figura 109 Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 04 e 05/jan/19. 129

Figura 110 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 04/jan/2019 através da LSTM 130

Figura 111 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 04/jan/19. 130

Figura 112 Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 04/jan/19. 131

Figura 113 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 05/jan/2019 através do método de Box & Jenkins 134

Figura 114 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 04 e 05/jan/19. 135

Figura 115 Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 05 e 06/jan/19. 135

Figura 116 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 05/jan/2019 através da LSTM 136

Figura 117 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 05/jan/19. 136

Figura 118 Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 05/jan/19. 137

Figura 119 Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 01/nov/2018 à 31/jan/2019 através do método de Box & Jenkins 141

Figura 120 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 01/nov/2018 e 31/jan/2019. . 141

Figura 121 Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 01/nov/2018 e 31/jan/2019. 142

Figura 122 Previsão 10 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento entre 01/nov/18 à 21/jan/2019 através da LSTM 143

Figura 123 Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 01/nov/2018 à 21/jan/2019. . 143

Figura 124 Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores diários observados no período de 01/nov/2018 à 21/jan/2019. 144

Figura A.1 Estratégia recursiva para previsão 5-passos-a-frente 160

Figura A.2 Estratégia direta para previsão 5-passos-a-frente 161

Figura A.3 Estratégia dirrec para previsão 5-passos-à-frente 162

Figura A.4 Estratégia mimo para previsão 5-passos-à-frente 163

Figura D.1 Estrutura da rede LSTM como 1 bloco de memória na camada escondida 170

Figura D.2 Estrutura do módulo de repetição da LSTM 171

Figura D.3 Estrutura do módulo de repetição da LSTM 172

Figura D.4 Estrutura do *input gate* da LSTM 172
Figura D.5 Estrutura da etapa de atualização da célula LSTM 173
Figura D.6 Estrutura da etapa de saída da célula LSTM 173

LISTA DE TABELAS

Tabela 1 Desvio observado entre a geração prevista e realizada 4

Tabela 2 Parâmetros adotados no VMD-SSA 66

Tabela 3 Análise da *fac* e *facp* para determinar o modelo de Box & Jenkins 69

Tabela 4 Parâmetros das componentes da série velocidades de vento 01/jan/2019 69

Tabela 5 Parâmetros das componentes da série velocidades de vento 02/jan/2019 70

Tabela 6 Parâmetros das componentes da série velocidades de vento 03/jan/2019 71

Tabela 7 Parâmetros das componentes da série velocidades de vento 04/jan/2019 72

Tabela 8 Parâmetros das componentes da série de velocidades de vento 05/jan/2019 73

Tabela 9 Parâmetros das componentes da série de velocidades de vento para o trimestre 01/nov/2018 a 31/jan/2019 74

Tabela 10 Parâmetros da série de velocidades de vento para os dias 01 a 05 de janeiro de 2019 75

Tabela 11 Parâmetros para a série de velocidades de vento do trimestre 01/nov/2018 a 31/jan/2019 75

Tabela 12 Parâmetros Rede LSTM 76

Tabela 13 Conjuntos de dados para a etapa de previsão 107

Tabela 14 Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 02/jan/2019 114

Tabela 15 Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 01/jan/2019 115

Tabela 16 Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 03/jan/2019 120

Tabela 17 Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 02/jan/2019 121

Tabela 18 Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 04/jan/2019 126

Tabela 19 Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 03/jan/2019 127

Tabela 20 Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 05/jan/2019 132

Tabela 21 Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas últimas horas do dia 04/jan/2019 133

Tabela 22 Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 06/Jan/2019 138

Tabela 23 Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 05/jan/2019 139

Tabela 24 Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA no trimestre 01/nov/2018 à 31/jan/2019 145

Tabela 25 Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM no trimestre 01/nov/2018 à 31/jan/2019 146

LISTA DE ALGORITMOS

Algoritmo 1 Pseudocódigo SSA 27
Algoritmo 2 Otimização Completa do VMD 40
Algoritmo 3 Coerência Espectral 41
Algoritmo 4 Modelo Híbrido VMD-SSA 56
Algoritmo E.1 Otimizador Adam 175

LISTA DE SIGLAS

ADAM Estimação adaptativa do momento do gradiente (do inglês, *adaptive moment estimation*)

AM Modulação em amplitude (do inglês, *amplitude modulate*)

ANFIS Sistema nebuloso de inferência neuro adaptativo (do inglês, *adaptive neuro fuzzy inference system*)

ANN Rede neural artifical (do inglês, *artificial neural network*)

AR Auto regressivo (do inglês, *auto regressive*)

ARIMA Auto regressivo integrado de médias móveis (do inglês, *auto regressive integrated moving average*)

ARMA Auto regressivo de médias móveis (do inglês, *auto regressive moving average*)

BA Algoritmo do morcego (do inglês, *bat algorithm*)

BIC Critério bayesiano de informação (do inglês, *bayesian information criterion*)

BLIMF Modo intrínseco Limitado em banda (do inglês, *band limited intrinsic mode functions*)

CG Gradiente conjugado (do inglês, *conjugate gradient*)

CI Inteligência computacional (do inglês, *computational inteligence*)

CNNSVM Máquinas convolucionais de vetor suporte (do inglês, *convolutional support vector machine*)

CSA Algoritmo de busca do cuco (do inglês, *cuckoo search algorithm*)

DBN Rede de crenças profundas (do inglês, *deep belief network*)

DBNN Rede neural bayesiana dinâmica (do inglês, *dynamic bayesian neural network*)

DBSCAN Método de clusterização baseado em densidade para aplicação com ruído (do inglês, *density-based spatial clustering of applications with noise*)

EA Algoritmo evolucionário (do inglês, *evolutionary algorithm*)

EEMD Grupo de decomposição de modo empírico (do inglês, *ensenble empirical mode decomposition*)

ELM Máquinas de aprendizado extremo (do inglês, *extreme learning machine*)

EMD Decomposição de modo empírico (do inglês, *empirical mode decomposition*)

EWT Transformada de ondeletas empírica (do inglês, *empirical wavelet transform*)

FEEMD Grupo de decomposição rápida de modo empírico (do inglês, *fast ensemble empirical mode decomposition*)

FFT Versão rápida da transformada de Fourier (do inglês, *fast Fourier transform*)

FM Modulação em frequênca (do inglês, *frequency modulate*)

FT Transformada de Fourier (do inglês, *Fourier transform*)

FVD Previsão do grau de validade (do inglês, *fast validity degree*)

GA Algoritmo genético (do inglês, *Genetic algorithm*)

GARCH Auto regressivo condicionalmente heterocedástico generalizado (do inglês, *generalized auto regressive conditional heteroscedastic*)

GPR Processo gaussiano regressivo (do inglês, *Guassian rrocess regresion*)

GRNN Rede neural de regresão geral (do inglês, *General regression neural network*)

GSA Algoritmo de busca gravitacional (do inglês, *Gravitational search algorithm*)

IMF Função de modo intrinseco (do inglês, *intrinsic mode functions*)

LSSVM Máquina de vetorsSuporte à mínimos quadrados (do inglês, *least squares support vector machine*)

LSTM Unidades de memória de longo-curto prazo (do inglês, *long short term memory*)

MAE Erro médio absoluto (do inglês, *mean absolute error*)

MAPE Erro médio percentual absoluto (do inglês, *mean absolute percentage error*)

MDN Rede de densidade de mistura (do inglês, *mixture density networks*)

MSE Erro médio quadrático (do inglês, *mean squared error*)

NARX Rede neural auto regressiva não-linear com entrada exógena (do inglês, *nonlinear auto regressive network with exogenous inputs*)

PAR(p) Periódico auto regressivo de ordem p (do inglês, *periodic auto regressive p-order*)

PCA Análise de componentes principais (do inglês, *principal components analysis*)

PSO Otimização por enxames de partículas (do inglês, *particle swarm optimization*)

RBFNN Rede neural de base radial (do inglês, *Radial basis function neural network*)

RMSE Raiz do erro médio quadrático (do inglês, *root mean squared error*)

RRMSE Erro médio quadrático relativo (do inglês, *relative root mean squared error*)

SDA Algoritmo de descida mais íngreme (do inglês, *steepest descent algorithm*)

SSA Análise espectral singular (do inglês, *singular spectrum analysis*)

STSNN Rede neurais de tempo estocástico (do inglês, *stochastic time strength neural network*)

SVD Decomposição de valor Singular (do inglês, *singular value necomposition*)

SVM Máquina de suporte vetor (do inglês, *suport vector machine*)

SVR Máquina de suporte vetorial para regressão (do inglês, *support vector regression*)

UHE Usinas hidrelétricas

VMD Decomposição de modo variacional (do inglês, *variational mode decomposition*)

WD Decomposição de ondeletas (do inglês, *Wavelet decomposition*)

WPD Decomposição de pacotes de ondeletas (do inglês, *wavelet package decomposition*)

LISTA DE SÍMBOLOS

$*$ Operador de convolução

. Multiplicação elemento-a-elemento no Matlab™

$\cdot$ Multiplicação algébrica entre dois números

δ Delta de *Dirac*

$\eta(t)$ Ruído branco Gaussiano

$\gg$ Muito maior que

$\hat{\omega}_k$ Frequência central do sinal de entrada após a aplicação da transformada de Fourier

$\hat{u}_k$ Sinal de entrada após a aplicação da transformada de Fourier

$\hat{y}_i$ i-ésimo valor previsto de y

$\hat{y}_{N+K}$ Série temporal temporal de tamanho $N+K$, com K passos-a-frente previstos

λ Multiplicador de lagrange

$\langle x, y \rangle$ Produto interno entre x e y

$\mathbb{C}^n$ Espaço vetorial complexo de ordem n

$\mathbb{R}$ Domínio dos números reais

$\mathbb{R}^N$ Conjunto de números reais de dimensão N

$\mathbf{H}(.)$ Transformada de Hilbert

$\mathbf{I}_n$ Matriz identidade de ordem n

$\mathbf{X}$ Matriz X

$\mathbf{X}$ Matriz de trajetórias *Hankel*

$\mathbf{Y}$ Matriz Y

$\mathbf{Y}^T$ Transposta *Hermitiana* da matriz Y

ω Frequência do sinal

ω_k Frequência central do sinal de entrada

∂_t Derivada parcial

$\Phi^+(t)$ Operador de modulação

$\Phi^-(t)$ Operador de demodulação

ϕ_i Fase inicial do sinal

π Número pi, equivalente a $\approx 3,1415$

ρ_{12}^{ω} Correlação ponderada ou *Weighted Correlation* (ω-*correlation*)

$\sum_{i=1}^{N}$ Somatório de todos os valores de $i = 1$ à N

$\mathbf{U}$ Autovetores à esquerda

$\mathbf{V}$ Autovetores à direita

$\tilde{Y}_N^K$ Série temporal filtrada

abs Valor do módulo de um número complexo

BW Largura de banda do sinal

BW_{AM} Largura de banda do sinal AM

BW_{FM} Largura de banda do sinal FM

c_{xy} Covariança cruzada

conj Complexo conjugado de um número no domínio $\mathbb{C}$

I Conjunto de componentes a serem reagrupadas

j Parte complexa de um número, cujo valor de $j^2 = -1$

L Comprimento da janela

L^2 Quadro da norma L

max Máximo valor

min Mínimo valor

N Comprimento da série temporal

r_{xy} Correlação cruzada ou *cross-correlation*

s_x Desvio Padrão de x

s_y Desvio Padrão de y

T Operador de transposição matricial

t Tempo (s)

u_k Sinal de entrada

X_r Série temporal de resíduos

y_i i-ésimo valor de y

Y_N Série temporal de tamanho N

SUMÁRIO

1 Introdução **1**

1.1 Estado da arte 5

1.2 Objetivos 16

1.2.1 Objetivo Geral 16

1.2.2 Objetivos Específicos 16

1.3 Organização da Tese 17

2 Fundamentação Teórica **21**

2.1 Análise Espectral Singular 21

2.1.1 Estágio de Decomposição 23

2.1.2 Estágio de Reconstrução 25

2.1.3 Correlação Ponderada 28

2.1.4 Correlação Cruzada 28

2.1.5 Exemplo de Aplicação SSA 29

2.1.6 Aspectos a serem destacados 36

2.2 Decomposição de Modo Variacional 36

2.2.1 Coerência Espectral 40

2.2.2 Exemplo de Aplicação VMD 41

2.2.3 Avaliação da Coerência 45

2.2.4 Aspectos a serem destacados 46

2.3 Resumo do Capítulo 47

3 Modelo Híbrido Proposto **48**

3.1 Modelo Híbrido VMD-SSA 48
3.1.1 Demodulação 50
3.1.2 Algoritmo do Modelo Proposto 54
3.2 Resumo do Capítulo 58

4 Aplicação do Modelo VMD-SSA 59
4.1 Experimentos Numéricos 59
4.1.1 Parâmetros do algoritmo VMD-SSA 64
4.1.2 Determinação dos parâmetros do modelo de Box & Jenkins 68
4.1.3 Parâmetros da rede LSTM 76
4.2 Resumo do Capítulo 77

5 Avaliação do Desempenho do Modelo Híbrido de Decomposição 78
5.1 Resultados da Decomposição 78
5.2 Resumo do Capítulo 104

6 Avaliação do Desempenho do Modelo Híbrido de Previsão 106
6.1 Resultados das Previsões 106
6.2 Métricas de Desempenho 108
6.2.1 Resultados das previsões para os dados observados em 01/jan/2019 109
6.2.2 Previsões para os dados observados em 02/Jan/2019 115
6.2.3 Previsões para os dados observados em 03/Jan/2019 121
6.2.4 Previsões para os dados observados em 04/Jan/2019 127
6.2.5 Previsões para os dados observados em 05/Jan/2019 133
6.2.6 Previsões para dados observados no trimestre 01/Nov/2018 à 31/Jan/2019 . 139
6.3 Resumo do Capítulo 146

7 Conclusão e Pesquisa Futura 148

7.1 Análise das Decomposições 148
7.2 Análise das Previsões 149
7.3 Pesquisa Futura 151

Referências 153

Apêndice A – Estratégias de Previsão 159
A.1 Estratégia Recursiva 159
A.2 Estratégia Direta 160
A.3 Estratégia DirRec 161
A.4 Estratégia MIMO 162
A.5 Estratégia DIRMO 163

Apêndice B – Modelagem Box e Jenkins 165
B.1 Critério de Informação Bayesiano - BIC 168

Apêndice C – Relação de Recorrência Linear - LRR 169

Apêndice D – Redes Neurais LSTM 170

Apêndice E – Otimizador Adam 174

Apêndice F – Publicações e Premiação 176
F.1 Trabalhos publicados 176
F.2 Premiação 177

1 INTRODUÇÃO

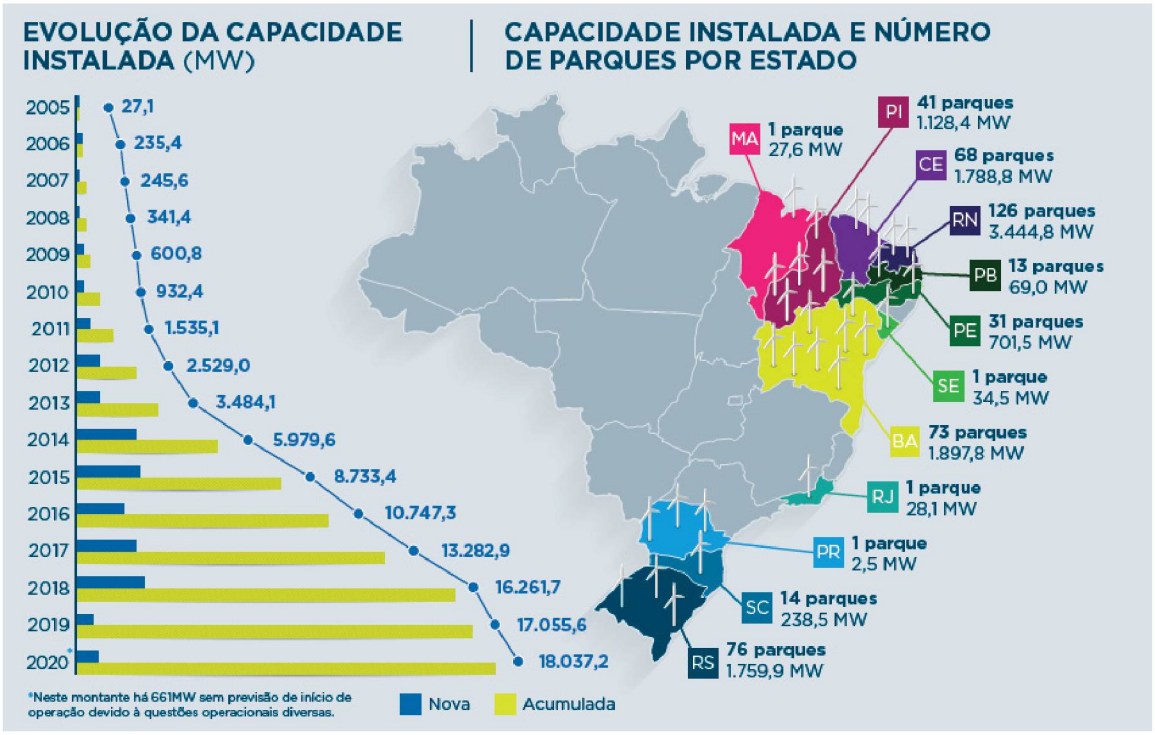
Geração eólica é uma das tecnologias de produção de energia elétrica ainda em desenvolvimento no Brasil, porém já apresenta grande penetração na matriz energética nacional, sendo uma das principais fontes de energia renovável (ABEEóLICA, 2018). Especificamente devido a necessidade de desenvolvimento da matriz energética, uma crescente integração de novos parques eólicos ao sistema interligado nacional (SIN) está sendo observada. O relatório de 2017 da Associação Brasileira de Energia Eólica (ABEEóLICA, 2017), aponta um crescimento expressivo da capacidade instalada de energia eólica conectada ao SIN nos últimos dois anos, além de apresentar os montantes já negociados em leilões para instalação futura, conforme dados da Figura 1. Esta integração, ao ser finalizada, representará um montante expressivo do total da capacidade de atendimento a demanda do SIN.

Em contrapartida, a matriz energética nacional possui na geração de base o sistema hidroelétrico, complementado pelo térmico, caracterizando desta forma um sistema hidrotérmico. Uma das principais características do sistema hidrotérmico é o critério de suprimento da demanda sujeito ao mínimo custo futuro, onde o déficit máximo admissível, ou seja, o risco de não suprimento, é de 5%. Com a inserção de fontes renováveis, tais como pequenas centrais hidrelétricas (PCHs), eólicas e solar, houve um expressivo crescimento da oferta, conforme Figura 2 (ABEEóLICA, 2018). Porém a confiabilidade do sistema ainda recai sobre as usinas com grandes reservatórios, chamadas de usinas hidrelétricas (UHE), pois a matriz de geração de energia elétrica constituída somente por fontes renováveis, sem qualquer tipo de armazenamento ou regularização da energia, não proporcionam garantias de suprimento.

Os parques de geração eólica são constituídos por um número razoável de aerogeradores e ao totalizar a potência injetada por um parque, a mesma se torna significativa para o centro de gravidade da carga do sistema onde o parque esta conectado. Mesmo com a crescente integração de parques eólicos aos sistemas de potência ocorrendo em vários países, a geração eólica ainda apresenta um nível de incerteza elevado, recebendo

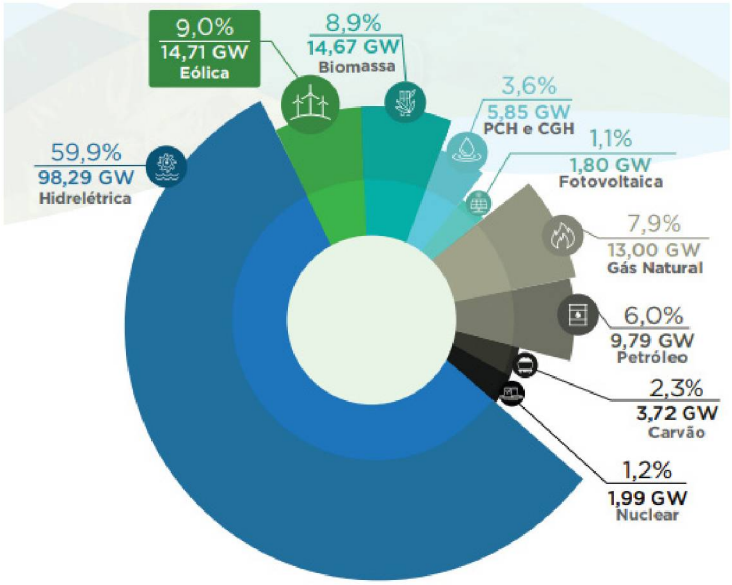
a classificação como fontes intermitentes de geração de energia.

Figura 1: Evolução da capacidade instalada de energia eólica



Fonte: O autor (2019) - adaptado de Abeeólica (2017).

Figura 2: Matriz energética brasileira



Fonte: O autor (2019) - adaptado de Abeeólica (2018).

A classificação como fonte intermitente (ou não confiável), atribuída a geração eólica atualmente na literatura, se dá devido as caóticas flutuações da velocidade do vento,

de forma que a energia eólica não tem capacidade de suprir de maneira estável a demanda do sistema de potência (LEI et al., 2009). A intermitência é motivada principalmente por contínuas e caóticas flutuações das velocidades do vento e a ausência de ferramentas que forneçam previsões coerentes. Esse fato tem sido observado em vários países onde a penetração da geração eólica na matriz energética é expressiva (LIU et al., 2013).

Um exemplo do impacto das fontes eólicas na programação diária de operação e confiabilidade do SIN, é apontado no boletim mensal de geração eólica de março de 2017 do operador nacional do sistema (ONS) (ONS, 2017). Pode-se observar que existem desvios negativos entre a geração programada prevista e a realizada da ordem de até -18,6% e desvios positivos de até 140,9%, conforme Tabela 1 (ONS, 2017). O erro médio percentual absoluto entre o programado previsto e o realizado no mês de março de 2017, considerando o conjunto de eólicas monitorada pelo ONS, foi da ordem de 38,0% (ONS, 2017).

Os erros nas previsões de velocidades de vento impactam de forma significativa a estimação da produção de um parque eólico, dado que a equação da relação de transformação da velocidade de vento em potência mecânica irá variar com o cubo da velocidade (KUSIAK et al., 2009), conforme Equação 1.1,

$$P = \frac{1}{2}\rho c_p A v^3 \tag{1.1}$$

onde ρ é densidade do ar seco (kg/m^3) medida à altura do cubo do rotor da turbina eólica e depende das condições atmosféricas da região, A representa a área de varredura da turbina (m^2), c_p é o coeficiente de potência de turbina, v velocidade do vento (m/s) e P potência em Watt (W).

Uma das causas da não confiabilidade é associada a dificuldade na obtenção de previsões de velocidades de vento, onde para obtê-las com acurácia é extremamente difícil, principalmente para planejamento de longo prazo (LIU et al., 2012). As velocidades do vento podem ser facilmente influenciadas por obstáculos e pela geografia do terreno. A velocidade também sofre variações com a altura, desta foma, a característica aleatória do vento é muito significante (LEI et al., 2009). Sendo necessário energia de reserva (ou reserva de capacidade hídrica) para atender possíveis déficits, quando as previsões sobre-estimam às velocidades e consequentemente a produção de energia eólica realizada é inferior a prevista.

Tabela 1: Desvio observado entre a geração prevista e realizada

Estado	Usina	Prog. MWm	Real. MWm	Desvio %	MAPE[1]	NMAPE[2]
CE	UEE Taiba Águia	3,9	3,6	-8,4	100,8	13,4
CE	UEE Taiba Andorinha	2,6	2,3	-10,4	94,5	14,4
CE	UEE Volta do Rio	1,9	1,5	-18,6	87,9	3,9
BA	Conj. Licínio de Almeida	29,4	32,1	9,3	44,0	13,9
BA	Conj. Planaltina	26,0	31,2	-20,2	38,3	14,7
PE	Conj. Caetés II	69,5	75,1	8,0	70,4	13,5
PE	Conj. Paranatama	39,5	41,7	5,6	63,8	15,2
PE	Conj. São Clemente	91,7	90,4	-1,4	61,2	13
PE	Conj. Tacarutu	39,7	32,4	-18,3	56,7	19,5
PI	Conj. Chapada I	29,8	50,9	70,8	238,5	17,9
PI	Conj. Chapada II	23,6	33,9	43,9	210,1	12,6
PI	Conj. Chapada III	24,9	35,9	44,5	217,1	11,8
PI	Conj. Chapada IV	20,3	35,6	75,2	240,8	9,5
RS	Conj. Atlantica	50,1	49,5	-1,2	48,3	13,1
RS	Conj. Livramento 2	65,4	65,1	-0,5	47,5	12,8
RS	Conj. Marmeleiro 2	20,4	49,1	140,9	250,9	32,7
RS	Conj. Quinta 138 kV	35,8	53,4	49,0	102,1	20,4
RS	Sta Vitória do Palmar	191,9	203,2	5,9	76,3	13,3
RS	Conj. Viamão 3	13,5	23,1	70,9	137,9	20,9
RS	Grupo Ventos da Lagoa	17,8	17,8	0,0	59,9	13,8
RS	Grupo Vento do Litoral	16,6	17,6	5,8	56,0	12,3
RS	Grupo Ventos do Sul	37,6	38,2	1,4	61,9	11,6
RS	Grupo Ventos dos Índios	14,5	16,4	13,6	62,6	13,0

Fonte: O autor (2019) - adaptado de (ONS, 2017).

[1] MAPE (*Mean Absolute Percentage Error*) – Erro médio percentual absoluto. Calculado através da expressão abaixo para todas as horas no mês de referência, desconsiderando os horários com geração programada igual a zero, $MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|Prog._{(i)} - Real_{(i)}|}{Prog._{(i)}}$.

[2] NMAPE (*Normalized Mean Absolute Percentage Error*) – Erro médio percentual absoluto normalizado: $NMAPE = \frac{1}{N}\sum_{i=1}^{N\cdot P_{inst}}|Prog._{(i)} - Real_{(i)}|$.

Por outro lado, quando as previsões sub-estimam a produção, além de um pla-

nejamento detalhado da operação do sistema hidrelétrico, visando minimizar o impacto no preço da energia, dado que devido possíveis cenários hidrológicos desfavoráveis, usinas térmicas com custo mais elevado devem ser acionadas para atender a demanda do sistema, porém a produção das eólicas real superam o valor previsto, comprometendo desta forma, a operação ótima sob ponto de vista de custo marginal do sistema.

Há ainda possíveis cenários onde existem limitação de fluxo de potência, restrições elétricas e hidráulicas, entre outras, onde um planejamento adequado deve ser feito afim de controlar a estabilidade do sistema frente as rampas de produção provocadas por rajadas de vento com pequena duração. Sendo importante informações coerentes de previsão de velocidades de vento, dado que alterações súbitas de produção são vistas como perturbações no sistema, influenciando na estabilidade rotórica, de frequência e de tensão. Diante do exposto, a necessidade de modelos mais precisos e confiáveis para a previsão de séries temporais de velocidades de vento é uma realidade (LEI et al., 2009). Pode-se concluir desta forma, que ainda existe necessidade de pesquisa e desenvolvimento de modelos aplicados à previsão de velocidades de vento, voltado ao mercado de geração de energia eólica no Brasil.

1.1 ESTADO DA ARTE

A previsão de séries temporais tem aplicações multidisciplinares tais como: mercados financeiros (HIRANSHA et al., 2018), hidrologia (YASEEN et al., 2019), previsão do tempo (MEHRKANOON, 2019), consumo de energia (XIAO et al., 2018), previsão de velocidades de vento para produção de energia eólica (MORENO; COELHO, 2018). Séries temporais coletadas através de realizações de sistemas reais normalmente são compostas por uma mistura de componentes, que são denominados como componentes sazonais, de tendência e de ruído aditivo (MORETTIN; TOLOI, 2006; CHATEFIELD, 2000). Consequentemente, modelar corretamente essas componentes é essencial para obter previsões do comportamento do sistema com coerência. A separação incorreta leva ao ajuste de modelos de previsão que podem subestimar, por exemplo, a componente de tendência. Porém quando se trata da modelagem de velocidades de vento para previsão da geração eólica, devido a natureza aleatória do fenômeno físico gerador da série, o grau de dificuldade em caracterizar cada uma das componentes de forma independente é elevado.

A hipótese a ser considerada nesta pesquisa estabelece que: *a utilização de técnica híbrida para a decomposição multiestágio da série temporal univariada em componentes*

aditivas, permite obter melhor separação das componentes, resultando em previsões de 1 *a K passos-a-frente com maior acurácia.*

Especificamente para a previsão de velocidades de vento, um número significativo de pesquisadores investigaram várias técnicas para a obtenção de modelos capazes de realizar previsões com maior precisão. Um crescimento expressivo se deu no desenvolvimento de modelos baseados em abordagens não paramétricas, ou técnicas orientadas a extrair conhecimento dos dados observados, principalmente devido ao fato de não ser necessário pressupostos de normalidade dos resíduos, de estacionariedade da série temporal ou de uma estrutura de modelo predefinida (GONG et al., 2016; CAVALCANTE et al., 2016; BARUNíK; KřEHLíK, 2016; ZIEL et al., 2016).

As abordagens que se enquadram nesse rol, incluem técnicas como máquina de suporte vetor (do inglês, *suport vector machine*, SVM), transformada de ondeletas empírica (do inglês, *empirical wavelet transform*, EWT), redes neurais artificiais (do inglês, artificial neural metwork, ANN), sistema nebuloso de inferência neuro adaptativo (do inglês, *adaptive neuro fuzzy inference system*, ANFIS), máquinas de aprendizado extremo (do inglês, *extreme learning machine*, ELM), entre outras. Apesar da qualidade dos resultados, cada técnica possui suas próprias vantagens e limitações, porém ao combiná-las com outra técnica, há um expressivo ganho de qualidade nas previsões. Pode-se citar como exemplo, o modelo de previsão inteligente híbrido que combina máquina de vetor suporte à mínimos quadrados (do inglês, *least squares support vector machine*, LSSVM) e o modelo de cadeias Markov, proposto por Wang et al. (2015). Neste modelo, o LSSVM é utilizado como previsor e tem seus parâmetros de ajuste otimizados pelo algoritmo de otimização por enxames de partículas (PSO) combinado ao algoritmo de busca gravitacional (do inglês, *gravitational search algorithm*, GSA). Os resultados relatados pelos autores mostraram que a abordagem superou outras técnicas em termos de qualidade de previsão.

Outra abordagem híbrida utilizada para realizar previsões de velocidade do vento, aplicando máquina de suporte vetorial para regressão (do inglês, *support vector regression*, SVR) com parâmetros ajustados por algoritmo genético (do inglês, *genetic algorithm*, GA), foi proposta por Bonfil et al. (2016). Segundo os autores, o modelo híbrido superou modelos de persistência e auto regressivos bem conhecidos como o auto regressivo de ordem p (do inglês, *auto regressive*, AR), auto regressivo de médias móveis (do inglês, *auto regressive moving average*, ARMA) e auto regressivo integrado de médias móveis (do inglês, *auto regressive integrated moving average*, ARIMA). Um conjunto de rede de densidade de mistura (do inglês, *mixture density networks*, MDN) foi aplicado por Men

et al. (2016), para prever a velocidade do vento e a produção de energia eólica com horizonte de 72h à frente, obtendo resultados promissores quando comparado aos resultados dos modelos de redes neurais auto regressiva não linear com entrada exógena (do inglês, *nonlinear auto regressive network with exogenous inputs*, NARX), ARMA e auto regressivo condicionalmente heterocedástico generalizado (do inglês, *generalized auto regressive conditional heteroscedastic*, GARCH).

Como exemplos de hibridização na etapa de filtragem pode-se citar o trabalho realizado por Wang e Hu (2015), que utiliza EWT para filtrar a série original e aplicá-la aos modelos ARIMA, ELM, SVM e LSSVM para obter previsões independentes e que servem como entrada para um modelo de processo gaussiano regressivo (do inglês, *guassian process regresion*, GPR), responsável por fazer combinações das mesmas para gerar a previsão de velocidade de vento para dois parques eólicos localizados na China. O modelo apresentado por Xiao et al. (2017) utiliza a combinação de quatro técnicas híbridas formadas por um conjunto extenso de algoritmos de decomposição, otimização e filtragem, a saber: ANN, ARIMA, algoritmo do morcego (do inglês, *bat algorithm*, BA), algoritmo de busca do cuco (do inglês, *cuckoo search algorithm*, CSA), gradiente conjugado (do inglês, *conjugate gradient*, CG), algoritmo evolucionário (do inglês, *evolutionary algorithm*, EA), conjunto de decomposição de modo empírico (do inglês, *ensenble empirical mode decomposition*, EEMD), decomposição de modo empírico (do inglês,empirical mode decomposition, EMD), grupo de decomposição rápida de modo empírico (do inglês, *fast ensemble empirical mode decomposition*, FEEMD), previsão do grau de validade (do inglês, *fast validity degree*, FVD), GA, rede neural de regressão geral (do inglês, *general regression neural network*, GRNN), PSO, rede neural de base radial (do inglês, *radial basis function neural network*, RBFNN), algoritmo de descida mais íngreme (do inglês, *steepest descent algorithm*, SDA), análise espectral singular (do inglês, *singular spectrum analysis*, SSA), SVM, decomposição de ondeletas (do inglês, *wavelet decomposition*, WD) e decomposição de pacotes de ondeletas (*wavelet package decomposition*, WPD). Esses algoritmos são aplicados para realizar previsões de velocidades de vento, cabendo destacar que os melhores resultados foram obtidos com as técnicas que utilizam SSA e BA em sua estrutura.

Enfim há várias propostas de combinações de modelos na etapa de aprendizado, ou na etapa de previsão do padrão de uma série temporal de velocidades de vento. Porém só recentemente que, pesquisas relacionadas a combinação de modelos de filtragem e decomposição, começaram a ser conduzidas. A grande motivação pela hibridização na etapa de decomposição se dá pela variabilidade das séries de velocidades de vento, associada a

maneira como a série é amostrada, principalmente em aplicações voltadas à geração de energia eólica, onde o intervalo mínimo de cada amostra é de 10 min, existindo desta forma, dificuldade em obter modelos capazes de realizar previsões *multi-step* consistentes. Cabe porém ressaltar, que a decomposição de séries temporais não é uma tarefa fácil, uma vez que não se conhece a influência de cada componente na série observada, por isso, normalmente não se tem garantia de total separabilidade das componentes, como será discutido mais adiante.

A técnica mais comumente encontrada na literatura, empregada na decomposição de séries temporais, é a análise espectral (RIOS; MELLO, 2012). Dentre os métodos que caracterizam o grupo de análise espectral, pode-se citar: análise de componentes principais (JOLLIFFE, 2002), análise espectral singular (BROOMHEAD; KING, 1986; GOLYANDINA et al., 2001), transformada de Fourier (COOLEY; TUKEY, 1965), transformada de ondeletas (MALLAT, 1989), decomposição de modo empírico (HUANG et al., 1998) e decomposição de modo variacional (DRAGOMIRETSKIY; ZOSSO, 2014). Atualmente as técnicas de análise espectral singular e decomposição de modo empírico tiveram um crescente aumento em popularidade, com aplicações nas mais diversas áreas do conhecimento, como será apresentado ao longo deste trabalho.

A análise de componentes principais (do inglês, *principal components analysis*, PCA) é uma técnica não paramétrica utilizada normalmente para reduzir a dimensionalidade de conjunto de dados, ou extrair informações relevantes de dados correlacionados, através da análise dos autovetores. Utilizando uma definição mais direta do mecanismo da PCA, pode-se dizer que a técnica busca descorrelacionar dados para facilitar a interpretação dos mesmos (JOLLIFFE, 2002). Desta última definição fica claro que só faz sentido aplicar PCA em dados correlacionados.

A técnica de análise espectral singular, ou simplesmente SSA, é relativamente nova, porém tornou-se popular na análise e previsão de séries temporais, principalmente em áreas relacionadas a biomedicina (SANEI; HASSANI, 2015). É uma técnica não paramétrica, desenvolvida para modelar séries temporais que podem ser curtas, não lineares e/ou não estacionárias, bem como ruidosas. O algoritmo SSA não depende da definição de funções a priori para modelar a série temporal, o mesmo gera um conjunto de componentes diretamente da série em estudo.

A principal ideia da SSA é fazer uma decomposição da série temporal original em uma soma de componentes independentes, através da aplicação de decomposição em valores singulares (do inglês, *singular value decomposition*, SVD). As componentes são

agrupadas de forma que representem a tendência, o comportamento oscilatório (componentes periódicos ou quase periódicos) e o ruído. Algumas dessas componentes são selecionadas e, em seguida, usadas para fazer a reconstrução da série filtrada, ou ainda pode-se extrair apenas a componente de tendência. O algoritmo SSA pode ainda ser aplicado para a etapa de previsões de séries temporais (GOLYANDINA et al., 2001).

Dentre os trabalhos mais recentes, relacionados à utilização da técnica SSA na forma híbrida e aplicados na previsão de séries temporais, podemos citar o modelo proposto por Ma et al. (2017) para previsão de velocidades de vento no curto-prazo, o mesmo aplica rede neural nebulosa dinâmica generalizada (GDFNN) integrada ao SSA, cujos parâmetros são otimizados pelo algoritmo de otimização baseado em tempestade cerebral (BSO).

Zhang et al. (2017) aplicou SSA combinada a SVM e CSA para previsão de carga em um sistema de potência; SSA combinada a modelos de volatilidade estocástica (do inglês, *stochastic volatility models*, SVM) foi aplicada por Arteche e García-Enríquez (2017) para estimar a volatilidade de séries no mercado financeiro. Um outro modelo que aplica SSA combinada com alisamento exponencial foi aplicado na previsão de séries de velocidades de vento por Papailias e Thomakos (2017); O modelo proposto por Moreno e Coelho (2018) combina a técnica SSA, utilizada na etapa de decomposição, e um sistema nebuloso de inferência neuro adaptativo para a etapa de previsão. Este modelo pode ser aplicado na previsão de velocidades de vento no curto-prazo, com bons resultados. Mi et al. (2019) propos um modelo híbrido que combina SSA com a técnica EMD para a filtragem da série temporal, e máquinas convolucionais de vetor suporte (CNNSVM) para a previsão *multi-step* de velocidades de vento, entre outros trabalhos.

No Brasil, somente a partir de 2013 é que a técnica SSA começou a ser explorada e aplicada em previsões de séries temporais. Porém, ainda é limitado o número de pesquisas que exploram tal técnica. O trabalho pioneiro no Brasil com a técnica SSA foi realizado por Junior (2013), que propôs em sua tese de doutorado, um modelo híbrido através da combinação de métodos preditivos que agrega cinco abordagens distintas e genéricas, do ponto de vista de modelagem: SSA, decomposição de ondeletas, redes neurais artificiais, programação matemática multiobjetivo *minimax*, com abordagem de programação por metas e método de simulação de quase Monte Carlo. O desempenho do modelo híbrido, em termos de modelagem, foi consideravelmente superior, em relação à todas as estatísticas de aderência consideradas pelo autor.

SSA também foi aplicada por Lopes (2014), para extrair as componentes de

tendências e modos oscilatórios que compõem o sinal da série de precipitação e vazão na Bacia do Rio São Francisco. Nesse trabalho o autor buscava detectar mudança hidroclimáticas na bacia hidrográfica. Cassiano (2014) em sua tese de doutorado, aplicou SSA combinada ao método de clusterização baseado em densidade (do inglês, *densityb based spatial clustering of applications with noise*, DBSCAN), para a previsão de energia eólica, obtendo resultados superiores a outras técnicas. Menezes (2014) aplicou SSA e modelos periódicos auto regressivos (do inglês, *periodic auto regressive*, PAR(q)) para gerar séries sintéticas de velocidades de vento, com o objetivo de auxiliar a tomada de decisão no planejamento da operação energética do sistema elétrico brasileiro. Outro modelo híbrido proposto por Vargas (2014), combinou a densidade condicional por *kernel* e SSA, aplicando-o à obtenção de previsão de geração eólica cujo resultados foram superiores aos modelos auto regressivos.

Apenas três publicações nacionais ao longo de 2016 deram enfoque a aplicação da SSA. Dantas (2016) aplicou SSA para a identificação, através de decomposição de sinais de consumo de energia elétrica, com objetivo de inferir o consumo dos dispositivos que compõem um sinal de consumo de energia elétrica. A abordagem proposta por Ayala (2016), utiliza SSA para obter a previsão da curva de rendimentos através da estrutura da taxa de juros da Colômbia, obtendo resultado significativo quando comparado com técnicas clássicas. Um modelo para previsão da inflação no Brasil utilizando SSA foi proposto por Matsuoka (2016). Nessa abordagem foi aplicado de forma inédita restrições de parâmetros na SSA e os resultados das previsões se mostraram superiores ao de outras técnicas empregadas pelo mesmo autor.

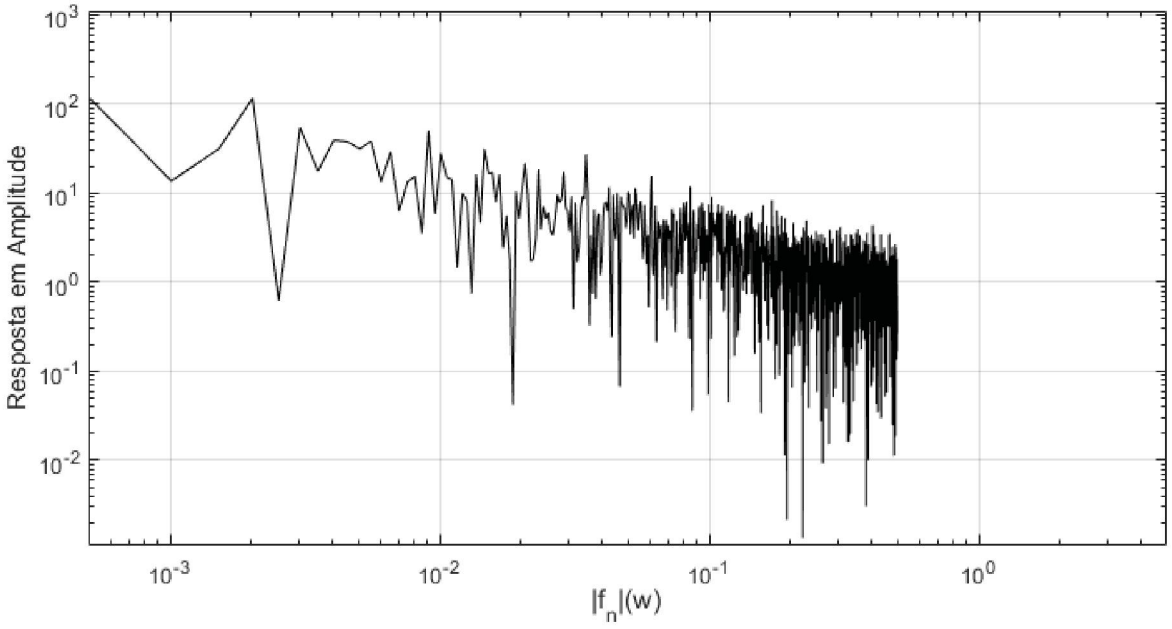
Já em 2017, o trabalho de doutorado publicado por Santos (2017), utiliza a técnica SSA de forma híbrida com redes bayesianas dinâmicas (do inglês, *dynamic bayesian neural networks*, DBNN) e PSO para a identificação de sistemas neurais, obtendo resultados promissores ao remover a não estacionariedade do sinal através da SSA, reduzindo desta forma o espaço de busca de DBNNs. E por último, o trabalho mais recente de autores brasileiros foi publicado no Periódico *Renewable Energy* no final de 2018, cujo modelo proposto combina SSA e ANFIS. O mesmo foi aplicado para obter a previsão *multi-step* de velocidades de vento para um parque éolica situado no Rio Grande do Norte (MORENO; COELHO, 2018).

Outro método de decomposição espectral citado anteriormente, que tem larga aplicação no estudo de sinais, é a transformada de Fourier (do inglês, *Fourier transform*, FT). Historicamente, a análise espectral de Fourier forneceu um método geral para exa-

minar as distribuições globais de energia-frequência, que basicamente, consiste em representar dados através de somas de senos e cossenos. Apesar da popularidade da aplicação de FT em séries temporais, a mesma só é válida sob condições gerais extremas, desta forma há restrições que devem ser observadas antes da aplicação de FT, tais como: o sistema deve ser linear; os dados devem ser estritamente periódicos ou estacionários; caso contrário, o espectro resultante pode não fazer muito sentido físico, ou de difícil interpretação como exemplo o espectro para o sinal original da série de velocidades de vento ou para as 5 componentes obtida pela técnica SSA e aplicação da Transformada de Fourier, representado nas Figuras 3 e 4, respectivamente, onde não é possível determinar qual a frequência dominante para cada componente e nem para o sinal original da série de velocidades do vento (TITCHMARSH, 1948 apud HUANG et al., 1998).

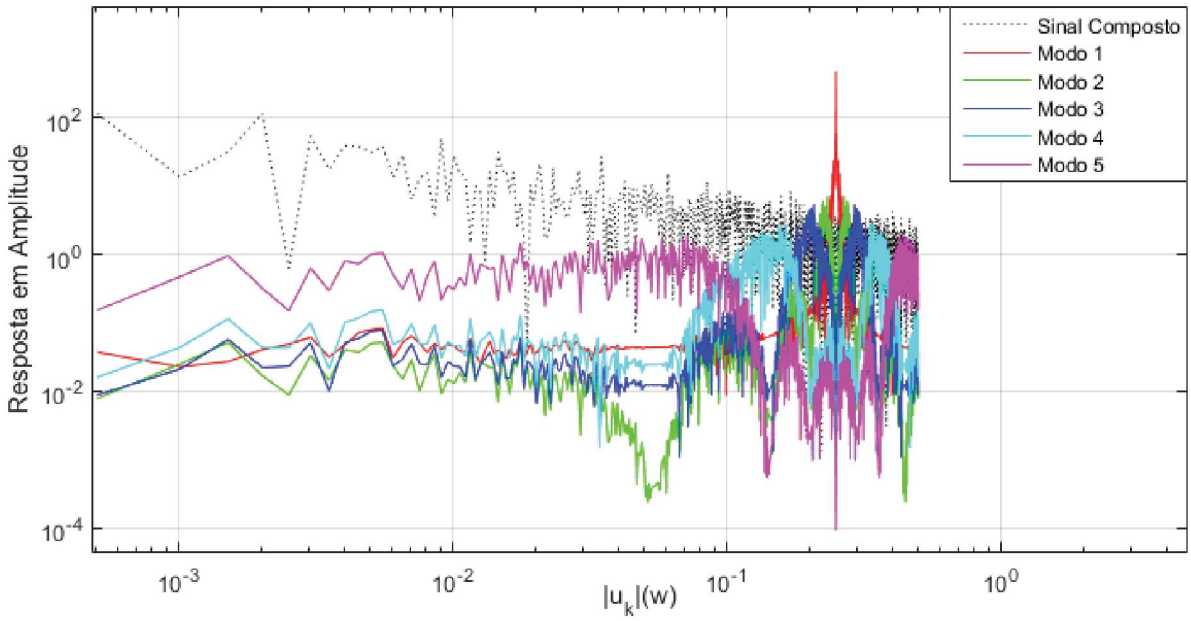
A transformada de ondeletas (WT), também citada como uma das técnicas populares dentro do grupo de análise espectral, consiste essencialmente em aplicar uma janela ajustável a FT. Esta abordagem é muito útil na análise de dados que apresentam mudança gradual de frequência. Apesar de algumas limitações, esta técnica se tornou bastante popular na análise de séries temporais. Uma das principais limitações é a sua natureza não adaptativa. Uma vez que a ondeleta básica é selecionada, a mesma terá que ser aplicada para analisar todos os dados (HUANG et al., 1998).

Figura 3: Espectro da série de velocidades de vento amostradas a cada 10min.



Fonte: O autor (2019).

Figura 4: Espectro de 5 componentes da série de velocidades de vento amostradas a cada 10min.



Fonte: O autor (2019).

O método de decomposição de modo empírico (EMD), última técnica do rol citado, foi proposto em 1998 por Huang et al. (1998). É uma técnica poderosa para separar as respostas transitórias de sistemas não lineares e não estacionários, em conjuntos finitos de componentes quase ortogonais, chamados de funções de modo intrínseco (do inglês, *intrinsic mode functions*, IMF), que representam a característica dinâmica em diferentes escalas de tempo. No entanto, uma deficiência da EMD é a mistura de dois ou mais componentes em um único IMF, o que pode afetar drasticamente o significado físico dos resultados da série decomposta (MOORE et al., 2018). Isso ocorre devido, principalmente, o algoritmo de EMD ser dependente de método recursivo e de métodos de interpolação para obter os limites superiores e inferiores do envelope (que é formado pelos máximos e mínimos locais da série temporal, respectivamente), bem como para a estimação do valor médio desse envelope, que ao ser subtraído da série original, o resultante é chamado de IMF (HUANG et al., 1998).

Devido as características intrínsecas do algoritmo EMD, é muito comum encontrar na literatura trabalhos que utilizam tal técnica na forma híbrida, ou seja, EMD é aplicada na etapa de filtragem e separação dos sinais, enquanto outras técnicas se encarregam de fazer a previsão. Como exemplos recentes podemos citar o trabalho de Jiang e Huang (2017), que aplicou EMD combinada a dois métodos de seleção de características, incluindo a divergência de *Kullback-Leibler* baseada em estimativa da densidade do núcleo

e a medida de energia, para reduzir o distúrbio de componentes mal filtradas, incorporou ainda LSSVM para a etapa de previsão de um passo a frente, e um modelo híbrido utilizando LSSVM e GARCH para corrigir a componente de erro resultante. Segundo reportaram os autores, os resultados de previsão obtidos foram satisfatórios quando comparados a outras técnicas. Qiu et al. (2017) aplicou EMD combinada a rede de crenças profundas (DBN) para efetuar previsões de carga no mercado de energia elétrica. Já Wang e Wang (2017) combinou EMD com *stochastic time dtrength neural network* (STSNN) para efetuar previsões de ações no mercado financeiro, entre outras milhares de aplicações.

Nesta tese é considerado um método alternativo de decomposição, chamado de decomposição de modo variacional (VMD), proposto em 2014 por Dragomiretskiy e Zosso (2014). Este método de decomposição permite extrair facilmente os modos de frequências semelhantes presente em uma série temporal, garantindo a quase separabilidade total dos modos (quase, pois é dependente do número de modos que se deseja separar, e na prática não se tem garantido a separação exata de sinais observados na vida real), o que não ocorre com a EMD, por exemplo.

Diferentemente da SSA, que tem fundamentação na análise de autovalores da SVD, a VMD está fundamentada no domínio de frequência. O algoritmo VMD aplica a transformada de Fourier na versão rápida (do inglês, *fast fourier transform*, FFT), em um modelo variacional não recursivo, para extrair k-modos limitados em banda (do inglês *band-limited intrinsic mode functions*, BLIMFs) simultaneamente. Em particular, o modelo VMD procura uma série de modos e suas respectivas frequências centrais, de forma que, os modos limitados em banda (BLIMFs) reproduzam de forma exata (ou quase exata no sentido de mínimos quadrados) o sinal de entrada.

A hibridização proposta consiste em obter um algoritmo multiestágio de decomposição de uma série temporal, que será realizada ao combinar a técnica VMD e a SSA. No segundo estágio de decomposição, será aplicado a técnica SSA para fazer uma nova decomposição dos modos extraídos, resultantes da decomposição no domínio da frequência, pela VMD. A escolha inicial da SSA é justificada por a mesma estar fundamentada na SVD, proporcionando desta forma, características peculiares a SSA, diferindo-a da decomposição espectral nos seguintes aspectos:

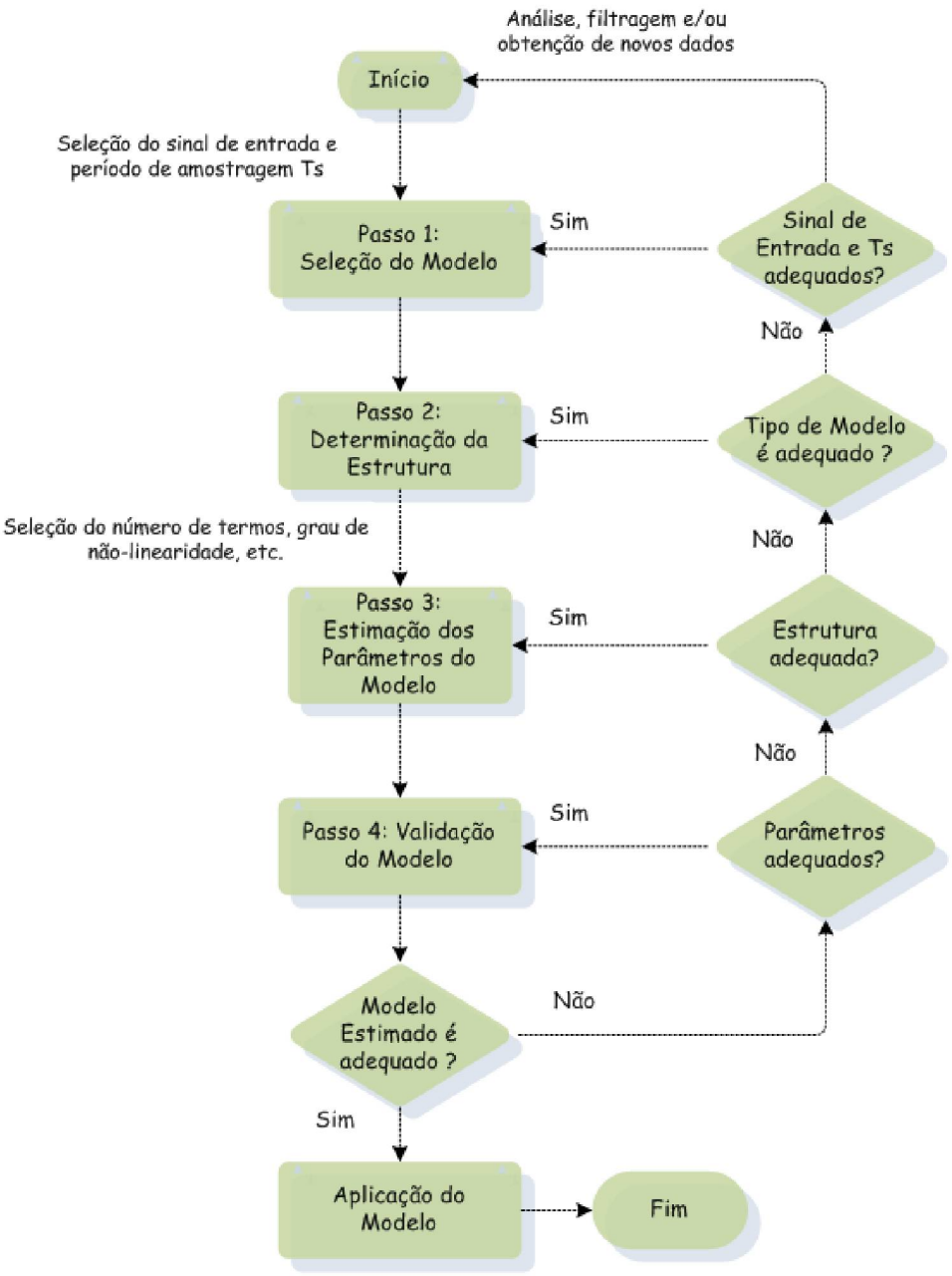
- ✓ Decomposição espectral baseia-se em apenas uma base de autovetores, enquanto SVD em duas diferentes bases: uma à direita (colunas de $V \in \mathbb{R}^n$) e outra à esquerda (colunas de $U \in \mathbb{R}^n$);

- ✓ SVD sempre utiliza bases de vetores ortonormais, enquanto que na decomposição espectral, nem sempre a base de vetores gerada é sequer ortogonal;
- ✓ Somente matrizes quadradas possuem decomposição espectral, ao contrário da SVD, que pode ser aplicada à qualquer matriz, inclusive retangular.

Pelo fato da SSA ser aplicada no segundo estágio de decomposição, espera-se conseguir garantir maior separabilidade das componentes, eliminando a necessidade de decomposição da série original em um elevado número de componentes, como ocorre por exemplo, quando se aplicada somente a técnica SSA na decomposição de uma série temporal (GOLYANDINA et al., 2001), no processo de identificação do comportamento da fonte geradora da série em questão. Normalmente quando se trata de identificação de sistemas, o processo é dividido em etapas como apresentado na Figura 5 (SANTOS, 2007). Porém quando se aplica técnicas orientadas a dados (*data-driven*) como o VMD e SSA, que não necessitam de modelo analítico para descrever o comportamento do sistema, o detalhamento das etapas iniciais do fluxograma são neglicenciadas, sendo despendido mais tempo na etapa de validação onde se utiliza medidas de erro para avaliar o desempenho da técnica.

Desta forma, a abordagem adotada utilizará indicadores normalmente aplicados para avaliar a qualidade do modelo, tais como: erro médio quadrático (do inglês, *mean squared error*, MSE); raiz do erro médio quadrático (*root mean squared error*, RMSE); erro médio absoluto (do inglês, *mean absolute error*, MAE); erro médio percentual absoluto (do inglês, *mean absolute percentage error*, MAPE) e raiz do erro médio quadrático relativo (do inglês, *relative root mean squared error*, RRMSE) para avaliar a acurácia das previsões.

Figura 5: Etapas na identificação de sistemas



Fonte: O autor(2019).

Cabe ressaltar ainda que não foi encontrado nenhuma pesquisa relativa a aplicação híbrida das técnicas VMD e SSA na literatura nacional ou internacional, quer seja em teses de doutorado ou artigos científicos. Destaque ainda para a escassez de pesquisas utilizando a combinação das duas técnicas, pois até o momento da realização deste trabalho, nenhuma publicação foi encontrada relativa ao assunto.

1.2 OBJETIVOS

Nesta seção é apresentado os objetivos da pesquisa, os mesmos são detalhados na forma de objetivo geral e específico, considerando que já foi postulada a hipótese à ser trabalhada, de forma à atingir tais objetivos, na seção anterior. Temas como decomposição de séries temporais, análise espectral, previsão de séries de velocidade de vento, inteligência computacional e modelos híbridos serão abordadas ao longo da tese.

1.2.1 OBJETIVO GERAL

O objetivo desta tese consiste em explorar a hibridização das técnicas de decomposição VMD e SSA, de tal forma que as componentes resultantes da série de velocidades de vento não apresentem correlação significativa, resultando em previsões com maior acurácia em termos de erro médio percentual absoluto (MAPE).

1.2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos são enumerados a seguir:

1. Desenvolver uma técnica de decomposição hibridizada, multiestágio através da aplicação de conceitos (ou das técnicas) de VMD e SSA, para formar a etapa de decomposição do modelo;

2. Fornecer através do modelo híbrido, componentes aditivas melhores separadas para a etapa de previsão da série temporal;

3. Avaliar acurácia do modelo em separar as componentes que constituem a série temporal, através da medida de correlação entre cada componente;

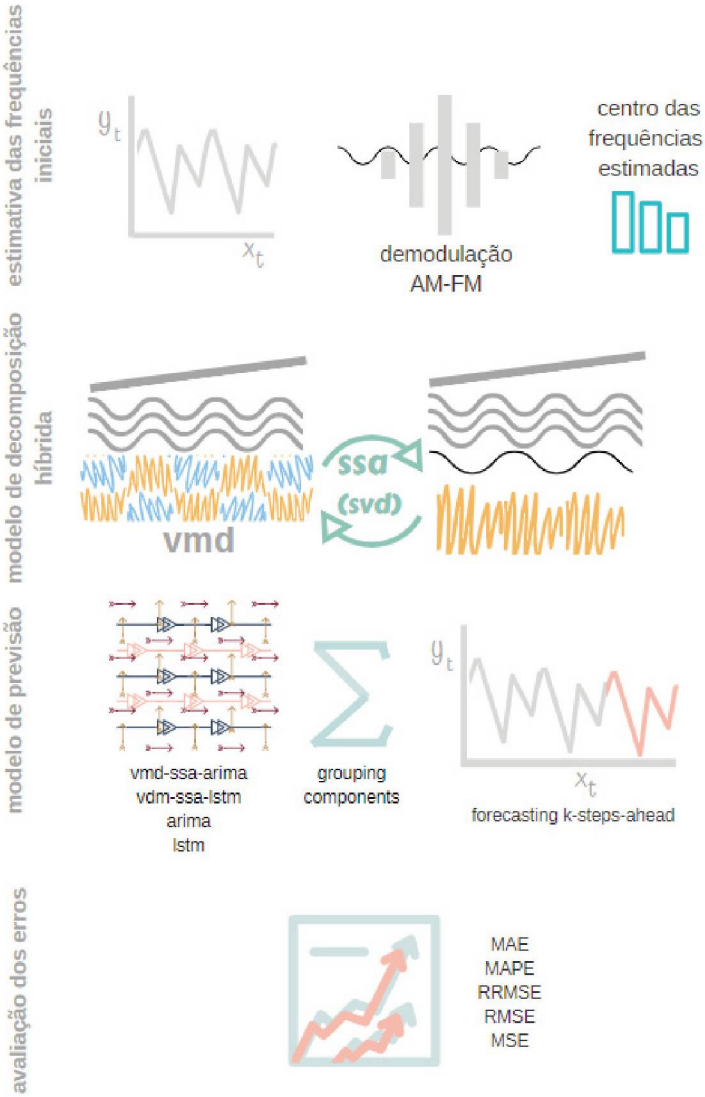
4. Aplicar de forma individualizada (em cada componente), modelos de previsão paramétricos (ARIMA) e não paramétricos (LSTM) para validar a hipótese da tese;

5. Avaliar acurácia das previsões resultantes, através de testes ou indicadores de qualidade (MSE, MAPE, RMSE, MAE, RRMSE);

6. Aplicar o modelo híbrido de previsão a um parque eólico em operação, localizado no Nordeste Brasileiro, avaliar o desempenho da técnica através da comparação das previsões *versus* realizado (ou observado), para um horizonte de 1 até k-passos-a-frente.

A Figura 6 apresenta a estrutura do modelo híbrido proposto nesta tese, composto pelas etapas de estimativa das frequências iniciais e pela etapa de hibridização do VMD-SSA. Já as etapas de previsão e de análise dos erros, incorporam o modelo para validar a hipótese formulada nesta tese.

Figura 6: Etapas do modelo proposto nesta tese



Fonte: O autor(2019).

1.3 ORGANIZAÇÃO DA TESE

Esta tese está organizada em 7 capítulos e 6 apêndices, conforme especificado a seguir:

- **Capítulo 1**: Introdução;

- **Capítulo 2**: Fundamentação Teórica;
- **Capítulo 3**: Modelo Híbrido Proposto;
- **Capítulo 4**: Aplicação do Modelo VMD-SSA;
- **Capítulo 5**: Avaliação do Desempenho do Modelo Híbrido de Decomposição;
- **Capítulo 6**: Avaliação do Desempenho do Modelo Híbrido de Previsão;
- **Capítulo 7**: Conclusão e Pesquisa Futura;
- **Apêndice A**: Estratégias de Previsão;
- **Apêndice B**: Modelagem Box & Jenkins;
- **Apêndice C**: Relação de Recorrência Linear - LRR;
- **Apêndice D**: Redes Neurais LSTM;
- **Apêndice E**: Otimizador Adam;
- **Apêndice F**: Publicações e Premiação.

O Capítulo 2 está divido em quatro seções, sendo as duas iniciais dedicadas exclusivamente a detalhar cada uma das técnicas utilizadas na hibridização, e as duas restantes apresentam os conceitos das estratégias de previsão e dos índices de desempenho. Na seção que trata da técnica SSA, é apresentado o formalismo matemático de cada etapa que constitui a técnica, porém sem ofuscar a fundamentação conceitual e a aplicação prática da técnica. Para tal, no final da seção é apresentado um exemplo de aplicação da técnica na decomposição de uma série temporal. Também é apresentado os conceitos normalmente utilizados para avaliar a qualidade dos resultados. A seção finaliza fazendo uma síntese, onde é destacado os principais aspectos da técnica, sejam eles positivos ou negativos.

Ainda no Capítulo 2, a seção que trata sobre a técnica VMD não apresenta um detalhe matemático tão elaborado como na SSA, mas apresenta os conceitos de forma que seja possível ao leitor associar e assimilar a fundamentação da técnica, que é baseada estritamente nos conceitos de mistura de sinais e deslocamento de sinais no domínio da frequência. A seção também apresenta o conceito de coerência espectral, muito útil para identificar frequências semelhantes entre dois sinais. É apresentado um exemplo de aplicação da VMD no final da seção, no exemplo é obtido a decomposição de um sinal

que contém duas frequências conhecidas a *priori*, a partir dessas componentes é analisado a coerência espectral entre ambos os modos extraídos e o sinal original. Tal como a seção anterior, a mesma finaliza fazendo uma síntese dos aspectos mais relevantes que envolvem a técnica VMD.

O Capítulo 3 aborda o modelo híbrido proposto VMD-SSA, detalhando cada etapa do algoritmo, além de apresentar a estrutura do modelo VMD-SSA para aplicação na etapa de decomposição de uma série temporal em várias componentes, e posterior aplicação na previsão de séries temporais. Também é apresentado o algoritmo do modelo desenvolvido resgatando a fundamentação dos conceitos de demodulação de sinais, útil para entendimento do modelo proposto.

No Capitulo 4 é apresentado o conjunto de dados a ser utilizado na aplicação do modelo. Discute-se o ajuste dos parâmetros das técnicas de decomposição SSA, VMD e das técnicas de previsão ARIMA e LSTM. O capítulo ainda apresenta o conjunto de modelos de previsão ARIMA(p,d,q) obtidos através do ajuste dos parâmetros (p,d,q) para cada uma das componentes; e na última seção do capítulo é apresentado os parâmetros de ajuste da rede neural LSTM.

O Capítulo 5 traz os resultados da decomposição obtidos com a aplicação do modelo híbrido proposto. Os resultados da decomposição são explorados e as correlações entre às componentes resultantes da decomposição analisadas.

O Capítulo 6 apresenta os resultados das previsões obtidas com o acoplamento à técnica híbrida de decomposição proposta (VMD-SSA). As técnicas de previsão utilizadas são em números de duas, cada uma pertencente à classe distinta, para avaliar o benefício do modelo híbrido. A primeira é a técnica ARIMA, pertencente à classe dos modelos paramétricos, e a última, uma rede neural LSTM, que pertence a classe dos modelos não paramétricos. Os resultados obtidos são comparados as técnicas ARIMA e LSTM, quando aplicadas de forma independentes.

O Capítulo 7 apresenta as conclusões e principais perspectivas, apontando as principais vantagens da hibridização proposta nesta tese para contextos específicos, além das possíveis extensões da técnica à outros problemas.

No Apêndice A é apresentado as 5 estratégias de previsões para k-passos-à-frente, detalhando as singularidades de cada estratégia, além de trazer representação gráfica que busca facilitar o entendimento do leitor com relação ao mecanismo básico e a identificação das principais diferenças entre as estratégias de previsão.

Na sequência, o Apêndice B apresenta o fundamento da modelagem de Box & Jenkins. Este tópico é de extrema relevância dado que, os modelos ARIMA(p,d,q) são uma classe de modelos paramétricos com larga aplicação na previsão de séries temporais.

O Apêndice C apresenta o fundamento de uma técnica de previsão normalmente incorporado a técnica SSA, chamado de relação de recorrência linear. Apesar da mesma não ser utilizada nesta tese, este tópico é importante para destacar o impacto da escolha dos parâmetros da técnica SSA.

No Apêndice D é apresentado os conceitos da rede neural LSTM. É apresentado com um pouco mais de detalhamento as etapas que constituem a célula de memória de longo e curto prazo, principal característica da rede neural LSTM.

O otimizador Adam é apresentado no Apêndice E, onde é dado ênfase aos conceitos e ao equacionamento matemático, além da representação do algoritmo do mesmo.

O Apêndice F contém uma lista com as publicações realizadas pelo autor desta tese durante a execução da pesquisa, também faz menção a premiação recebida por melhor artigo publicado em congresso.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação teórica dos dois métodos de decomposição aplicados nesta pesquisa. Inicialmente é apresentado a técnica SSA, com detalhamento do formalismo matemático de cada etapa que constitui a técnica. É discutido a fundamentação conceitual, além da apresentação da aplicação prática da técnica no final da seção. Também é apresentado os conceitos normalmente utilizados para avaliar a qualidade dos resultados, tais como a correlação ponderada e correlação cruzada. A seção finaliza com uma síntese sobre os aspectos da técnica.

Na seção seguinte é apresentado a fundamentação teórica da técnica VMD, que é baseada estritamente nos conceitos de mistura de sinais e deslocamento de sinais no domínio da frequência. A seção também apresenta o conceito de coerência espectral, muito útil para identificar frequências semelhantes entre dois sinais. É apresentado um exemplo de aplicação da VMD no final da seção. No exemplo é obtido a decomposição de um sinal que contém duas frequências conhecidas a *priori*, a partir destas componentes é analisado a coerência espectral entre ambos os modos extraídos e o sinal original. Tal como a seção anterior, a mesma finaliza fazendo uma síntese dos aspectos mais relevantes que envolvem a técnica VMD.

2.1 ANÁLISE ESPECTRAL SINGULAR

A análise espectral singular (do inglês, *singular spectrum analysis*, SSA) é uma técnica de análise e previsão de séries temporais. SSA visa a decomposição da série original em uma soma de pequeno números de componentes interpretáveis, tais como: a componente de tendência que varia lentamente, componentes oscilatórios e um "ruído de estrutura menor" (ELSNER; TSONIS, 1996). A técnica é fundamentada na decomposição de valores singulares (SVD) de uma matriz de trajetórias obtida através da série temporal. Não há necessidade de modelos paramétricos ou pressupostos de estacionariedade, ou ainda que os ruídos sejam descorrelacionados. Essas características permitem que a

SSA tenha uma ampla gama de aplicabilidade (GOLYANDINA et al., 2001; ELSNER; TSONIS, 1996; SANEI; HASSANI, 2015).

O marco inicial da técnica SSA foi em 1986 por meio do trabalho de Broomhead e King (1986). Porém somente em 2001, com a publicação do livro *Analysis of Time Series Structure* por Golyandina et al. (2001) é que a técnica se popularizou. A técnica SSA consiste na aplicação de quatro passos, que são enumerados e definidos a seguir, mantendo a nomenclatura em inglês: *embedding step, svd step, groupping step* e *diagonal averaging step* representados na Figura 7 e detalhados a seguir:

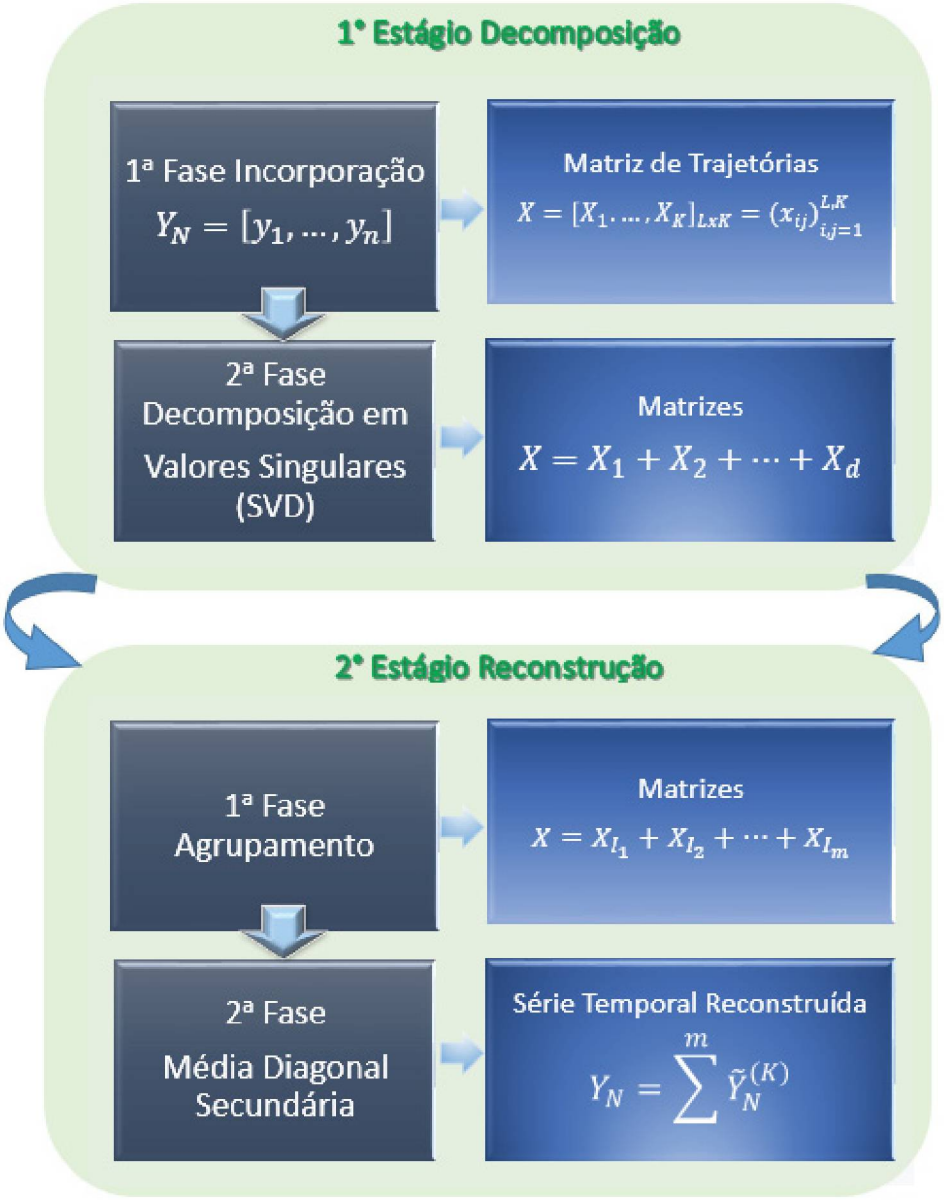
- ✓ ***Embedding Step*** **- Incorporação**: a série temporal unidimensional é representada como uma série multidimensional cuja dimensão é chamada de comprimento da janela L. As séries temporais multidimensionais formam a matriz de trajetória e o único parâmetro desta etapa é L. Normalmente L tem valor restrito a $L < \frac{N}{2}$, onde N é o comprimento da série temporal;

- ✓ ***SVD Step*** **- Decomposição**: é a decomposição do valor N singular da matriz de trajetórias em uma soma das matrizes bi-ortogonais de *rank* unitário;

- ✓ ***Grouping Step*** **- Agrupamento**: corresponde a dividir as matrizes calculadas no *SVD step* em vários grupos, e então realizar a soma das matrizes dentro de cada grupo. O resultado do passo é uma representação da matriz de trajetórias como uma soma de várias matrizes resultantes;

- ✓ ***Diagonal Averaging Step*** **- Média Diagonal Secundária**: é uma operação linear que mapeia a matriz de trajetórias formada pela série inicial para a própria série inicial. Desta forma, obtemos uma decomposição da série inicial em várias componentes aditivas.

Figura 7: Estágios da SSA



Fonte: O autor(2019).

2.1.1 ESTÁGIO DE DECOMPOSIÇÃO

Os dois primeiros passos (*embedding* e *svd*) formam o estágio de decomposição da série temporal, e os dois últimos passos, formam o estágio de reconstrução no algoritmo SSA. O poder da técnica SSA baseia-se na análise das componentes extraídas na etapa de decomposição. Por exemplo, se não estivermos interessados em certos aspectos das componentes da série temporal, nós os somamos junto as componentes de ruído. No entanto, deve ser observada a restrição de que isso só pode ser feito se as componentes forem fortemente separáveis. O procedimento matemático para realizar as duas etapas é

descrito abaixo, considerando a série temporal $Y_N = [y_1, \dots y_N] \in \mathbb{R}^N$ de comprimento N. A explicação detalhada completa pode ser encontrada em Golyandina et al. (2001), faz-se referência a ela para maiores detalhes dos fundamentos matemáticos.

Dada a série temporal Y_N com cardinalidade igual a N e $F\colon \mathbb{R}^N \to \mathbb{R}^{L\times K}$ um mapa invertível, define-se incorporação como um procedimento onde dado um valor L para o comprimento da janela, a série temporal $Y_N \in \mathbb{R}^N$ é transformada através do mapa F, em uma matriz $\mathbf{X} = [X_1, \dots, X_K]_{L\times K} \in \mathbb{R}^{L\times K}$, de tal forma que $X_j = [y_j, \dots, y_{j+L-1}]^T \in \mathbb{R}^L, j = 1, \dots, K$; e T denota a operação de transposição. Ou seja, $Y_N \in \mathbb{R}^N \xrightarrow{F} \mathbf{X} \in \mathbb{R}^{L\times K}$, onde $K = N - L + 1$. O resultado desta fase é a matriz de trajetórias conforme Equação 2.1, que tem característica de ser uma matriz *Hankel*, ou seja, todos os elementos da diagonal $i + j = constante$ são iguais. Os vetores X_j, com $j = 1, \dots, K$ na equação são chamados de vetores $L_{defasados}$.

$$\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_K]_{L\times K} = (x_{ij})_{i,j=1}^{L,K} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_N \end{bmatrix} \tag{2.1}$$

Após a obtenção da matriz de trajetórias $\mathbf{X}$, a mesma é decomposta em uma soma de matrizes elementares bi-ortogonais de posto 1. Cabe aqui uma observação importante, apesar da aplicação ser a base para a técnica SSA, a matriz de trajetórias pode ser utilizada para unificar vários procedimentos de análise de séries temporais, tais como filtragem e modelagem auto regressiva (SANEI; HASSANI, 2015). Para ilustrar tal afirmação, considere β representando uma matriz linha de quantidades desconhecidas, se fixarmos a dimensão do vetor $L \times 1$, pode-se obter entre outras, o seguinte exemplo:

- Para $L = 2$ e $\beta = [-1, 1]$ pode-se obter as primeiras diferenças das realizações, ao fazer $\beta\mathbf{X}$;
- Para todo $L \geq 2$ e $\beta = [\frac{1}{L}, \frac{1}{L}, \dots, \frac{1}{L}]$ pode-se obter a média móvel de ordem L para as realizações ao fazer $\beta\mathbf{X}$.

Feito esse breve comentário, retoma-se a definição, onde deve-se tomar $S = \mathbf{X}\mathbf{X}^T$, de forma que S seja um operador normal, compacto e positivo semi-definido, ou seja, tomando $\Psi(S)$ como o *espectro* de S e $\{U_l U_l^T\}_l^L = 1$ uma resolução de identidade sobre o espaço de *Hilbert* $(\mathbb{R}^L, <\ .\ >)$ associada ao operador S, onde U_l é o autovetor associado

ao autovalor $\lambda_l \in \Psi(S)$, de modo que $\lambda_l \geq 0 \ \forall \ l$. Definindo $d = max(i, \text{ tal que } \lambda_i > 0) = \text{rank } \mathbf{X}$. Se denotarmos $V_i = \dfrac{\mathbf{X}^T U_i}{\sqrt{\lambda_i}}$, a matriz $\mathbf{X}$ pode ser expandida através de decomposição em valores singulares (SVD). Como esta condição suficiente é verdadeira, tem-se a decomposição SVD conforme Equação 2.2

$$\mathbf{X} = \sum_{i=1}^{d} \mathbf{X}_i = \mathbf{X}_1 + \ldots + \mathbf{X}_d \tag{2.2}$$

onde $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$, $(i = 1, \ldots, d)$. Vale lembrar que as matrizes $\mathbf{X}_i$ possuem *rank* 1. Outra observação importante é que a decomposição SVD da Equação 2.2 é ótima no sentido de que entre todas as matrizes $\mathbf{X}^{(r)}$ com *rank* $r < d$, a matriz $\mathbf{X}^{(r)} = \sum_{i=1}^{r} X_i$ resulta na melhor aproximação da matriz de trajetórias $\mathbf{X}$, de forma que $\| \mathbf{X} - \mathbf{X}^{(r)} \|$ é mínima.

As matrizes $\mathbf{X}_i$ são matrizes elementares, já os vetores U_i e V_i são identificados como vetores singulares da matriz $\mathbf{X}$. Outras nomenclaturas mais específicas podem ser encontradas, onde nomeia-se U_i como vetor singular à esquerda (*left sigular vector*), ou autovetor principal, ou ainda, EOF que é a abreviação de *factor empirical orthogonal functions*. Enquanto que V_i, frequentemente chamado de componentes principais, ou autovetores do lado direito da matriz de trajetória (*right singular vector*). A tripla $(\sqrt{\lambda_i}, U_i, V_i)$ é chamada de *i-th eigentriple* (ou auto-triplas) da matriz $\mathbf{X}$; $\sqrt{\lambda_i}$ são os valores singulares; e o conjunto de $\{\sqrt{\lambda_i}\}_{i=1}^{d}$ são frequentemente chamados de *spectrum* da matriz $\mathbf{X}$.

2.1.2 ESTÁGIO DE RECONSTRUÇÃO

O primeiro passo do estágio de reconstrução inicia com o agrupamento (*grouping step*), que corresponde a dividir a matriz elementar em vários grupos e realizar a soma das matrizes dentro desses grupos. Para facilitar o entendimento, considere $I = \{i_i, \ldots, i_p\}$ como um grupo de índices $i = 1, \ldots, p$; então a matriz $\mathbf{X}_I$ correspondente ao grupo I é definida como $\mathbf{X}_I = \mathbf{X}_{i1} + \ldots + \mathbf{X}_{ip}$. A divisão do conjunto de índices $\{1, \ldots, L\}$ em subconjuntos disjuntos $I_1, \ldots, I_m$ corresponde a representação $\mathbf{X} = \mathbf{X}_{i1} + \ldots + \mathbf{X}_{im}$. O procedimento de escolha dos conjuntos $I_1, \ldots, I_m$ é chamado de agrupamento.

Para um dado grupo I, a contribuição da componente $\mathbf{X}_I$ considera que as contribuições acumuladas transmitidas pelas primeiras r matrizes elementares, equivale a fração de explicabilidade associada à aproximação ótima da matriz trajetória pela matriz

de posto r representada na Equação 2.3

$$contr_i = \frac{\sum_{i \in I}^{r} \lambda_i}{\sum_{i=1}^{d} \lambda_i} \tag{2.3}$$

As etapas de SVD e agrupamento apoiam-se na propriedade denominada separabilidade, que caracteriza o quão bem os diferentes componentes podem ser separados uns dos outros em termos de dependência linear. A decomposição e reconstrução da série original terá êxito se os seus componentes aditivos forem separáveis um do outro, sendo este um dos pontos que pode tornar a técnica de SSA ineficiente. A ausência de separabilidade aproximada é frequentemente observada em série com estrutura complexa, como por exemplo a série de velocidades de vento. Recomenda-se que após ter selecionado os autovalores correspondentes às componentes de ruído e sinal (para o exemplo de estar decompondo a série em duas componentes), avaliar a efetividade da separação através da correlação ponderada (ω correlação será discutida mais adiante), que é uma função que quantifica a dependência linear entre duas componentes.

Na etapa de reconstrução da série, o último passo consiste em converter as matrizes $\mathbf{X}_{i1}, \ldots, \mathbf{X}_{im}$ em novas séries de comprimento N, através do processo de média diagonal ou *hankelização* (GOLYANDINA et al., 2001), para que possam ser convertidas novamente em uma série temporal. Para maior entendimento, considere a matriz $\mathbf{Y}$ de dimensão $L \times K$, com elementos y_{ij}, onde $1 \leq i \leq L$ e $1 \leq j \leq K$. Tomando $L^* = min(L, K)$, $K^* = max(L, K)$ e $N = L + K - 1$, finalmente fazendo $y^*_{ij} = y_{ij}$ se $L < K$, caso contrário $y^*_{ij} = y_{ji}$. A média diagonal transforma a matriz $\mathbf{Y}$ de dimensão $L \times K$ em uma série temporal $y_1, \ldots, y_N$ através da Equação 2.4

$$y_k = \begin{cases} \dfrac{1}{k} \displaystyle\sum_{m=1}^{k} y^*_{m,k-m+1} & \text{para } 0 \leq k < L^*, \\ \dfrac{1}{L^*} \displaystyle\sum_{m=1}^{L^*} y^*_{m,k-m+1} & \text{para } L^* \leq k < K^*, \\ \dfrac{1}{N-k+1} \displaystyle\sum_{m=k-K^*+1}^{N-K^*+1} y^*_{m,k-m+1} & \text{para } K^* \leq k < N \end{cases} \tag{2.4}$$

A Equação 2.4 aplicada a matriz $\mathbf{X}_{Ik}$ produz a série $\tilde{X}^{(k)} = (\tilde{x}_1^{(k)}, \ldots, \tilde{x}_N^{(k)})$, isto é equivalente a decomposição da série inicial $x_1, \ldots, x_N$ em uma soma de m séries reconstruídas, como apresentado na Equação 2.5. O Algoritmo 2 apresenta os passos da SSA.

$$x_n = \sum_{k=1}^{m} \tilde{x}_N^{(k)} \qquad (n = 1, 2, \ldots, N) \tag{2.5}$$

Algoritmo 1 Pseudocódigo SSA

Entrada: Y_N {Série temporal de tamanho N}
Saída: $\tilde{Y}_N^{(K)}$ {Série temporal filtrada}

Parâmetros:
L {Comprimento da janela}
I {Conjunto de componentes a serem reagrupadas}
$\mathbf{X}$ {Matriz de trajetórias (Hankel)}
{Embedding Step: primeiro passo da SSA}
{Verificar a restrição do tamanho máximo da janela}
se $L > \frac{N}{2}$ **então**
    Limita a janela em $L = N - L$
fim se
$K = N - L + 1$ {Dimensão da matriz de trajetórias}
para $i = 1$ **até** K **faça**
    Obter a matriz Hankel $X(1 : L, i) = x(i : L + i - 1)$, conforme Eq. 2.1
fim para
{SVD Step: segundo passo da SSA}
$\mathbf{S} = \mathbf{X}\mathbf{X}^T$
$[\mathbf{U}, autoval] = eig(\mathbf{S})$ {Autovetores à esquerda}
$[d, i] = sort(-diag(autoval))$ {Ordena em ordem crescente os autovalores e obtem seus indíces}
$\mathbf{U} = \mathbf{U}(:, i)$ {Ordena as colunas de forma que os autovetores estejam corretamente associados a ordem dos autovalores}
$\mathbf{V} = \mathbf{X}^T\mathbf{U}$ {Autovetores à direita}
{Grouping Step: terceiro passo da SSA}
Faz a transposta $\mathbf{V}_t = \mathbf{V}^T$
$\mathbf{rca} = \mathbf{U}(:, I)\mathbf{V}_t$
{Reconstruction Step: quarto passo da SSA}
$Lp = min(L, K)$
$Kp = max(L, K)$
para $k = 0$ **até** $Lp - 2$ **faça**
    para $m = 1$ **até** $k + 1$ **faça**
        Obter a matriz Hankel $\tilde{x}(k + 1) = \tilde{x}(k + 1) + (\frac{1}{k+1}) \times rca(m, k - m + 2)$
    fim para
fim para
para $k = Lp - 1$ **até** $Kp - 1$ **faça**
    para $m = 1$ **até** Lp **faça**
        Obter a matriz Hankel $\tilde{x}(k + 1) = \tilde{x}(k + 1) + (\frac{1}{Lp}) \times rca(m, k - m + 2)$
    fim para
fim para
para $k = Kp$ **até** N **faça**
    para $m = k - Kp + 2$ **até** $N - Kp + 1$ **faça**
        Obter a matriz Hankel $\tilde{x}(k + 1) = \tilde{x}(k + 1) + (\frac{1}{N-k}) \times rca(m, k - m + 2)$
    fim para
fim para
$\tilde{Y}_N^K = \sum_{k=1}^{m} \tilde{x}_k^m \qquad (k = 1, \ldots, N)$

Fonte: O autor(2019).

2.1.3 CORRELAÇÃO PONDERADA

Ao analisarmos as propriedades da técnica SSA, separabilidade é o que caracteriza quão bem as diferentes componentes podem ser separadas uma das outras. Desta forma, a decomposição de uma série temporal Y_N, através da técnica SSA, só será bem sucedida se as componentes aditivas resultantes da série são aproximadamente separadas uma das outras.

Uma maneira de medir a separabilidade, ou a dependência entre duas séries Y_N^1 e Y_N^2, é através da aplicação da correlação ponderada (*weighted correlation* ou ω-*correlation*), que é dada pela Equação 2.6 (ESQUIVEL, 2012),

$$\rho_{12}^{\omega} = \frac{\left\langle Y_N^1, Y_N^2 \right\rangle_{\omega}}{\|Y_N^1\|_{\omega}\|Y_N^2\|_{\omega}} \tag{2.6}$$

onde $\|Y_N^i\|_{\omega} = \sqrt{\left\langle Y_N^i, Y_N^i \right\rangle_{\omega}}$; $\left\langle Y_N^i, Y_N^j \right\rangle_{\omega} = \sum_{k=1}^{N} \omega_k y_k^i y_k^j$ com $(i, j = 1, 2)$ e assumindo que $L \leq \frac{N}{2}$ de tal forma que $\omega_k = min\{k, L, N-k\}$.

Se o valor absoluto das ω-correlações for pequeno, então o valor correspondente das componentes da série são quase ω-ortogonais, mas, se for grande, então as componentes estão longe de ser ω-ortogonal e, portanto, não estão bem separadas. Em linhas gerais, é necessário que a correlação ponderada seja nula ou que esteja próxima de zero, indicando que as componentes não estão altamente correlacionadas e não formam um mesmo grupo de auto-triplas. Então, se dois componentes reconstruídos possuem correlação zero, isso significa que esses dois componentes estão totalmente separados. Grandes valores de correlação entre componentes reconstruídos indicam que os componentes devem ser reunidos em mesmo um grupo.

2.1.4 CORRELAÇÃO CRUZADA

Correlação cruzada é uma medida do grau de similaridade entre duas séries temporais X_1 e Y_1, ou conjunto de dados, em função do deslocamento de uma em relação a outra. A correlação cruzada também é conhecida como produto interno de janela deslizante, a mesma é definida pela Equação 2.7 (CHATEFIELD, 2000),

$$r_{xy} = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} X_1(t)Y_1(t+\tau)dt \tag{2.7}$$

onde r_{xy} é a correlação cruzada entre X_1 e Y_1; τ é o avanço de tempo. Uma maneira direta de estimar a função de correlação cruzada é utilizar a definição de covariança cruzada (*cross covariance*) c_{xy}, que é definida da seguinte forma: considere n pares de valores correspondente as séries temporais X_1 e Y_1, ou seja, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ disponíveis para análise. Pode ser mostrado, conforme proposto em (JENKINS; WATTS, 1968), que uma estimativa do coeficiente da covariança cruzada c_{xy} no *lag* k é dada pela Equação 2.8, tal que

$$c_{xy} = \begin{cases} \frac{1}{n}\sum_{t=1}^{n-k}(x_t - \bar{x})(y_{t+k} - \bar{y}), & k = 0, 1, 2, \ldots, n \\ \frac{1}{n}\sum_{t=1}^{n+k}(y_t - \bar{y})(x_{t-k} - \bar{x}), & k = 0, -1, -2, \ldots, -n \end{cases} \tag{2.8}$$

onde $\bar{x}, \bar{y}$ são as médias amostrais das séries temporais X_1, Y_1, respectivamente. Similarmente, a estimativa do coeficiente da correlação cruzada r_{xy} para o *lag* k é dada pela Equação 2.9

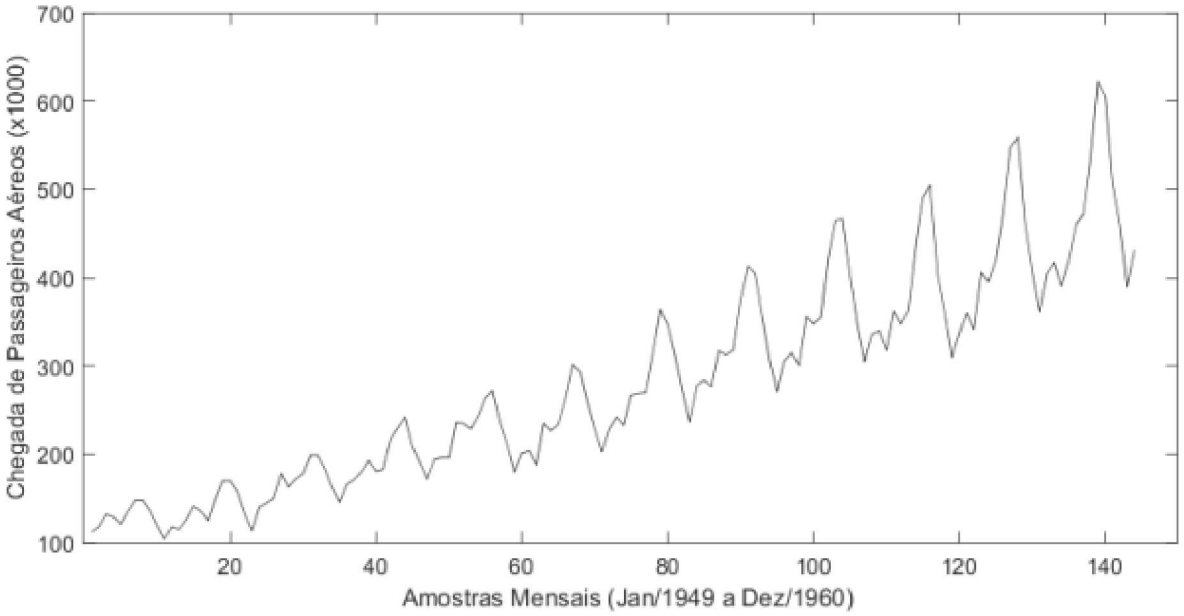
$$r_{xy}(k) = \frac{c_{xy}(k)}{\sqrt{c_{xx}(0)}\sqrt{c_{yy}(0)}} = \frac{c_{xy}(k)}{s_x s_y} \quad k = 0, \pm 1, \pm 2, \ldots, \pm n \tag{2.9}$$

onde s_x, s_y representam o desvio padrão de x e y.

2.1.5 EXEMPLO DE APLICAÇÃO SSA

Para ilustrar a técnica SSA, nesta seção é apresentado um exemplo da aplicação na decomposição e reconstrução de uma série temporal de dados históricos que representam a chegada de passageiros em aeroportos nos USA no período de Jan/1949 à Dez/1960 (BOX et al., 1976). A série é apresentada na Figura 8, e possui $N = 144$ dados amostrais. Pode-se observar claramente um comportamento sazonal e de tendência da série.

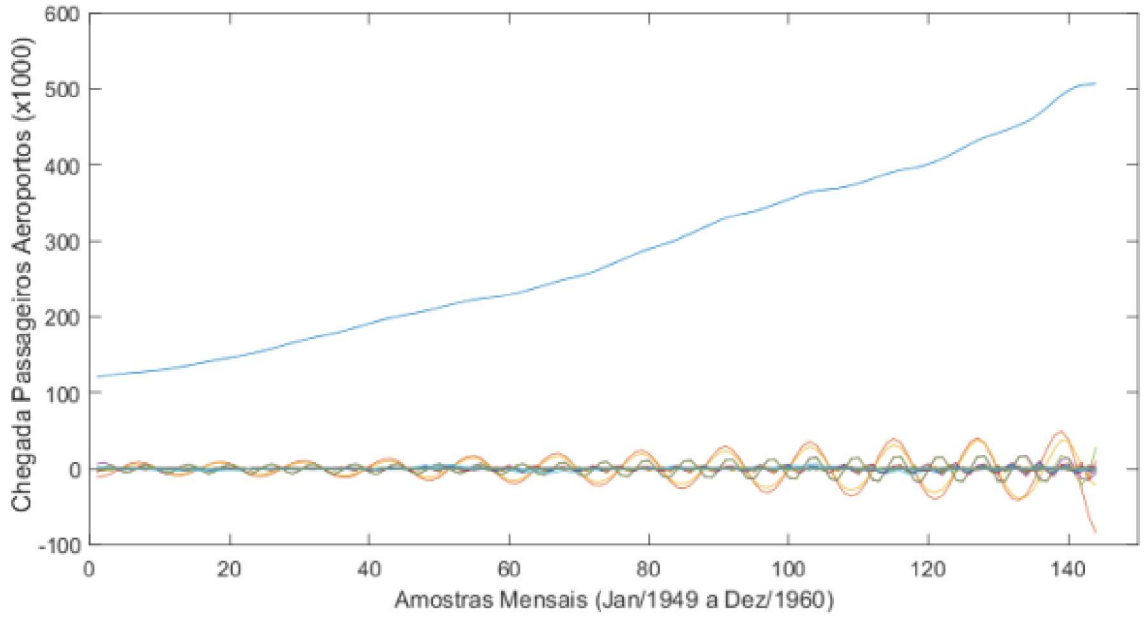
Figura 8: Série chegada de passageiros em aeroportos nos USA



Fonte: O autor(2019) - adpatado de (BOX et al., 1976).

Ao aplicar as duas primeiras fases do primeiro estágio da técnica SSA, que são responsáveis pela decomposição da série, com uma janela $L = 20$, de forma que a matriz de trajetórias tenha a dimensão LxK, onde $K = N - L + 1 = 144 - 20 + 1 = 125$, dessa forma a dimensão da matriz *Hankel* será 20x125. Esta escolha foi arbitrária, apenas para exemplificar a aplicação da técnica. Após a decomposição, obteve-se as componentes apresentadas na Figura 9.

Figura 9: Decomposição da série com $L = 20$

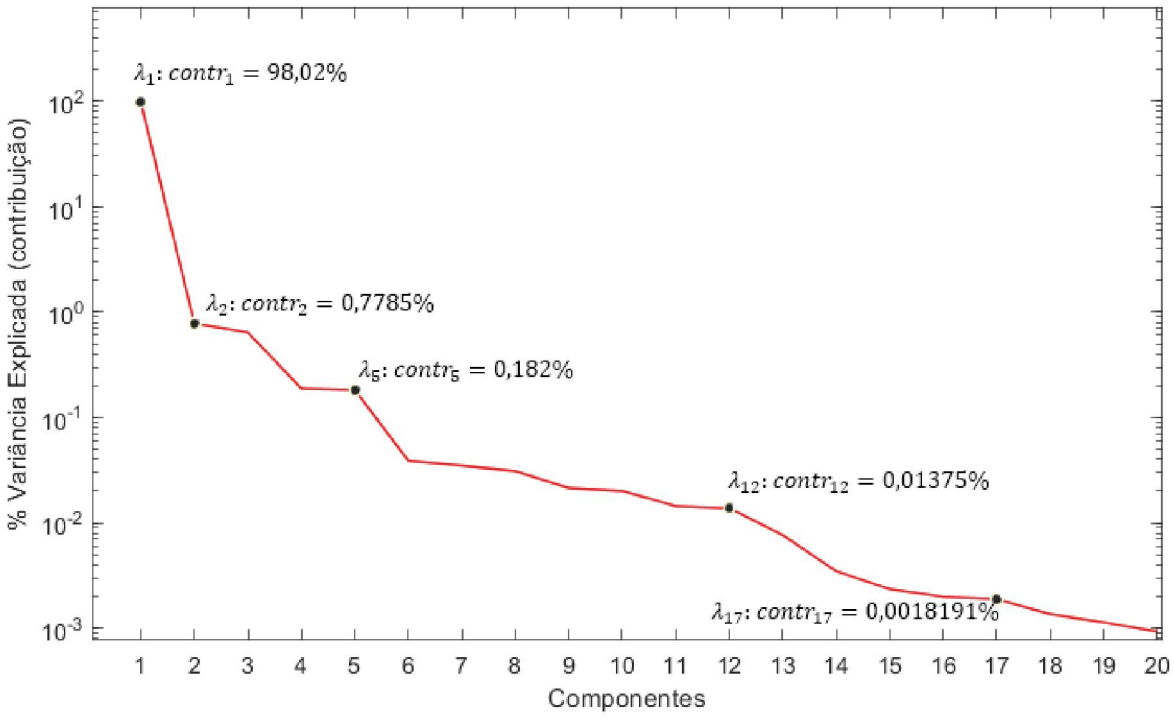


Fonte: O autor(2019).

O algoritmo da contribuição percentual de cada componente na explicação da variância da série temporal é apresentada na Figura 10. Nesta representação, as componentes mais significativas para a reconstrução do sinal, aparecem com valores singulares elevados. Porém, segundo Golyandina et al. (2001), como regra, uma série de ruído puro produz uma sequência decrescente de valores singular. Se tal ruído é adicionado a um sinal, descrito por algumas ou um conjunto de auto-triplas de vetores com valores singulares elevados, então uma quebra no espectro de autovalores pode distinguir o sinal das auto-triplas do ruído. Por outro lado, planos explícitos (uma reta horizontal) nos espectros de autovalores indicam os números ordinais das auto-triplas pareadas, significando que as mesmas podem ser agrupadas.

Pode-se observar na Figura 10 que 0 autovalor λ_1 da primeira componente explica $98,02\%$ da variância, enquanto que soma das contribuições dos demais autovalores $\lambda_2, \lambda_3, \ldots, \lambda_20$ associada a cada componente contribuem com os 2% restantes. Apesar de ser uma informação relevante, não fica claro como as componentes devem ser agrupadas na etapa de reconstrução da série, pois tem-se valor singular elevado para a primeira auto-tripla e valor quase desprezível para as demais, aparentemente apenas a primeira auto-tripla deveria ser utilizada na reconstrução do sinal original.

Figura 10: Contribuição de cada componente na explicação da variância da série

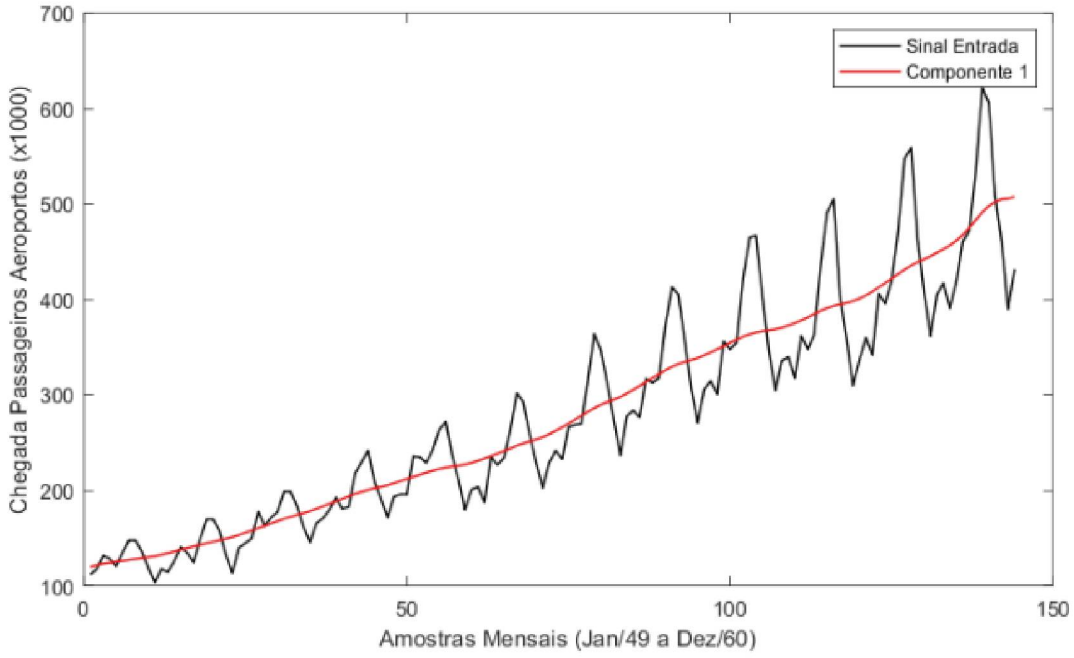


Fonte: O autor(2019).

A próxima etapa consiste em determinar a forma como as componentes serão agrupadas. Arbitrariamente será adotada as 5 primeiras componentes como independentes e as demais, a partir da sexta até a vigésima componente, serão agrupadas em uma única componente denominada de ruído. Esta escolha nos levará ao próximo passo da análise que constitui em checar a separabilidade dessas componentes através da correlação ponderada, também será aplicado a correlação cruzada para ilustrar o grau de separabilidade alcançado pela SSA.

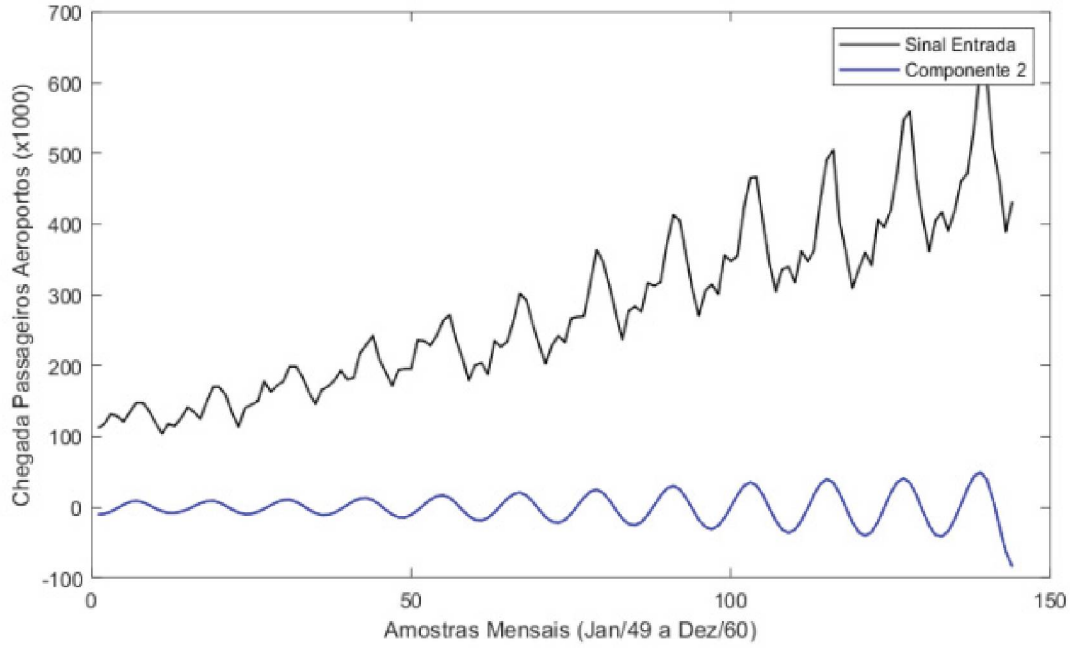
As componentes reconstruídas e agrupadas são apresentadas nas Figuras 11 à 15. Pode-se observar na Figura 11 que a primeira componente representa de forma satisfatória a tendência da série original. Enquanto as demais apresentam um perfil oscilatório de sazonalidade. A questão remanescente é: A separabilidade das componentes é perfeita ? Inspecionando visualmente a Figura 11 à 15 não se pode afirmar com clareza. Porém ao fazer a análise das correlações ponderadas e/ou das correlações cruzadas é possível ter uma idéia mais concreta da separabilidade das componentes.

Figura 11: Primeira componente reconstruída



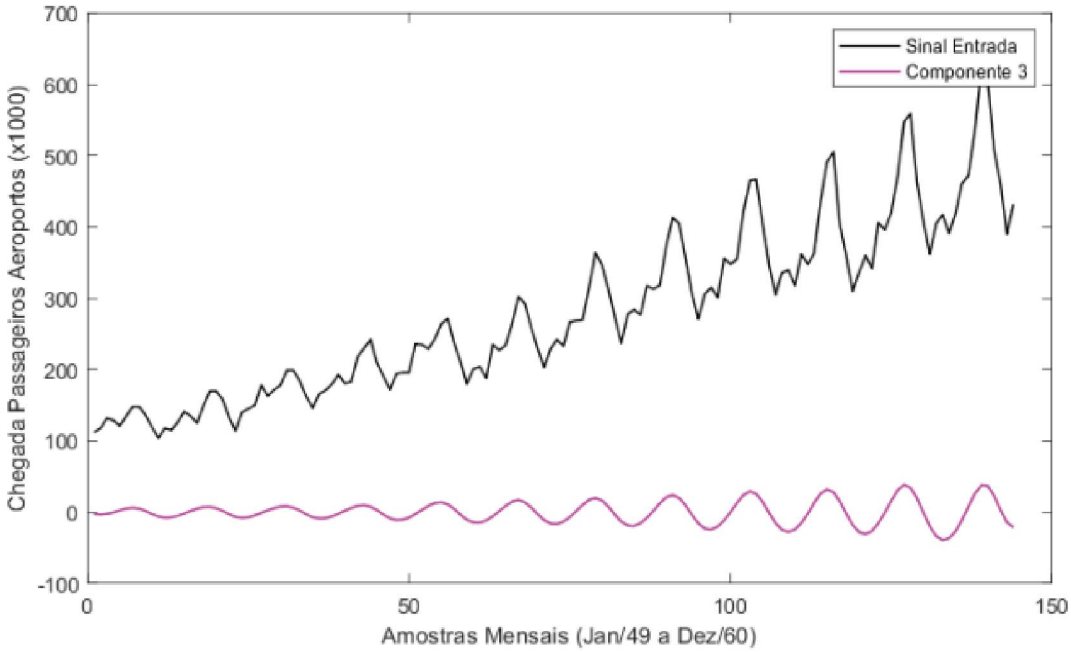
Fonte: O autor(2019).

Figura 12: Segunda componente reconstruída



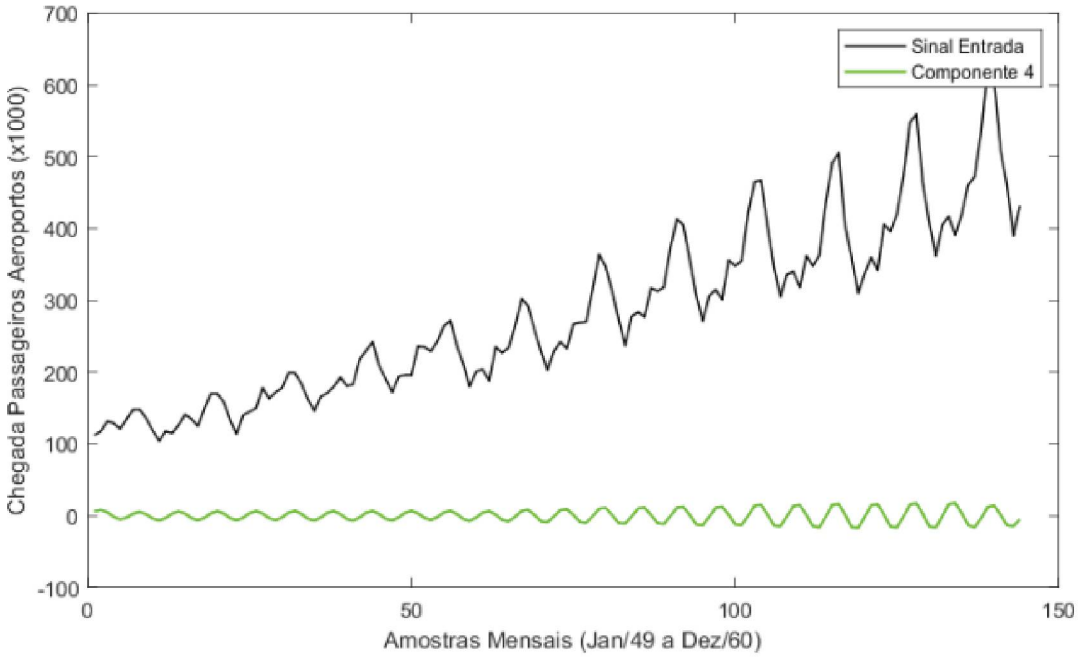
Fonte: O autor(2019).

Figura 13: Terceira componente reconstruída



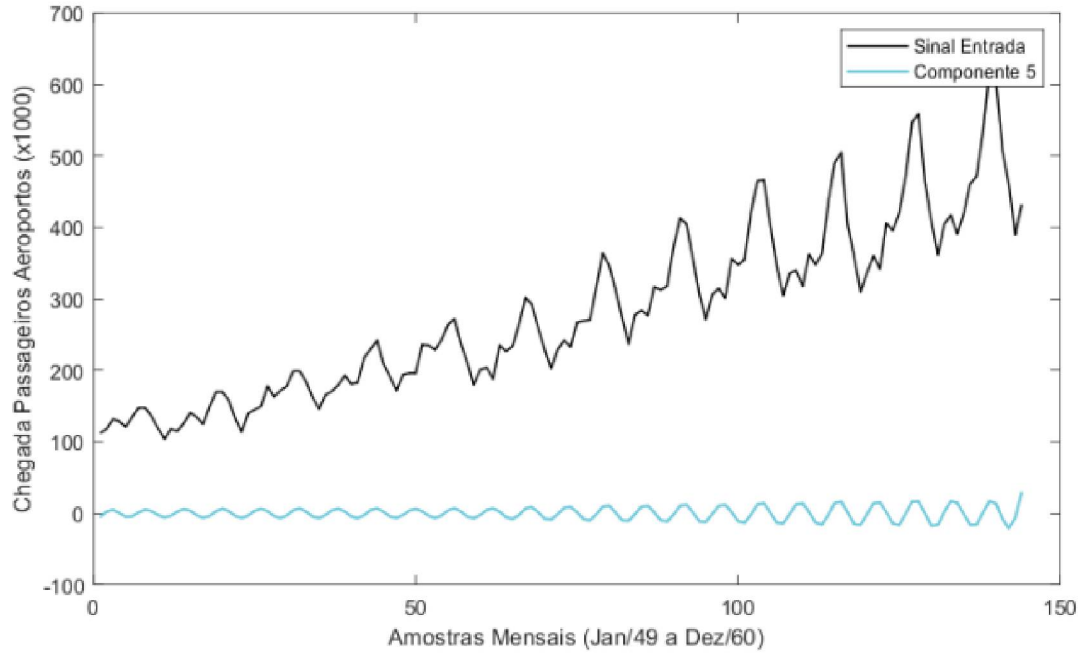
Fonte: O autor(2019).

Figura 14: Quarta componente reconstruída



Fonte: O autor(2019).

Figura 15: Quinta componente reconstruída

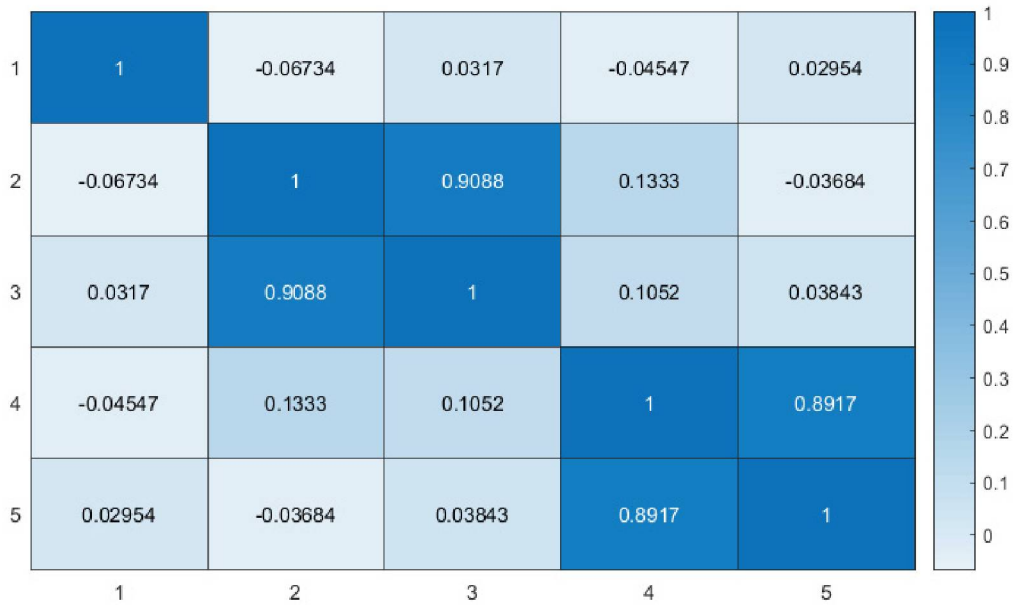


Fonte: O autor(2019).

A Figura 16 apresenta a matriz de correlação ponderada na forma de um *heatmap*. É possível observar que a componente 1 é praticamente ortogonal as demais, ou seja, apresenta correlação próxima de zero. Já as componentes 2 e 3 apresentam correlação

elevada, significando que as mesmas devem ser agrupadas em uma mesma componente. O mesmo se pode dizer sobre as componentes 4 e 5, ou seja, devem ser agrupadas pois compartilham informações altamente correlacionadas.

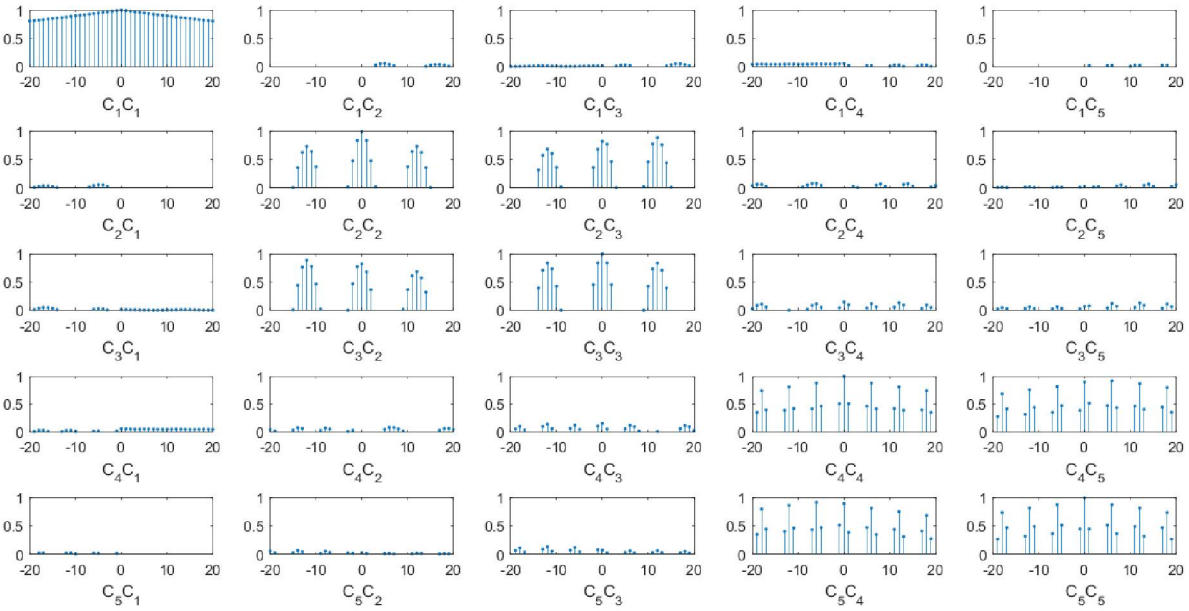
Figura 16: Correlação ponderada das componentes obtidas por meio do SSA



Fonte: O autor(2019).

Ao fazer uso da correlação cruzada, pode-se ter a mesma informação fornecida pela correlação ponderada, porém a similaridade está associada também ao *lag* máximo onde a correlação deixa de existir ou tende a valores muito pequenos, como pode ser visto na Figura 17.

Figura 17: Correlação cruzada das componentes obtidas por meio do SSA



Fonte: O autor(2019).

2.1.6 ASPECTOS A SEREM DESTACADOS

Pode-se observar que a técnica SSA por ser não paramétrica, não necessita de pressupostos básicos tais como normalidade dos resíduos, de estacionariedade da série temporal ou de uma estrutura de modelo predefinida. Esses normalmente são pressupostos necessários na análise de séries temporais quando se utiliza técnicas clássicas, sendo portanto, uma das qualidades da técnica SSA. Porém deve-se destacar que a maior dificuldade na aplicação da técnica é a impossibilidade de se garantir a separabilidade das componentes. O grau de separabilidade decorre da escolha correta do número mínimo de componentes nas quais a série original será decomposta. Outra dificuldade é a definição do número de componentes a ser selecionadas na etapa de agrupamento, de forma que informações do comportamento da série original não sejam perdidos.

2.2 DECOMPOSIÇÃO DE MODO VARIACIONAL

A decomposição de modo variacional (do inglês, *variational mode decomposition*, VMD) é uma técnica adaptativa e não-recursiva aplicada para a decomposição de sinais, a mesma foi proposta por Dragomiretskiy e Zosso (2014). Através da VMD um sinal de entrada é decomposto em um conjunto de modos de pulsação u_k (ou componentes)

limitados em banda, onde cada modo é compacto em torno de uma frequência central de pulsação w_k (entenda-se aqui como largura de banda em torno de uma frequência central), determinada ao longo da decomposição. Esses modos de pulsação ou sub-sinais são definidos na literatura como BLIMFs, acrônimo de *band-limited intrinsic mode functions*. A largura de banda para cada modo u_k pode ser estimada através do seguinte procedimento (DRAGOMIRETSKIY; ZOSSO, 2014):

- ✓ Para cada modo u_k compute o sinal analítico associado por meio da transformada de Hilbert, para obter um espectro de frequência unilateral, ou seja, um espectro que consiste somente de frequências não negativas;
- ✓ Ao misturar com um sinal exponencial sintonizado com a respectiva frequência central estimada, o espectro de frequência de u_k é traduzido para as regiões da banda base;
- ✓ A largura de banda de u_k é estimada através do alisamento Gaussiano do sinal demodulado (por exemplo: quadrado da norma L^2 do gradiente).

Cabe aqui definir melhor o que é a mistura de sinais no domínio da frequência. Mistura é o processo de combinação de dois sinais não-lineares, introduzindo assim termos de frequência cruzada na saída, que para recuperar o sinal original deve-se aplicar um processo de filtragem na frequência desejada. O mais simples misturador é o operador de multiplicação, como por exemplo, a relação trigonométrica $2cos(2\pi f_1 t)cos(2\pi f_2 t) = cos\left[(2\pi(f_1 + f_2)t\right] + cos\left[2\pi(f_1 - f_2)t\right]$ é uma mistura de sinais. Já ao multiplicar um sinal analítico $f_A(t)$ com um exponencial puro $e^{-j\omega_0 t}$ se produz um deslocamento na frequência do sinal, resultando em um sinal constituído por uma única frequência, como por exemplo, o sinal apresentado no domínio de Fourier, através da Equação 2.10

$$f_A(t)e^{-j\omega_0 t} \overset{\mathcal{F}}{\longleftrightarrow} \hat{f}_A(\omega) * \delta(\omega + \omega_0) = \hat{f}_A(\omega + \omega_0) \tag{2.10}$$

onde δ é o delta de *Dirac*; $\omega_0 = 2\pi f_0$ é a frequência do sinal exponencial puro; $j^2 = -1$ é o operador complexo (ou parte imaginária de um número complexo); ω a frequência original do sinal que sofreu o deslocamento; e $*$ é o operador de convolução. Com esse conceito em mente, defini-se o problema variacional restrito que é representado na Equação 2.11, cujo termo $\|\|$ representa a norma L^2, u é o sinal de entrada, u_k denota as componentes obtidas, ω_k é o centro de frequência estimada do sinal e o termo $\left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right]$

transforma $u_k(t)$ em um sinal analítico, resultando em um espectro de frequência unilateral (*single-side*), tal que

$$\begin{aligned} &\underset{\{u_k\},\{\omega_k\}}{\text{minimizar}} \quad \left\{ \sum_{k=1}^{K} \left\| \partial_t \left\{ \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\} \right\|^2 \right\} \\ &\text{sujeito à} \quad \sum_{k=1}^{K} u_k = u \end{aligned} \tag{2.11}$$

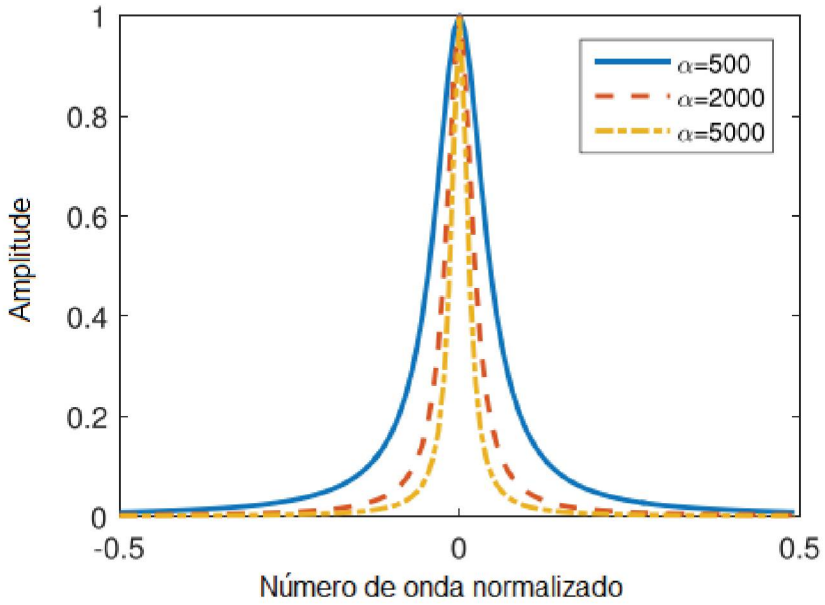
Para resolver este problema de otimização com restrições, é adotado o uso do Lagrangeano aumentado, tornando-o irrestrito. A solução do problema original dado pela Equação 2.11 pode ser obtida como sendo o ponto de sela $L(\{u_k\}, \{w_k\}, \lambda)$, onde λ é o multiplicador de Lagrange. Este ponto de sela é resultante de um processo iterativo, obtido como sendo a solução de uma sequência de sub-problemas. Especificamente para nth iteração, a kth componente do sinal e sua frequência central ω_k podem ser atualizados no domínio Espectral de Fourier. Este problema é apresentado na forma de duas equações relativas a otimização dos modos u_k e dos centros das frequências ω_k, respectivamente através das Equações 2.12 e 2.13,

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{u}(\omega) - \sum_{i<k} \hat{u}_i^{n+1}(\omega) - \sum_{i>k} \hat{u}_i^{n}(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{2.12}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 \, d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 \, d\omega} \tag{2.13}$$

onde $\hat{u}_k, \hat{u}$ e $\hat{\lambda}$ representam a TF para u_k, u e λ, respectivamente. Analisando ainda a Equação 2.12, a mesma pode ser considerada um filtro passa-baixa de *Wiener* centrado em torno da frequência ω_k, com largura de banda definida por α. A largura de banda do filtro será reduzida na proporção que o valor de α aumentar, ou seja, para α assumindo grandes valores a largura de banda do filtro será mínima e mais ruído será filtrado, como representado na Figura 18 (YU; MA, 2018) .

Figura 18: Influência de α no filtro de Wiener



Fonte: O autor(2019) - adaptado de (YU; MA, 2018).

Ao analisar a Equação 2.13, pode-se verificar que os valores iniciais das frequências centrais podem ser obtidos como as frequências de pico do espectro de Fourier, e o número de componentes do sinal (isto é, K) também pode ser estimado de acordo com o número dos picos dominantes no espectro.

Para um completo detalhamento do formalismo matemático que resultou nesta forma fechada das equações, recomenda-se consultar Dragomiretskiy e Zosso (2014). De acordo com os conceitos apresentados, onde se definiu o problema variacional com restrições, podemos representá-lo no Algoritmo 2, que é o VMD na forma irrestrita utilizando as Equações 2.12 e 2.13 para um erro de convergência ϵ.

Algoritmo 2 Otimização Completa do VMD

1: Inicialize: $\hat{u}_k^1, \omega_k^1, \hat{\lambda}^1, n \leftarrow 0$
2: **repita**
3: $n \leftarrow n + 1$
4: **para** $k = 1$ **até** K **faça**
5: Atualize $\hat{u}_k$ para todo $\omega \geq 0$
6: $\hat{u}_k^{n+1} \leftarrow \dfrac{\hat{u}(\omega) - \sum_{i<k} \hat{u}_i^{n+1} - \sum_{i>k} \hat{u}_i^n + \frac{\hat{\lambda}^n}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2}$
7: Atualize ω_k
8: $\omega_k^{n+1} \leftarrow \dfrac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 \, d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 \, d\omega}$
9: **fim para**
10: Aplicar *Dual Ascent* para todo $\omega \geq 0$
11: $\hat{\omega}^{n+1} \leftarrow \hat{\omega}^n + \tau(\hat{f} - \sum_k \hat{u}_k^{n+1})$
12: **até** *Convergência* : $\sum_k \| \hat{u}_k^{n+1} - \hat{u}_k^n \|_2^2 \, / \, \| \hat{u}_k^n \|_2^2 < \epsilon$

Fonte: O autor(2019).

2.2.1 COERÊNCIA ESPECTRAL

O domínio espectral de *Fourier* é muito útil para identificação de similaridade entre dois sinais. Como estamos tratando da separabilidade de sinais, é importante analisar o conteúdo de cada componente obtida através da separação. Desta forma, se as mesmas compartilham alguma frequência em comum significa que o sinal não está totalmente separado.

Para obter a coerência espectral entre dois sinais faz-se necessário obter a TF para cada sinal. Os coeficientes de *Fourier* de uma série temporal são então multiplicados pelo complexo conjugado dos coeficientes de *Fourier* da segunda série temporal, e a sua magnitude conjunta é tomado como a coerência entre os dois sinais. Ou seja, esta operação nada mais é que o produto interno *hermitiano*, que pode ser definido como descrito na Equação 2.14 (COHEN, 2014),

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i = \mathbf{Y}^T \mathbf{X} = \mathbf{Y}^T \mathbf{I}_n \mathbf{X} \qquad \forall (x, y) \in \mathbb{C}^n \tag{2.14}$$

onde $\mathbf{Y}^T$ é a transposta *hermitiana* da matriz coluna $\mathbf{Y}$; $\mathbf{I}_n \in \mathbb{M}_n(\mathbb{R})$ é a identidade de ordem n.

Deixando o formalismo matemático da álgebra linear, o procedimento para obtenção da coerência espectral utilizando a TF no ambiente *Matlab*™ é muito simples, conforme Algoritmo 3. Porém a simples multiplicação dos coeficientes de *Fourier* de dois sinais espectrais não é uma medida apropriada de coerência, devido influência da amplitude da cada série temporal, independentemente da relação real das duas séries temporais entre si (COHEN, 2014). Desta forma, uma medida interpretável de coerência espectral é obtida fazendo a normalização com à amplitude individual dos dois sinais.

Algoritmo 3 Coerência Espectral

1: Entradas: $Sinal_A, Sinal_B, n$
2: Obter a TF do $Sinal_A$ tal que $SigAx = \dfrac{TF(Sinal_A)}{n}$
3: Obter a TF do $Sinal_B$ tal que $SigBx = \dfrac{TF(Sinal_B)}{n}$
4: Realizar a multiplicação elemento-a-elemento dos coeficientes de $SigAx$ pelo complexo conjugado dos coeficientes de $SigBx$
5: $SpecX = abs(SigAx. \cdot conj(SigBx)).^2$
6: Obter a normalização de $SpecX$ pela magnitude dos sinais
7: $SpecXCoe = SpecX./(SigAx. \cdot SigBx)$

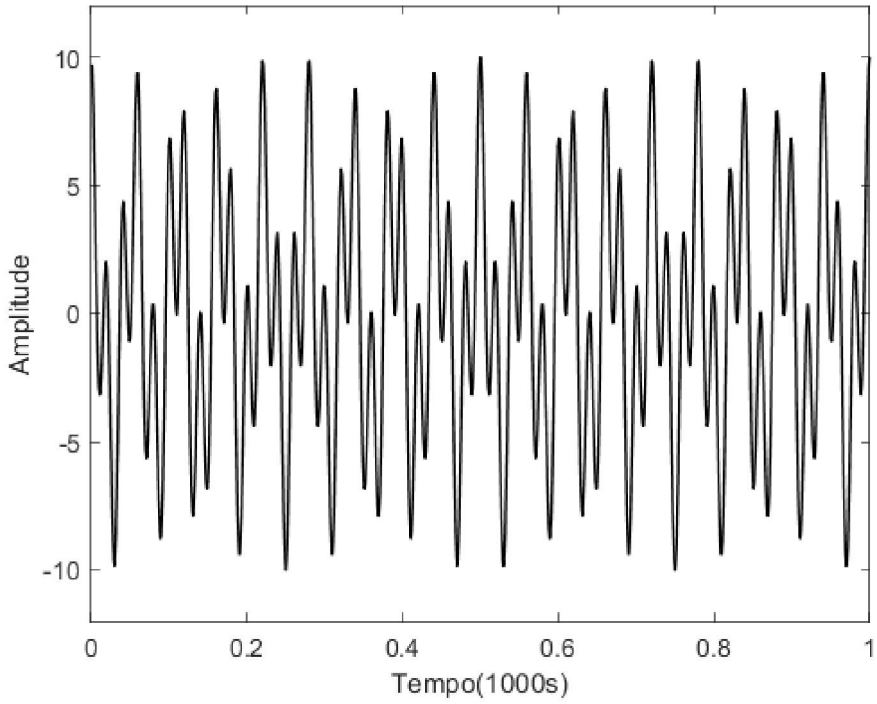
Fonte: O autor(2019).

onde $Sinal_A$, $Sinal_B$ são dois sinais no domínio do tempo com comprimento n; $SigAx$, $SigBx$ é os sinais no espectro de *Fourier*; *abs* é o valor do módulo do sinal complexo; *conj* é a operação de complexo conjugado no domínio dos complexos $\mathbb{C}$; os símbolos ./ e . significam divisão e multiplicação elemento-a-elemento no ambiente *Matlab*™, respectivamente.

2.2.2 EXEMPLO DE APLICAÇÃO VMD

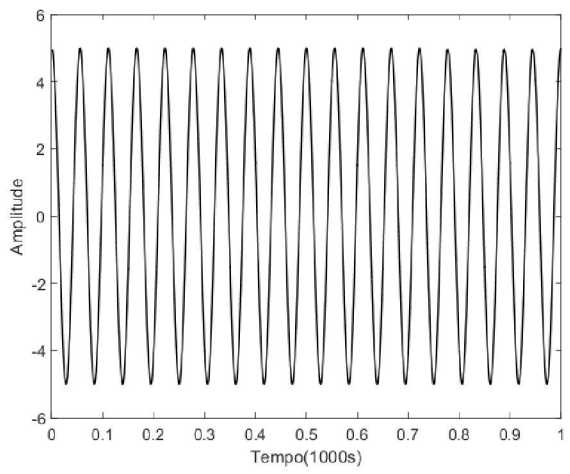
Afim de ilustrar a aplicação do algoritmo VMD na separação de sinais limitados em banda, para uma melhor compreensão dos conceitos apresentados na seção anterior, considere a série temporal da Figura 19, que é composta pela mistura de dois sinais A e B representados na Figura 20, que contém as frequências de 18 e 50Hz, ou seja, o resultante Sinal C contém a mistura de ambas as frequências.

Figura 19: Sinal C - composto por conteúdo nas frequências 18 e 50Hz

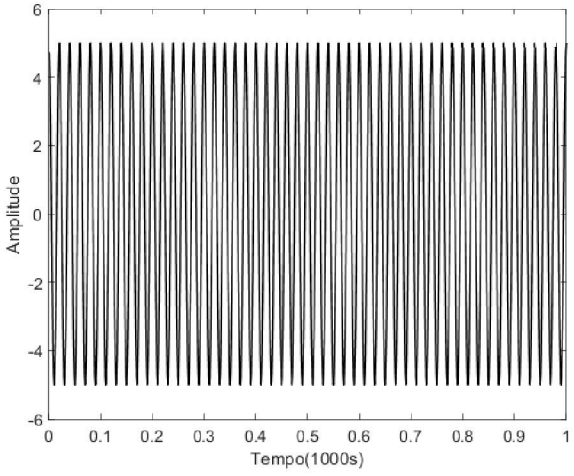


Fonte: O autor(2019).

Figura 20: Sinais misturados que originaram o Sinal C



(a) Sinal A - conteúdo na frequência de 18Hz

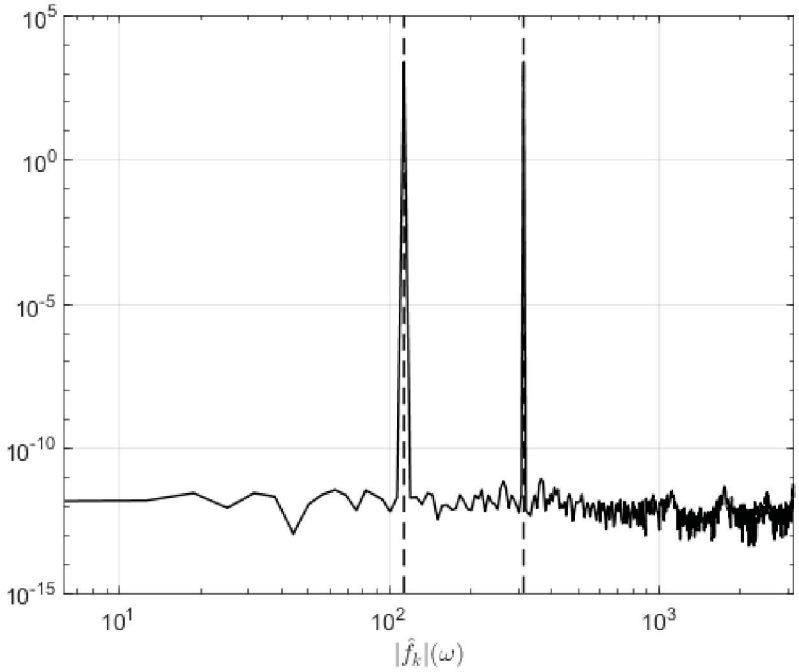


(b) Sinal B - conteúdo na frequência de 50Hz

Fonte: O autor(2019).

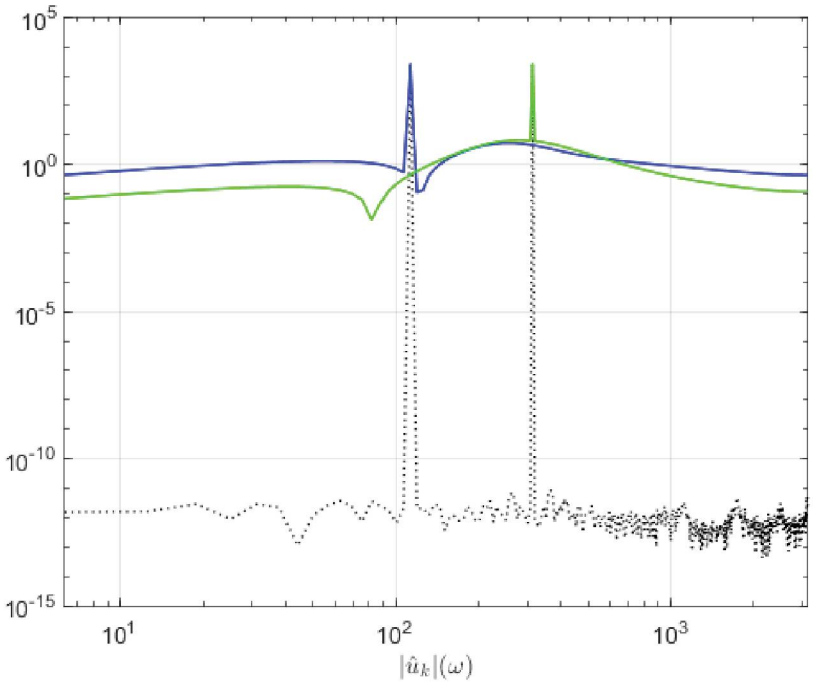
Ao analisar o espectro do sinal de entrada na Figura 21, pode-se observar que as frequências de 18 e 50Hz estão destacadas, não existindo outros picos. Desta forma, ao aplicar o algoritmo VMD, espera-se que os modos estimados estejam centrados nessas duas frequências. De fato, pode-se observar na Figura 22 que o algoritmo detectou e estimou corretamente os centros dessas duas frequências.

Figura 21: Espectro do sinal C



Fonte: O autor(2019).

Figura 22: Espectro do sinal C e modos estimados pelo algoritmo VMD

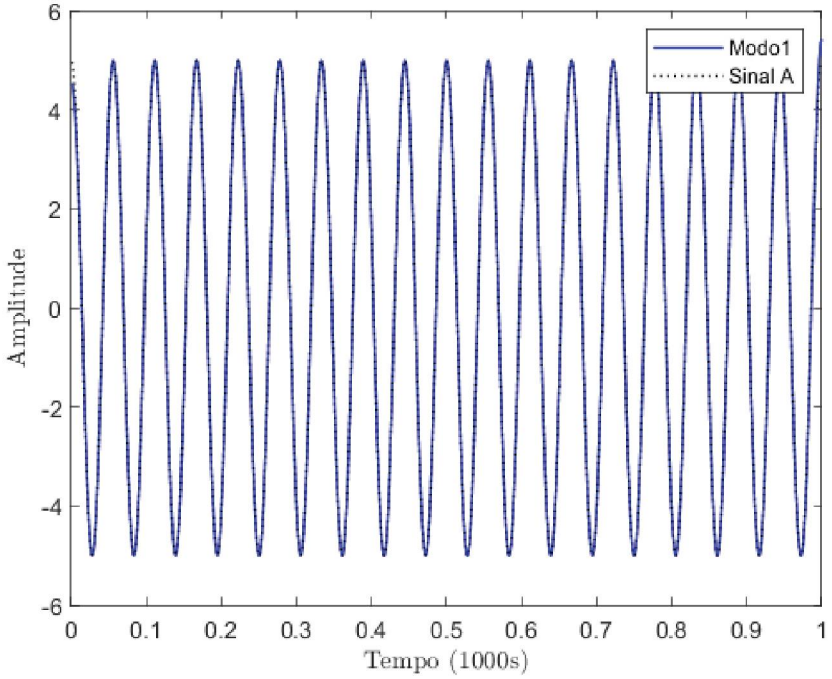


Fonte: O autor(2019).

A etapa final do VMD consiste em recuperar o sinal no domínio do tempo, em sua banda base. As Figuras 23 e 24 mostram a comparação dos sinais recuperados com

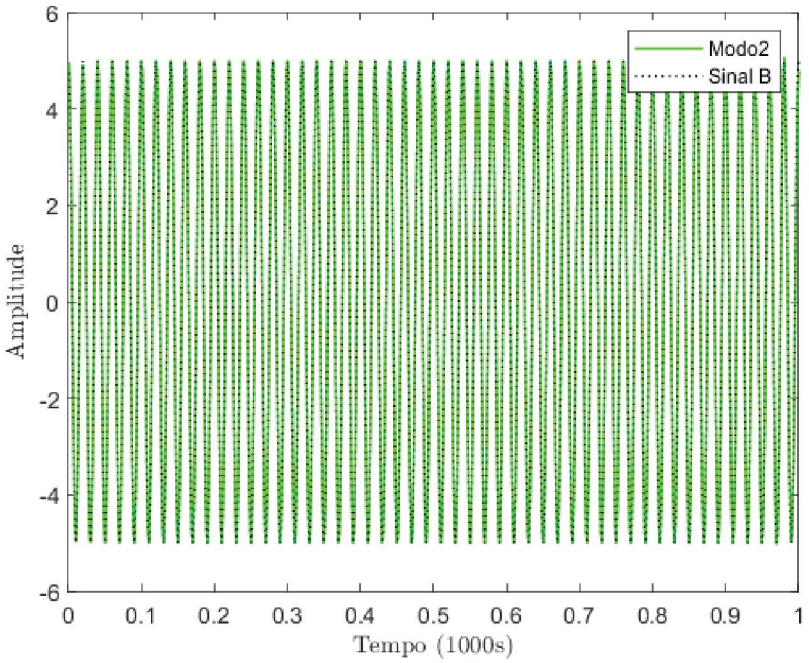
os sinais originais A e B.

Figura 23: Modo 1 de $18Hz$ reconstruído e comparado ao sinal A



Fonte: O autor(2019).

Figura 24: Modo 2 de $50Hz$ reconstruído e comparado ao sinal B

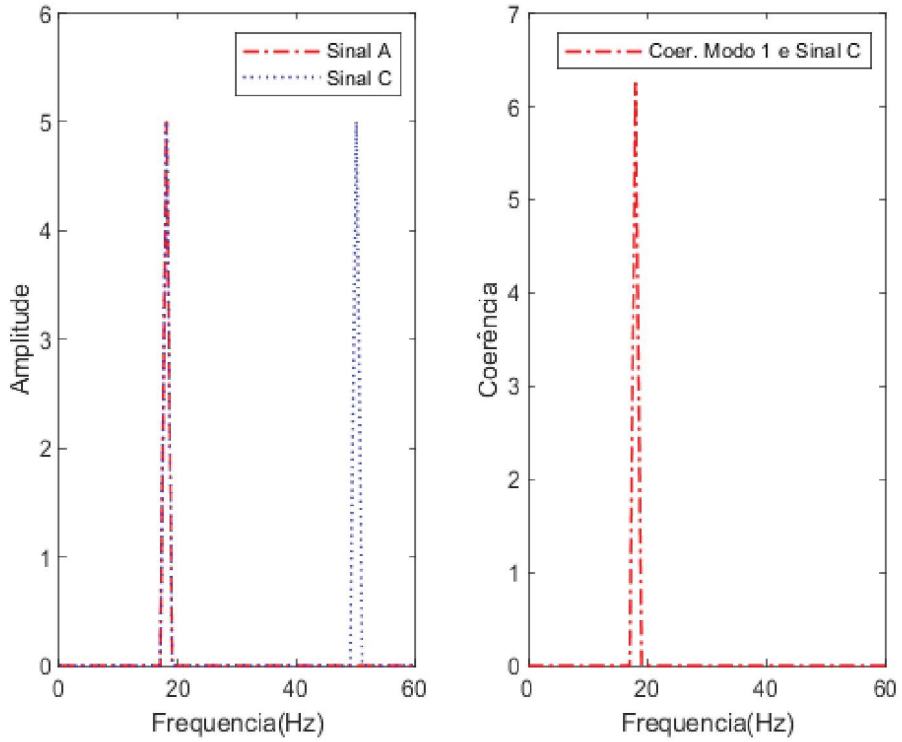


Fonte: O autor(2019).

2.2.3 AVALIAÇÃO DA COERÊNCIA

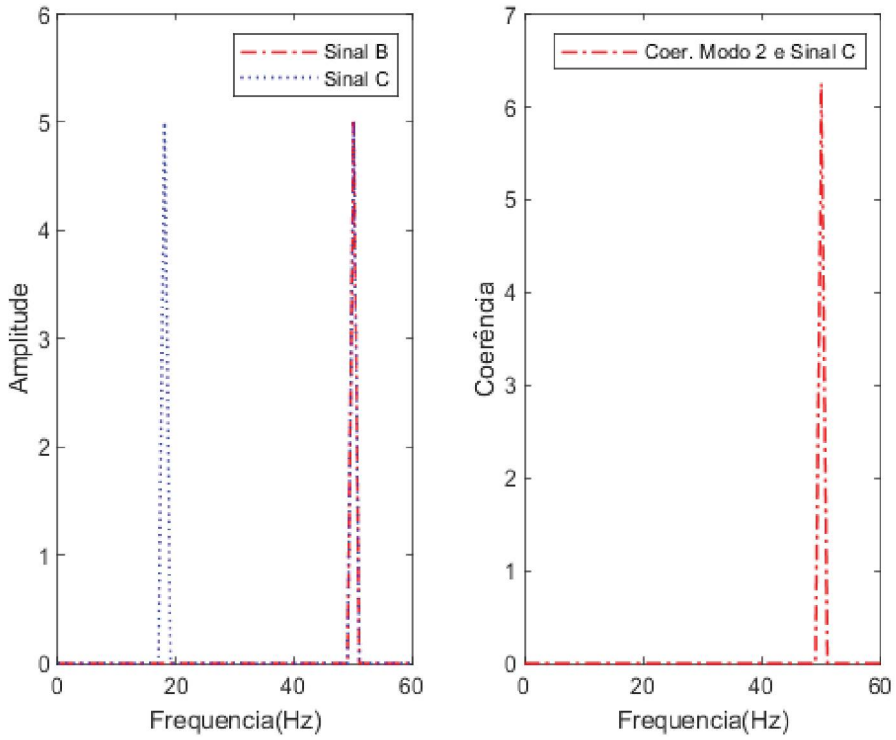
Para concluir o exemplo de aplicação da técnica VMD será avaliado a coerência entre os sinais reconstruídos após a aplicação da técnica e o sinal composto. O quê se espera obter com esta avaliação é a indicação de que os dois sinais possuem apenas uma única frequência em comum, sendo esta a única propriedade que compartilham. Desta forma, ao avaliar a coerência entre o sinal reconstruído do modo 1 e o sinal C, apenas a frequência de $18Hz$ é a propriedade compartilhada entre ambos, conforme Figura 25. Já para o sinal reconstruído do modo 2 apresentado na Figura 26, existe coerência apenas na frequência de $50Hz$.

Figura 25: Coerência entre o modo 1 de $18Hz$ e o sinal C



Fonte: O autor(2019).

Figura 26: Coerência entre o modo 2 de $50Hz$ e o sinal C



Fonte: O autor(2019).

2.2.4 ASPECTOS A SEREM DESTACADOS

O algoritmo VMD tem a capacidade de extrair componentes de um sinal limitados em banda, independentemente da sua complexidade e de quão misturados estão. A única ressalva é para modos muito próximos ao limite da frequência de amostragem do sinal (*Nyquist frequency*). A técnica VMD evita efetivamente a propagação de erros decorrentes do tradicional procedimento de filtragem recursiva presente no EMD, e portanto, alcança resultados mais precisos na reconstrução das componentes. Outro ponto vantajoso da técnica, é que para a remoção total do ruído de um sinal, basta apenas eliminar o multiplicador de *Lagrange* do algoritmo, pois ele é o responsável por garantir a fidelidade do sinal reconstruído ao sinal original.

Pode-se observar também que um problema presente na técnica SSA, por exemplo, da não garantia de separabilidade, que compromete a etapa de previsão, não ocorre neste caso, principalmente pelo fato do VMD não utilizar conceitos da álgebra linear, onde os critérios de ortogonalidade devem ser satisfeitos. No VMD os modos são atualizados por filtragem simples através de um filtro de *Wiener*, e isso ocorre diretamente no domínio de Fourier com um sintonizador ajustado à frequência central atual. Então as

frequências centrais são atualizadas como o centro de gravidade do espectro de potência do modo, e finalmente, o multiplicador Lagrange, que impõe a reconstrução exata do sinal, é atualizado aplicando o *dual ascent*. Porém existem algumas desvantagens, uma delas está na própria fundamentação da técnica VMD, que considera as componentes do sinal como sendo limitadas em banda e separadas no domínio da frequência. Este pressuposto torna a VMD menos efetiva quando aplicada em sinais não estacionários de banda-larga, especialmente quando os espectros das componentes estão sobrepostos.

2.3 RESUMO DO CAPÍTULO

Neste capítulo foi apresentado a fundamentação das técnicas SSA e VMD, além de discutir os conceitos de correlação e coerência espectral. Na seção em que a técnica SSA foi tratada, foi dado ênfase ao formalismo matemático de cada etapa que constitui a técnica; também foi apresentado uma aplicação prática da técnica. O capítulo também destaca algumas limitações da SSA, principalmente pela mesma ser altamente dependente da escolha correta dos parâmetros que estão ligados ao comprimento da janela L, ao número de componentes a ser decomposta a série temporal e a forma como elas serão agrupadas na etapa de reconstrução do sinal. Não há uma definição clara na literatura sobre o melhor meio de selecionar as auto-triplas e agrupar as componentes para reconstruir o sinal, sendo esta uma das maiores dificuldades na aplicação da técnica.

O capítulo também apresentou os conceitos da técnica VMD, de forma que seja possível ao leitor associar e assimilar a fundamentação da técnica, que é baseada estritamente nos conceitos de mistura de sinais e deslocamento de sinais no domínio da frequência. Também foi apresentado o conceito de coerência espectral, muito útil para identificar frequências semelhantes entre dois sinais. Como forma de estimular o entendimento da técnica, foi apresentado um exemplo de aplicação da VMD. No exemplo é obtido a decomposição de um sinal que contém duas frequências conhecidas a *priori*, e a partir dessas componentes é analisado a coerência espectral entre ambos os modos extraídos e o sinal original. Os resultados obtidos no exemplo mostram o potencial da técnica, porém também revela suas limitações, tais como: necessidade de conhecimento das frequências que compõe o sinal a *priori*, a técnica só pode ser aplicada à sinais limitado em banda (*narrow band*), o VMD é altamente dependente do parâmetro α que controla a largura de banda do filtro de Wiener, e não há ainda um detalhamento ou fundamentação conceitual para facilitar a escolha de tal parâmetro.

3 MODELO HÍBRIDO PROPOSTO

Este capítulo apresenta a estrutura do modelo híbrido proposto para aplicação na etapa de decomposição de uma série temporal em várias componentes, e posterior aplicação na previsão de séries temporais. Também é apresentado o algoritmo do modelo desenvolvido, além da fundamentação dos conceitos de demodulação de sinais, útil para entendimento do modelo proposto.

3.1 MODELO HÍBRIDO VMD-SSA

Dirigindo-se à decomposição de séries temporais normalmente encontrada na literatura de previsão de séries temporais, geralmente as componentes sazonais e de tendência são removidas, resultando apenas em uma componente, normalmente com flutuações aleatórias porém preservando a variância e média ao longo do tempo, ou seja, é uma componente estacionária. Desta forma, o objetivo comum na decomposição de séries temporais é tornar um sinal não estacionário em estacionário (BENDRE; MANTHALKAR, 2019; TESSIER; ARMSTRONG, 2015; MORETTIN; TOLOI, 2006; CHATEFIELD, 2000).

Tomando como referência o que foi apresentado no capítulo anterior, apesar da técnica SSA ter se popularizado e várias aplicações práticas podem ser encontradas na literatura, a mesma ainda sofre algumas limitações. Um aspecto crítico da SSA é a reconstrução da série temporal original, onde é utilizado vários pressupostos sobre sua estrutura subjacente para descartar o ruído ou determinar a maneira de combinar as várias componentes. Essa reconstrução depende da escolha das componentes da decomposição que compõe a matriz de covariância da matriz de trajetória. Na maioria das aplicações, esta seleção é baseada no conhecimento e experiência prévia do pesquisador e em uma variedade de regras práticas, tais como a análise gráfica dos autovalores da matriz de covariância (GOLYANDINA et al., 2001). Porém deve-se ressaltar que, a própria natureza da estrutura do autovalor da matriz de covariância implica que esses autovalores diminuem de forma exponencial, sendo desta forma, uma abordagem não tão bem funda-

mentada. Além disso, falta um raciocínio teórico sobre o porquê que certos pares devem ser selecionados, pode ser aconselhável, por exemplo, considerar a utilização de toda essa informação, ponderada adequadamente pelos dados da série original (PAPAILIAS; THOMAKOS, 2017).

É possível encontrar na literatura trabalhos propondo a utilização de técnicas de clusterização para a seleção correta dos pares autovalor-autovetor, como por exemplo, análise de componentes principais, análise de agrupamentos por clusterização hierárquica integrada com análise de componentes principais, ou ainda a análise gráfica de vetores singulares, ou o uso de no mínimo 50% a 100% das auto-triplas, etc. (TEIXEIRA et al., 2005; PAPAILIAS; THOMAKOS, 2017). Não há desta forma, um método fundamentado que garanta a melhor escolha das auto-triplas para a etapa de agrupamento.

Esse problema é acentuado quando se analisa sinais com comportamentos complexos, como por exemplo, séries de velocidades de vento, que normalmente apresentam alta variabilidade nas componentes de frequência mais elevadas, além de mudanças abruptas no padrão de tendência. Uma possível solução seria realizar a decomposição do sinal no maior número de componentes possíveis, ou seja, ir até o limite de $L \leq \frac{N}{2}$. Porém, a dificuldade na realização de análise gráfica dos autovalores aumenta substancialmente, considerando o fato de que uma série de velocidades de vento tem amostras obtidas a cada intervalo de 10min, uma série mensal tem comprimento N aproximado de 4320 amostras (para o caso de 30 dias), desta forma, teria-se mais de 2000 componentes para serem analisadas, e partir desta análise, decidir quais auto-triplas devem ser agrupadas e a quais grupos deveriam pertencer, inviabilizando desta forma o método.

O modelo híbrido proposto neste trabalho combina as técnicas VMD e SSA conforme Figura 29 e Algoritmo 4. A técnica VMD possui como característica, a mais importante a ser considerada ao incorporá-la no modelo híbrido, a extração de componentes da série temporal no domínio da frequência, de forma que cada componente esteja realmente centrada em frequências distintas, delimitadas por uma largura de banda. Desta forma, o problema de separabilidade existente na SSA é amenizado ao aplicar no estágio inicial o VMD como algoritmo de decomposição. Porém se a série temporal não é estacionária, a mesma possuirá propriedades de sinais banda larga (*wide band*).

As propriedades de banda larga dos sinais não estacionários são induzidas principalmente por seu comportamento de modulação em frequência. Uma saída para esse problema é tornar o sinal estacionário, obtendo um sinal limitado em banda (*narrow band*) para ser aplicado ao VMD e obter a extração de cada modo. Porém pode-se perder

informações relacionadas a componentes periódicas presente no sinal, por exemplo, ao aplicar técnicas de diferenciação para remover a não estacionariedade do sinal antes de realizar a decomposição.

Na abordagem proposta neste trabalho, para contornar o problema de aplicabilidade do VMD apenas em sinais limitados em banda, o algoritmo VMD sofreu pequenas alterações para poder ser aplicado a sinais não estacionários, de forma que seja preservada as características originais de não estacionariedade do sinal ao reconstruir o sinal original. Como o VMD busca identificar o centro de frequências, ou a banda base de um sinal, para então extrair os modos relativos a tais frequências (ver Seção 2.2), essa característica do VMD foi utilizada como estágio de identificação para o modelo proposto.

A idéia do estágio de identificação é ter uma estimativa das frequências presentes na série temporal considerando uma largura de banda limitada, em outras palavras estamos apenas identificando as componentes AM, sem nos preocuparmos com as componentes que se sobrepõe no domínio da frequência (lembrando que essa é uma característica dos sinais não estacionários AM-FM). As frequências obtidas nesta estimativa, que podem ser sinais harmônicos ou ainda em oposição de fase, mas não sinais compostos AM-FM, são então utilizadas como referência para o estágio de decomposição do VMD que estará sintonizado nessas frequências, porém antes de realizar decomposição, o comportamento não estacionário é removido através da demodulação do sinal, tem-se então o sinal limitado em banda e uma componente residual.

Conforme demonstrado em Chen et al. (2017), um sinal *wide band* demodulado de forma ideal será *narrow band* se a frequência do operador de demodulação se igualar a frequência presente no sinal a ser demodulado, ou seja, se $f_d(s) = f(s)$. Desta forma, faz-se necessário definir o operador de demodulação para melhor entendimento do modelo proposto.

3.1.1 DEMODULAÇÃO

A abordagem utilizando demodulação do sinal original não estacionário foi aplicada por (CHEN et al., 2017) para separar sinais coletados de máquinas rotativas com velocidades variáveis e se mostra bastante promissora. A ideia de demodulação para remover a componente FM, resultando em um sinal limitado em banda pode ser encontrada também em Wang et al. (2014). Para melhor entendimento de como a demodulação remove o componente FM de um sinal AM-FM, considere o sinal não estacionário definido na Equação 3.1 (maior detalhamento em Chen et al. (2017)).

$$g(t) = \sum_{i=1}^{k} C_i(t) cos\Big(2\pi \int_0^t f_i(s)ds + \phi_i\Big) + \eta(t) \tag{3.1}$$

onde k é o número de modos (ou componentes) do sinal; $C_i(t)$ é a amplitude instantânea do i^{th} modo de oscilação do sinal; $f(t)$ é a frequência instantânea (IF); ϕ é a fase inicial do sinal e $\eta \approx \mathcal{N}(0, \sigma^2)$ denota ruído branco Gaussiano com média $\mu = 0$ e variância σ^2.

Não há uma definição matemática formal da representação da largura de banda de um sinal AM-FM, utilizamos aqui a abordagem proposta em Wang et al. (2014) e Chen et al. (2017), que considera que o sinal não estacionário esta mais relacionado ao comportamento das componentes AM e FM que compõe o sinal. Desta forma, a representação da largura de banda é a adotada na Equação 3.2,

$$BW = BW_{AM} + BW_{FM} \tag{3.2}$$

onde $BW_{AM} = 2F_a$ e denota a largura de banda do sinal AM; F_a representa a máxima frequência da amplitude instantânea de $C_i(t)$; BW_{FM} é a largura de banda do sinal FM, determinada pela Regra de Carson. Se considerarmos o efeito de cada tipo de sinal na determinação da largura de banda do sinal não estacionário, na prática o que se tem de forma empírica, segundo Chen et al. (2017), é que a largura de banda do sinal não estacionário é determinada predominantemente pela largura de banda do sinal FM, dado que $BW_{FM} \gg BW_{AM}$.

Definido o conceito de largura de banda, as etapas da demodulação de um sinal AM-FM podem ser então exploradas. Considere a definição de sinal analítico aplicado ao sinal AM-FM, de forma que podemos representar $g(t)$ como $g_a(t)$, conforme Equação 3.3

$$\begin{aligned} g_a(t) &= g(t) + j\mathbf{H}(g(t)) \\ &\approx C_i(t) e^{\left(j2\pi \int_0^t f_i(s)ds + \Phi_i\right)} \end{aligned} \tag{3.3}$$

onde $j^2 = -1$ e $\mathbf{H}(.)$ denota a transformada de Hilbert. Pode-se então definir o par de operadores, seguindo a nomenclatura adotada por (CHEN et al., 2017), o operador de demodulação Φ^- e de modulação Φ^+ é dado por:

$$\Phi^{-}(t) = exp\Big(- j2\pi\Big(\int_0^t f_d(s)ds\Big) - f_c t\Big) \tag{3.4}$$

$$\Phi^{+}(t) = exp\Big(j2\pi\Big(\int_0^t f_d(s)ds\Big) - f_c t\Big) \tag{3.5}$$

onde f_d denota o operador da função de frequência e f_c é a frequência da portadora do sinal, ou frequência central. O sinal demodulado $g_a^d(t)$ é obtido pela multiplicação do sinal $g_a(t)$ pelo operador de demodulação Φ^{-}, tal como representado na Equação 3.6,

$$\begin{aligned} g_a^d(t) &= g_a(t)\Phi^{-}(t) \\ &= g_a(t)exp\Big(j(2\pi f_c t + 2\pi \int_0^t (f(s) - f_d(s))ds + \phi)\Big) \end{aligned} \tag{3.6}$$

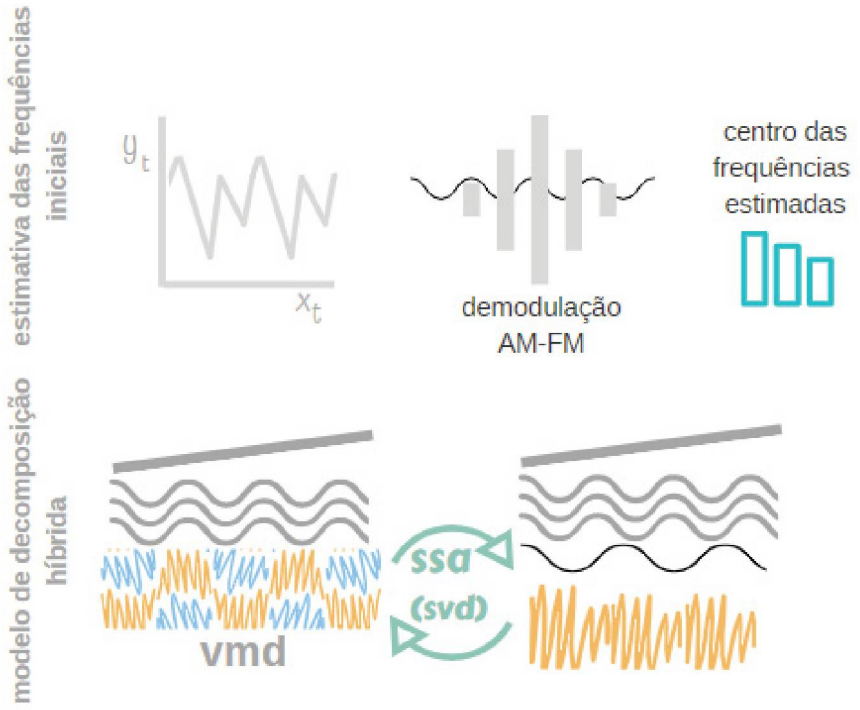
Desta forma $g_a^d(t)$ será um sinal puramente AM, com frequência centrada em f_c considerando $f_d(s) = f(s)$. A largura de banda do sinal demodulado $g_a^d(t)$ será um sinal limitado em banda (*narrow band*) com $BW_{min} = BW_{AM}$. O sinal analítico original pode ser recuperado, bastando multiplicar $g_a^d(t)$ por $\Phi^{+}(t)$, ou seja, aplicar a modulação ao sinal demodulado conforme Equação 3.7,

$$g_a(t) = g_a^d(t)\Phi^{+}(t) \tag{3.7}$$

Em linhas gerais, neste modelo utilizado para decomposição da série temporal, o algoritmo de demodulação é aplicado em dois estágios, conforme Figura 27 de tal forma que:

- No primeiro estágio a função principal é eliminar o componente FM, resultando em um sinal $g_a^d(t) = g_a(t)exp\Big(j(2\pi f_c t + 2\pi \int_0^t (f(s) - f_d(s))ds + \phi)\Big)$ limitado em banda, com $BW = BW_{AM}$ e centrado em f_c;
- No segundo estágio, que consiste na aplicação do algoritmo VMD original proposto por Dragomiretskiy e Zosso (2014), o sinal $g_a^d(t)$ é demodulado em uma banda base de forma que $g_a^b(t) = g_a^d(t)exp(-2j\pi f_c t)$, resultando no deslocamento da frequência do sinal $g_a^d(t)$ sem alterar a largura de banda (ver Seção 2.2);

Figura 27: Estrutura do modelo proposto

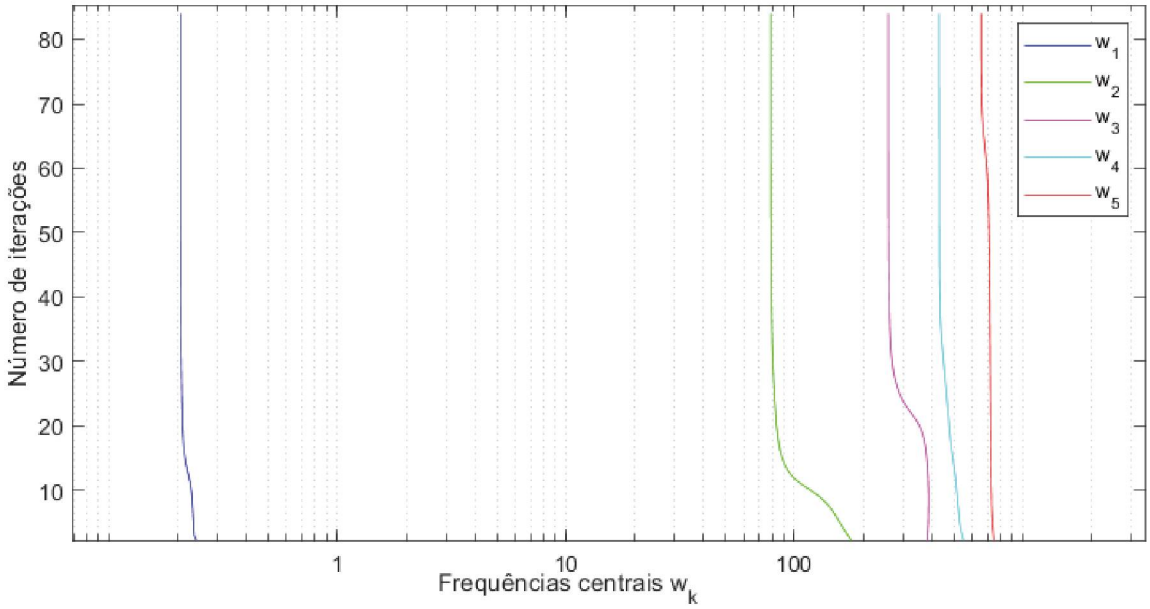


Fonte: O autor(2019).

O sinal residual será *wide band*, e é obtido através do resultado da diferença do sinal original e o sinal demodulado, o mesmo é aplicado então no próximo estágio que é decomposição no domínio do tempo e filtragem através da SSA, conforme detalhado no Algoritmo 4.

Desta forma, a alteração realizada no algoritmo VMD consiste em inserir um estágio de demodulação inicial após a identificação das frequências. Ao aplicarmos esta etapa de identificação, o algoritmo só é interrompido quando os centros das frequências para o conjunto dos k modos selecionados são encontrados, ou quando o mínimo valor de erro no processo de convergência é atingido. Para exemplificar este procedimento a Figura 28 apresenta um exemplo de convergência após 90 iterações, para a obtenção do centro de frequência de cinco (5) modos para uma série de velocidades de vento com 1000 amostras. Com este exemplo pode-se observar a rapidez com que o algoritmo converge durante o processo de otimização da identificação dos modos de frequências contidos no sinal.

Figura 28: Centro de frequências para cinco modos convergência após 90 iterações



Fonte: O autor(2019).

3.1.2 ALGORITMO DO MODELO PROPOSTO

Ao combinar as duas técnicas VMD e SSA em um processo multiestágio, o VMD é utilizado como algoritmo de decomposição no primeiro estágio, extraindo apenas as componentes limitadas em banda (*narrow-band*) e a SSA no segundo estágio de forma sequencial, como algoritmo de decomposição e de alisamento da componente residual. Desta forma os passos da SSA são aplicados de forma descontinuada para cada sequência de filtragem da componente residual, ou seja, após o VMD extrair as k-componentes de interesse e fazer a reconstrução das mesmas no domínio do tempo, o resíduo dessas componentes são processadas pelo SSA de forma sequencial, isto é, são novamente decompostas em subcomponentes; tem-se então d-subcomponentes, das quais $d-1$ são decompostas novamente pelo SSA, até que todas componentes sejam processada pelo SSA, que é aplicado neste estágio até o passo de agrupamento.

O algoritmo SSA, até a etapa de agrupamento, funciona como algoritmo de decomposição e também como alisador das componentes. Para tal, basta selecionar apenas as $d-1$ subcomponentes para serem agrupadas. Desta forma, parte do sinal original não fará parte das componentes agrupadas, obtendo componentes alisadas (ou suavizadas) após a etapa de agrupamento. Quando todas as $d-1$ foram processadas pelo SSA finaliza então o estágio de decomposição sequencial ou multiestágio. A partir desse ponto é iniciado o passo de previsão, ou seja, o algoritmo de *forecasting* é aplicado em cada

componente de forma independente e então, após ter obtido a previsão para o horizonte à frente desejado, o último passo da SSA é finalizado (que é a reconstrução da série através da soma das componentes obtidas ao aplicar as médias diagonais), reconstruindo dessa forma a série temporal. Maior detalhamento é apresentado no Algoritmo 4, enquanto que a estrutura do modelo proposto é apresentado na Figura 29. Para a etapa de *forecasting*, qualquer técnica de previsão pode ser aplicada, esse assunto será tratado mais adiante.

O valor de N no Algoritmo 4 representa o número de elementos (ou amostras) da série temporal. Conforme descrito em Golyandina et al. (2001), para séries curtas e com estrutura complexa, o valor de N de ser maior que 50, caso contrário a escolha da janela L pode ser tornar um problema difícil de ser contornado, principalmente pelo fato de que a série não estacionária possui múltiplos períodos e que devem estar contidos na escolha da janela L, que por sua vez esta associada à restrição de assumir valores tais que $L < \frac{N}{2}$ (GOLYANDINA et al., 2001).

Algoritmo 4 Modelo Híbrido VMD-SSA

Entrada: Série Temporal X

Entrada: N número de elementos na série temporal

Entrada: Número de k modos para a decomposição no domínio da frequência

Entrada: Número de d subcomponentes para decomposição no domínio do tempo

Saída: Resultados das previsões

1: Identificar as frequências dos $k - modos$ desejados presentes na série temporal {Algoritmo Demodulação AM-FM}

2: Decompor X no domínio da frequência em $k - modos$ de forma que os modos obtidos sejam limitados em banda {Algoritmo VMD}

3: Reconstruir cada modo de forma independente, obtendo x_k componentes no domínio do tempo {Algoritmo VMD}

4: Obter a componente residual X_r de forma que $X_r = X - \sum_{i=1}^{k} x_i$
{*Início da Decomposição Sequencial SSA*}

5: **para** $i = 1$ **até** número de subcomponentes d **faça**

6: Fazer a Decomposição da componente X_r, obtendo d subcomponentes {*Embedding e SVD steps no SSA*}

7: Realizar o Agrupamento selecionando apenas $(d - 1)$ subcomponentes de forma a obter X^i filtrada{*Grouping step no SSA*}

8: Atualizar a componente residual X_r, de forma que $X_r^{(i)} = X_r^{(i-1)} - X^{(i)}$

9: **fim para**
{*Fim da Decomposição Sequencial SSA*}

10: Construir a matriz de componentes, que deve ter a dimensão $C_{N\times(k+2)}$
{*Início da Etapa de Previsão*}

11: Aplicar operador de atraso (Lag) em cada componente da matriz $C_{N\times(k+2)}$

12: Realizar previsão para cada componente, aplicando a técnica desejada

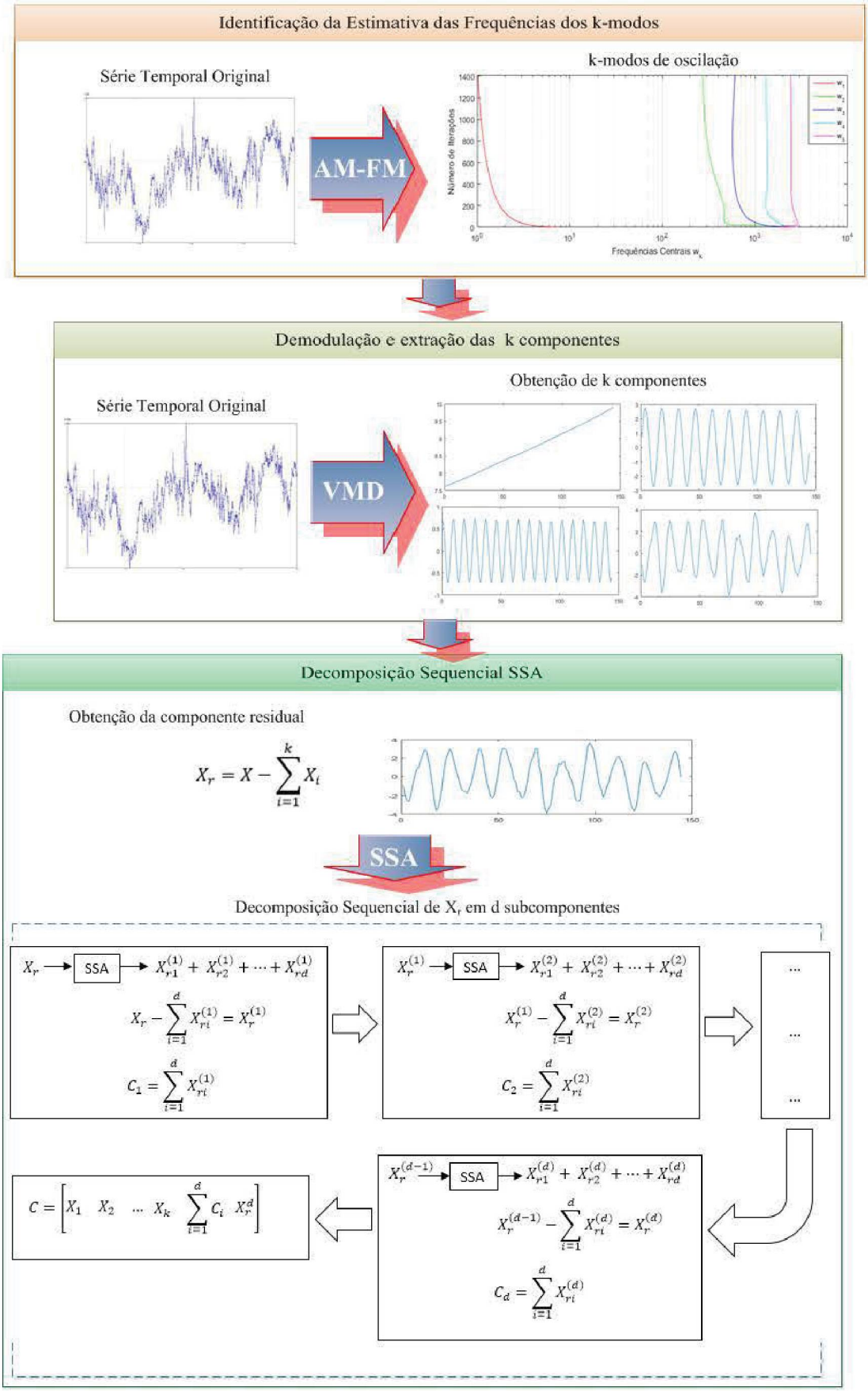
13: Reconstruir a série temporal (real e prevista) através da soma das $(k+2)$ componentes {*Reconstruction step no SSA*}

14: Avaliar a qualidade das previsões

15: Apresentar os resultados das previsões

Fonte: O autor(2019).

Figura 29: Estrutura do modelo proposto sem a etapa de previsão



Fonte: O autor(2019).

3.2 RESUMO DO CAPÍTULO

O modelo híbrido proposto foi introduzido neste capítulo, onde buscou-se detalhar cada etapa do algoritmo. O capítulo também apresentou a estrutura do modelo VMD-SSA, para aplicação na etapa de decomposição de uma série temporal em várias componentes, e posterior aplicação na previsão de séries temporais. Também foi apresentado o algoritmo do modelo desenvolvido, resgatando a fundamentação dos conceitos de demodulação de sinais, de forma a facilitar o entendimento para o leitor.

4 APLICAÇÃO DO MODELO VMD-SSA

Este capítulo discute a aplicação do modelo proposto, visando avaliar a capacidade do mesmo em realizar a separação das componentes de forma que as mesmas sejam descorrelacionadas. O modelo de decomposição multiestágio é aplicado a diferentes série de velocidades de vento. Os resultados da decomposição são avaliados quanto ao grau de correlação que guardam entre cada componente, onde o potencial da técnica em obter componentes descorrelacionadas é avaliado do ponto de vista quantitativo.

4.1 EXPERIMENTOS NUMÉRICOS

As séries temporais de velocidades de vento, utilizadas com a abordagem proposta, são originárias de medições realizadas a 80 metros de altura do solo, amostradas em intervalos de 10 minutos. As mesmas são provenientes de uma torre de monitoramento anemométrico instalada em parque eólico construído na região dos municípios de Guanambi e Caetité, no estado da Bahia, disponibilizada pela empresa Centrais Eólicas Guanambi, referente ao Conjunto Eólico Licínio de Almeida e Planaltina.

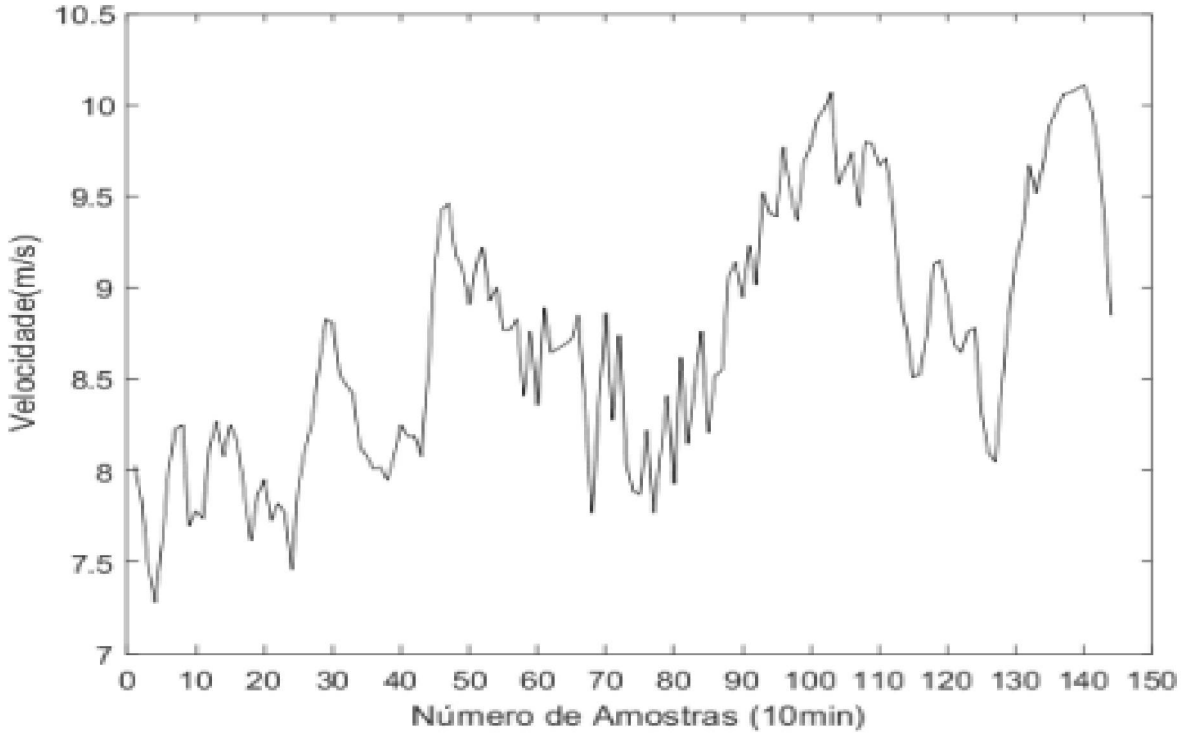
As séries disponíveis são referente aos meses de novembro de 2018 a janeiro de 2019. Esse conjunto de dados foi agrupado de forma a compor um conjunto de 6 (seis) séries de velocidades de vento, em que 5 séries diárias com extensão de 144 amostras formam uma parte do conjunto, e a outra parte contém todos os dados relativos ao trimestre, porém integralizados em velocidades médias diárias, sendo desta forma o conjunto composto por:

- Amostras a cada 10 minutos, totalizando 144 amostras/dia, compondo uma série diária. Estas séries serão aplicadas para obtenção da previsão de 1 à 12 passos-à-frente (10 min à 2 horas). Serão avaliadas as 5 séries diárias de forma individual, ou seja, um experimento para cada um dos dias do intervalo de 01 a 05 de janeiro de 2019;

- Amostras a cada 1 dia, relativas a média diária dos 144 períodos de observações, totalizando 92 amostras para o trimestre composto pelos meses de novembro/2018 a janeiro/2019. Esta série será aplicada para obter previsões de 1 à 12 dias à frente.

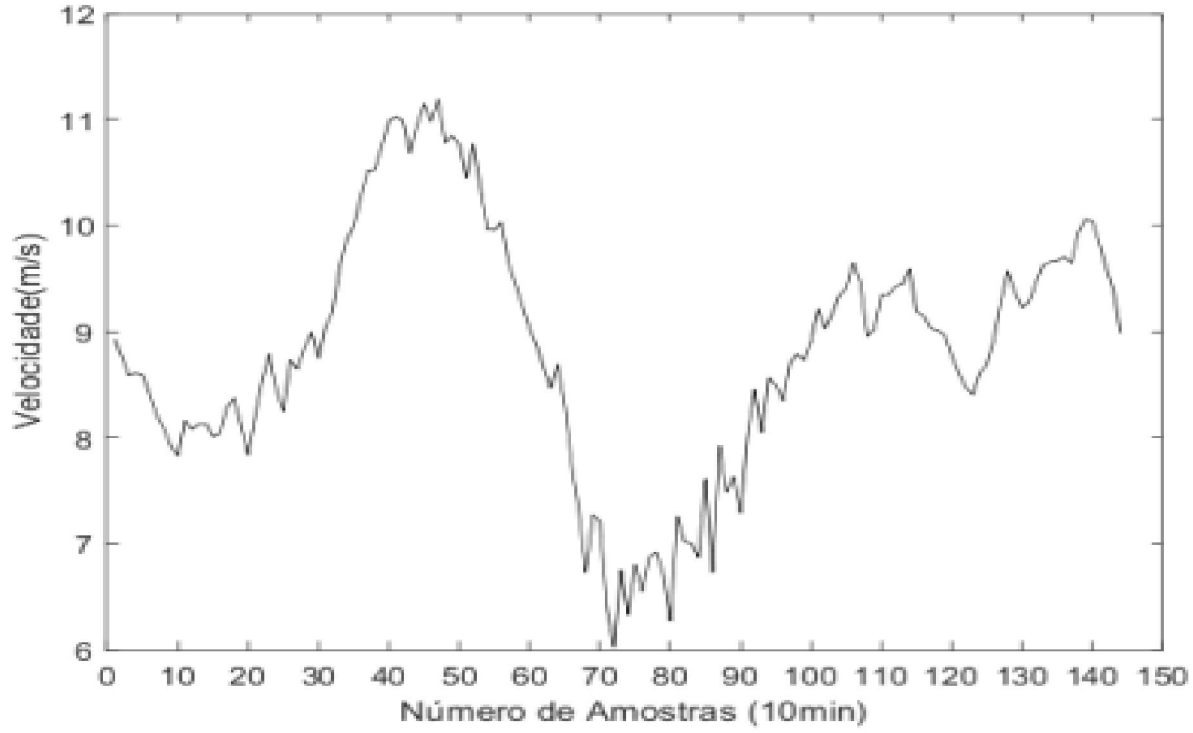
O comportamento destas séries estão representados nas Figuras 30 a 35. Pode-se observar que a grande maioria das séries adotadas são curtas, pois foi observado que o comportamento da série de velocidades de vento é influenciada por mecanismos com *lags* não muito extenso, ou seja, os lags significativos estão no intervalo de 1 a 25 amostras, como exemplo, o resultado da função de autocorrelação parcial (*facp*) para uma série contendo 144 amostras na Figura 36. Para maiores detalhes sobre a *facp* ver APÊNDICE B.

Figura 30: Série de velocidades de vento amostrada a cada 10min - 01 de janeiro 2019



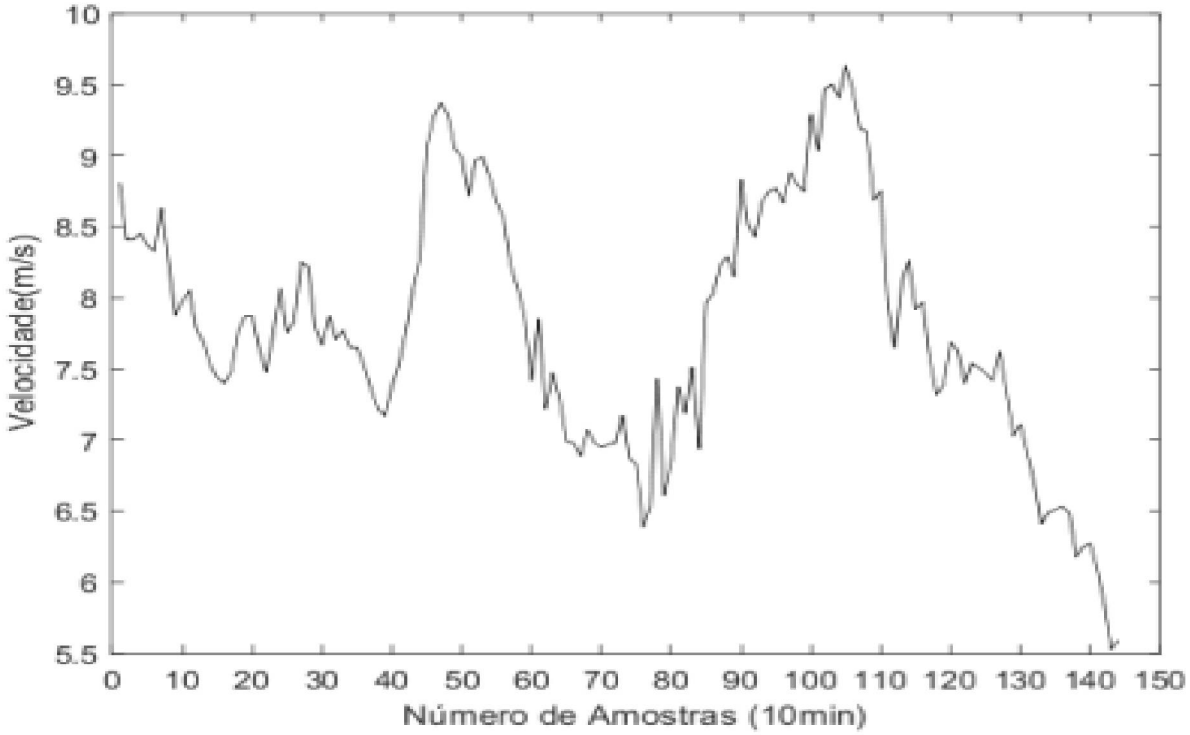
Fonte: O autor(2019).

Figura 31: Série de velocidades de vento amostrada a cada 10min - 02 de janeiro 2019



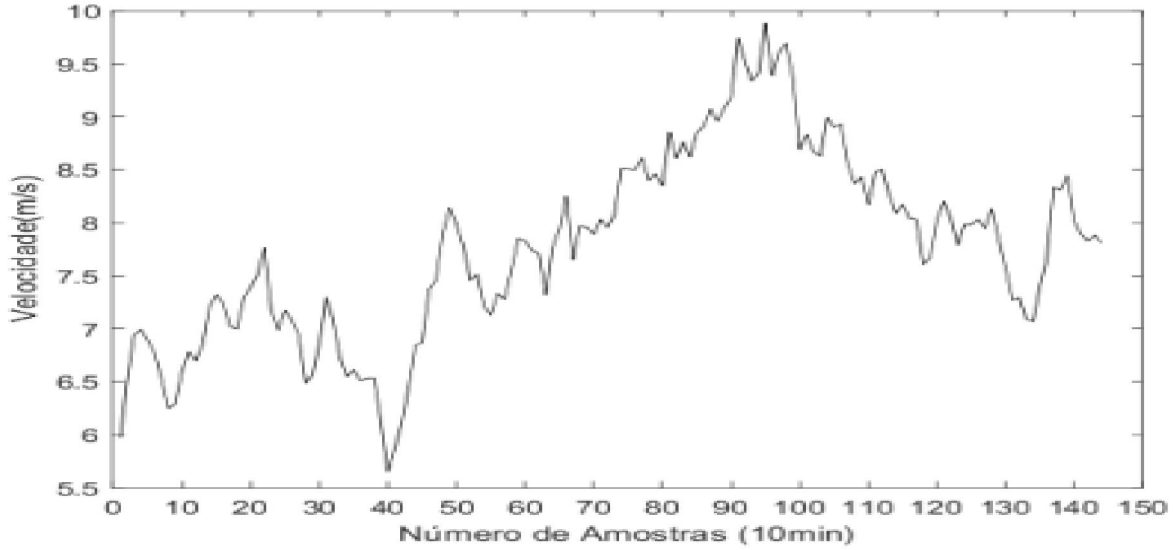
Fonte: O autor(2019).

Figura 32: Série de velocidades de vento amostrada a cada 10min - 03 de janeiro 2019



Fonte: O autor(2019).

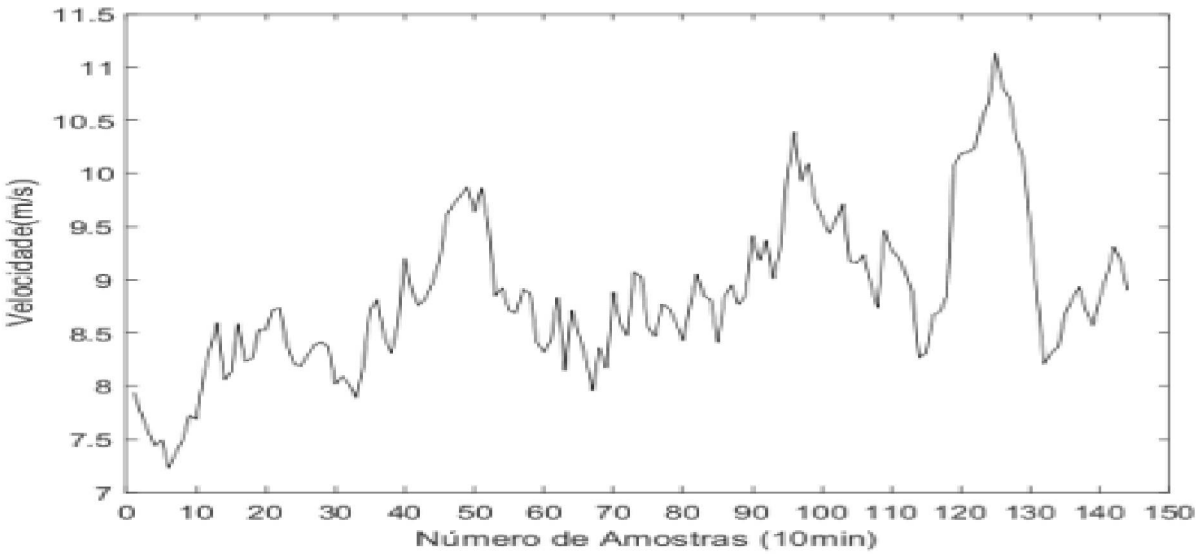
Figura 33: Série de velocidades de vento amostrada a cada 10min - 04 de janeiro 2019



Fonte: O autor(2019).

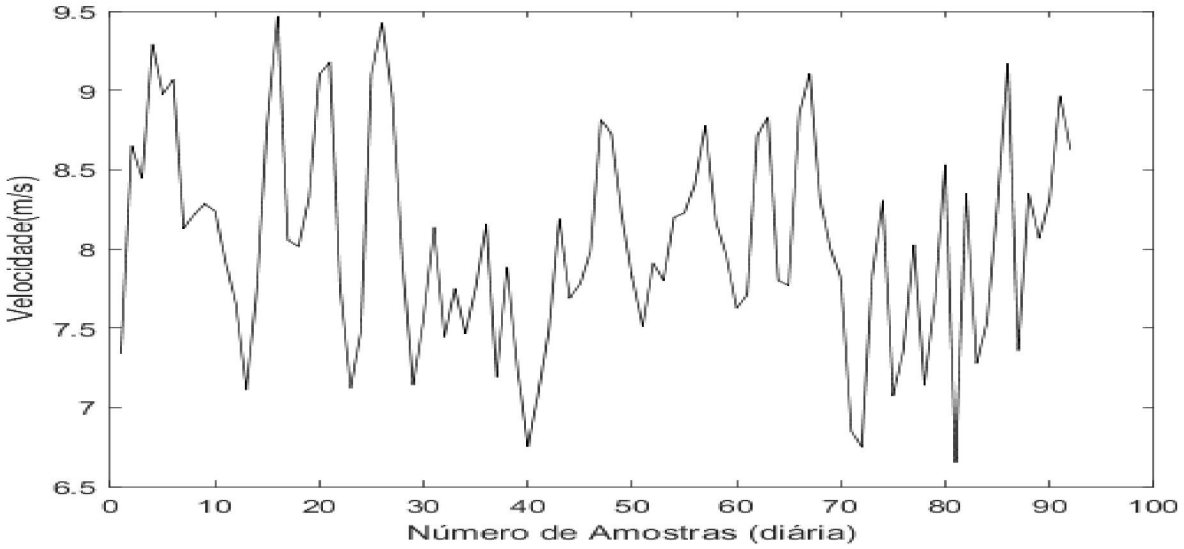
Importante observação a ser feita é com relação as variações abruptas de velocidades de vento contida nas séries representas nas Figuras 30 à 33. As variações possuem amplitudes entre 2 m/s (Fig. 30) até 5 m/s (Fig. 33).

Figura 34: Série de velocidades de vento amostrada a cada 10min - 05 de janeiro 2019



Fonte: O autor(2019).

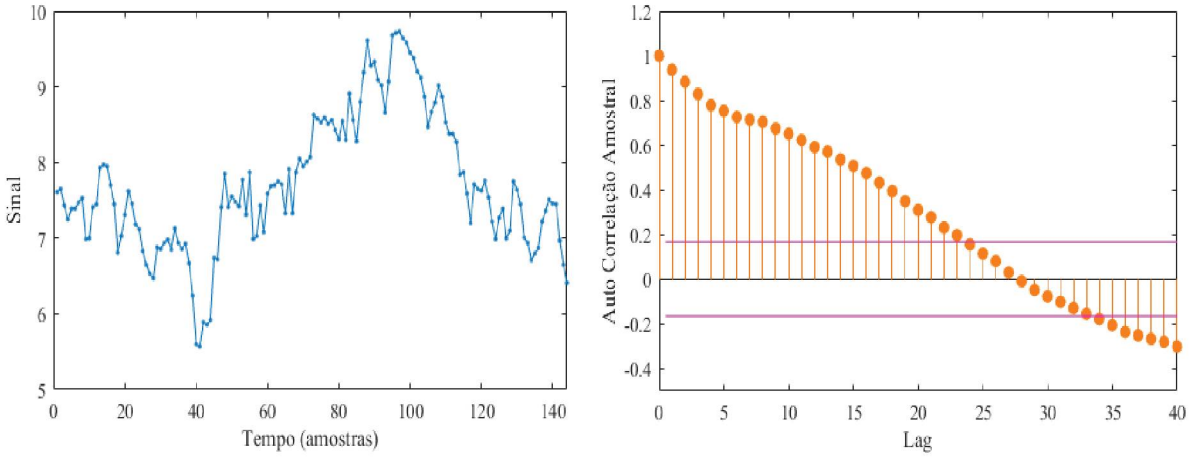
Figura 35: Série de velocidades de vento amostrada a cada 1 dia - trimestre composto pelos meses de novembro 2018 a janeiro 2019



Fonte: O autor(2019).

As séries de velocidades representadas nas Figuras 34 e 35 representam o caso extremo nornalmente verificado nessa região da Bahia, onde as variações de amplitudes tem um crescimento quase que vertical e rapidamente decrescem também em uma queda vertical. Para a série diária representada na Figura 35 esse efeito é mais destacado, mesmo considerando que os valores ali representados são médias das 144 amostras coletadas ao longo de um dia.

Figura 36: Função de autocorrelação parcial - série de velocidades de vento com 140 amostras



Fonte: O autor(2019).

4.1.1 PARÂMETROS DO ALGORITMO VMD-SSA

A seleção de parâmetros adotada neste trabalho seguiu as recomendações de Dragomiretskiy e Zosso (2014) para a técnica VMD, e Golyandina et al. (2001) para a SSA, respectivamente. A técnica VMD possui apenas dois parâmetros que necessitam ser selecionados com cautela para não comprometer a qualidade dos modos extraídos. O primeiro é o valor de α, que controla a largura de banda BW do modo a ser extraído, o segundo é o número de modos K a serem extraídos, que tem vínculo com a reconstrução do sinal.

O valor de α tem relação com a variação da largura de banda BW, como descrito anteriormente na Seção 2.2. De modo que, reduzindo o valor de α há uma aumento na largura de faixa do filtro de Wiener, resultando em aumento da largura de banda em torno da frequência central no sinal filtrado. Proporcionando desta forma, extrair componentes com alta fidelidade em relação ao sinal original, porém altamente correlacionadas. O que se deseja é o oposto, objetivando facilitar dessa forma, o ajuste do modelo de previsão para as componentes oscilatórias presente na série temporal. Ao elevar o valor de α há uma redução da largura de banda do sinal, permitindo que o nível de ruído incorporado ao modo estimado da componente em questão não esteja presente no sinal filtrado, além de garantir que a componente extraída não possua correlação (ou possua fraca correlação) com as componentes remanescentes.

Cabe ressaltar que para séries temporais de velocidades de vento, se utilizarmos α com valor muito pequeno, ao recuperar o sinal no domínio do tempo, têm-se componentes fortemente correlacionadas, o que não é o foco deste trabalho, pois o quê se deseja é obter componentes que representem o comportamento oscilatório e de tendência de maneira que sejam descorrelacionadas ou que possuem fraca correlação. Obtendo componentes com este comportamento, os níveis de ruído associado as suas amplitudes são mínimas, ou seja, as variações de amplitude e frequência devem ficar restritas a componente residual X_r. Neste trabalho, o algoritmo proposto possui um estágio de filtragem sequencial da componente X_r através da técnica SSA, desta forma, o valor de α será escolhido de forma que esteja dentro dos limites inferior e superior que serão detalhados a seguir.

Dragomiretskiy e Zosso (2014) estabelece um limite superior para o valor de $\alpha_{max} = 10.000$, caso se queira reconstruir os modos com mínimo erro em relação ao sinal original. Porém este não é o interesse nesta pesquisa, pois uma série temporal pode não ser perfeitamente limitada em banda, e o que se busca é uma estimava das frequências contidas no sinal, que serão separadas em uma etapa seguinte, como já discutido em capítulo anterior. Desta forma, como a proposta deste trabalho é construir um modelo híbrido, e

que o mesmo proporcione melhor separabilidade das componentes de um sinal, enquanto que as etapas de reconstrução e filtragem do sinal serão feitos através da SSA, o limite superior não é a melhor opção. Neste caso, será utilizado a frequência de amostragem do sinal (f_{sample}) para obter o limite inferior, respeitando o critério Nyquist (NYQUIST, 2002) conforme Equação 4.1, de forma que α_{min} (associado a maior largura de banda) possa ser determinado através da aplicação da Equação 4.3, onde N é o número de amostras da série temporal. Desta forma, tem-se um intervalo limite para poder arbitrar o valor de α,

$$BW < \frac{f_{sample}}{2} \tag{4.1}$$

$$f_{sample} = \frac{1}{N} \tag{4.2}$$

$$\alpha_{max} \geq \alpha \geq \alpha_{min} > BW \tag{4.3}$$

Outro parâmetro que deve ser definido no VMD é o número de modos ou componentes que se deseja extrair. Como não se conhece a *priori* quais frequências estão presente no sinal, essa escolha é arbitrária. Porém pensando na etapa de previsão, relativo ao modelo proposto, cabe definir alguma restrição neste sentido. Desta forma, considerando o fato que deverá ter um modelo de previsão associado a cada componente, quanto maior o número de componentes C_i maior o número de modelos necessários. Levando ainda em conta o fato de que se deseja aplicar um modelo de previsão paramétrico, o número de conjunto de parâmetros C_p (que contém n_p parâmetros) necessários cresce na proporção do número de componentes C_i multiplicada pelo número de modelos M_i, conforme Equação 4.4. Enquanto que o número de total de parâmetros N_p cresce proporcionalmente a n_p, conforme Equação 4.5

$$C_p = C_i M_i \tag{4.4}$$

$$N_p = n_p C_i M_i \tag{4.5}$$

Desta forma, arbitrariamente definimos neste trabalho como 5 (cinco) o número de componentes a serem extraídas mais a componente residual, após vários testes e verificações de correlações entre componentes, onde foi observado que para as séries de

velocidades em questão aumentar o número de componentes superando o valor de 5, não há ganhos significativos na redução da correlação, como será discutido mais adiante. Para a SSA, o único parâmetro a ser determinado é o valor de L, que representa o número de componentes nas quais se deseja decompor a série original. Em Golyandina et al. (2001) há um limite superior estabelecido para o valor de L, de tal forma que, deve-se observar a desigualdade apresentada na Equação 4.6, caso N seja impar, fazer $(N-1)/2$. Definimos o limite inferior como sendo 2, ou seja, o número mínimo de componentes a ser extraído é 2. Um bom método, segundo Golyandina et al. (2001), é utilizar L igual ou superior ao número de amostras onde ocorre mudança de tendência da série, ou seja, que minimiza o valor da correlação ρ entre componentes extraídas,

$$L \leq \frac{N}{2} \tag{4.6}$$

Desta forma os parâmetros para cada série temporal são apresentados na Tabela 2.

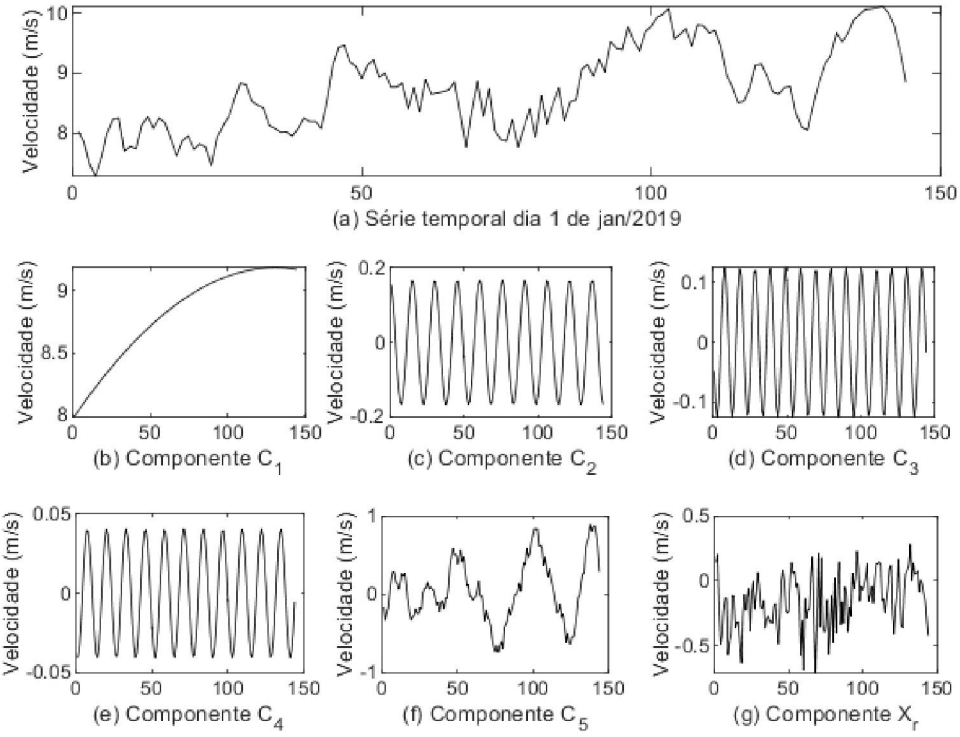
Tabela 2: Parâmetros adotados no VMD-SSA

Série Temporal Vel. Vento	Janela L (SSA)	α (VMD)
144 amostras (base 10min)	72	1000
92 amostras (base diária)	46	1000

Fonte: O autor(2019).

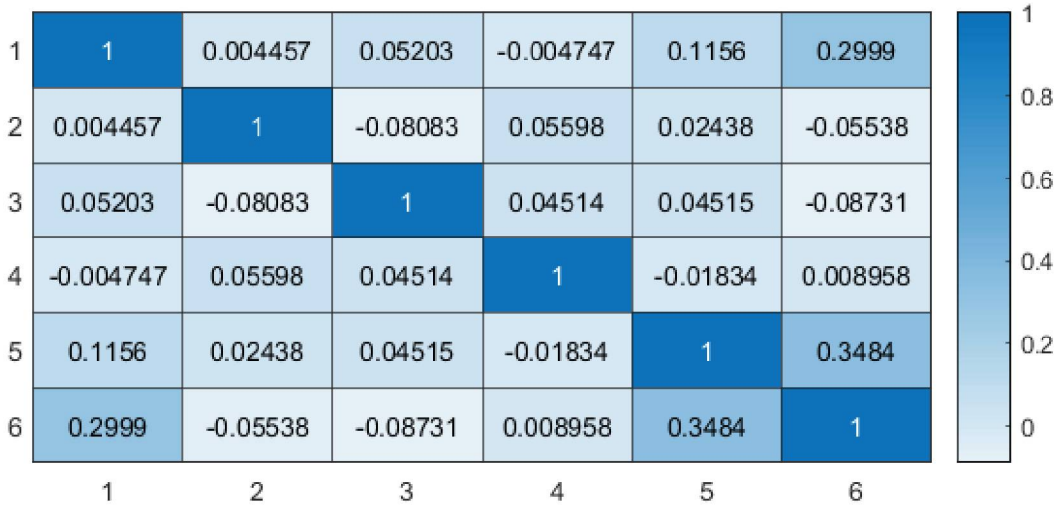
Para ilustrar a capacidade do modelo proposto em realizar a decomposição da série temporal de velocidades de vento, a Figura 37 de (a) a (g) apresenta a série de dados para o dia 1 de janeiro de 2019 e suas componentes obtidas pelo modelo VMD-SSA. A Figura 38 apresenta o grau de correlação entre as componentes C_1 a X_r resultantes da decomposição VMD-SSA. Esse conjunto de componentes será então utilizado como entrada para o modelo de previsão.

Figura 37: Decomposição VMD-SSA da série temporal velocidades de vento (10min) de 01/jan/2019



Fonte: O autor(2019).

Figura 38: Correlação entre componentes obtidas pela decomposição VMD-SSA para série vel. vento(10min) de 01/jan/2019



Fonte: O autor(2019).

4.1.2 DETERMINAÇÃO DOS PARÂMETROS DO MODELO DE BOX & JENKINS

Como citado anteriormente, o modelo de Box & Jenkins será utilizado para a etapa de previsão, representando o grupo de técnicas paramétricas. O mesmo foi desenvolvido utilizando o software Matlab™, e é fundamentado nos passos apresentados na Figura 5. Desta forma, é necessário avaliar cada passo e tomar a decisão de seguir em frente ou reajustar o modelo. As principais etapas para a previsão de séries temporais, de maneira genérica, e que estão contidas no processo de identificação de sistemas descrito na Figura 5, e que foram adotadas para modelar as séries que fazem parte desta pesquisa são (maiores detalhes no Apêndice B):

- ✓ Identificar se a série possui componentes periódicas (sazonalidade): através da aplicação da transformada de Fourier (FT) é feito análise dos picos do espectro, identificando a componente principal das harmônicas em destaque e se há sazonalidade;
- ✓ Remover a sazonalidade S através dos operador de atraso B, conforme Equação B.4;
- ✓ Identificar se a série possui tendência: através da análise gráfica;
- ✓ Remover a componente de tendência através do operador de diferenças $\nabla^d y_t = (1 - B)^d y_t$;
- ✓ Identificar qual modelo de Box & Jenkins representa melhor o comportamento da série, através da análise da função de auto correlação (*fac*) ou auto correlação parcial (*facp*) descrita na Equação B.5 ou B.8;
- ✓ Construir o modelo;
- ✓ Aplicar o modelo para obter as previsões de $1, \ldots, K$ passos-à-frente;
- ✓ Avaliar as previsões através dos indicadores de desempenho do modelo previsor (ver Seção 6.2).

A identificação de qual modelo de Box & Jenkins representa melhor o comportamento da série, consiste primariamente em determinar se o mecanismo formador da série possui comportamento que está associado à um modelo auto regressivo AR(p) e/ou

de média móvel MA(q), ou ambos. Ao analisar o comportamento da (*fac*) e da (*facp*), pode-se determinar a ordem q e p, tal como apresentado na Tabela 3.

Tabela 3: Análise da *fac* e *facp* para determinar o modelo de Box & Jenkins

Modelo	Função Auto Correlação (*fac*)	Função Auto Correlação Parcial (*facp*)
AR(p)	Decaimento exponencial ou seno amortecido	Tende próximo a zero após o *lag p*
MA(q)	Tende próximo a zero após o *lag q*	Decaimento exponencial ou seno amortecido
ARMA(p,q)	Decaimento exponencial ou seno amortecido	Decaimento exponencial ou seno amortecido

Fonte: O autor(2019).

Como para cada série de velocidades, ao aplicar o modelo de decomposição, resultou em seis componentes distintas, deve ajustar um modelo paramétrico para cada componente, ou seja, a etapa de previsão é realizada para cada componente. Os modelos em destaque são apresentados nas Tabelas 4 à 11. Cada um desses modelos representam o mecanismo gerador do comportamento da série. Os mesmos foram aplicados às respectivas componentes para obter as previsões ou novas realizações de 1 até 12 passos-à-frente. Os resultados finais da decomposição são apresentados e discutidos no próximo capítulo.

Tabela 4: Parâmetros das componentes da série velocidades de vento 01/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Componente	Parâmetro	Valor	Erro	Estatística t	BIC
C_1 ARIMA(1,0,0)	Constante	0,11519	0,017233	6,684	n.a
	AR{1}	0,98782	0,0019023	519,29	-1200,54
C_2 ARIMA(8,0,0)	Constante	-3,378e-08	0,29126	-1,16e-07	n.a
	AR{8}	-441,05e-03	481,3e-3	-916,37e-3	-1938,53
C_3 ARIMA(13,0,0)	Constante	6,0846e-07	0,059451	1,0235e-05	n.a
	AR{13}	-0,83037	0,049218	-16,871	-1940,54
C_4 ARIMA(13,0,0)	Constante	4,2448e-08	0,12091	3,5107e-07	n.a
	AR{13}	-0,51027	0,78327	-0,65146	-1928,53
C_5 ARIMA(5,1,0)	Constante	0,000873	0,0081422	0,10722	n.a
	AR{5}	0,19875	0,093478	2,1262	0,033491
X_r ARIMA(4,1,0)	Constante	-2,4972e-03	15,604e-03	-160,033-03	n.a
	AR{4}	-31,023e-2	68,015e-03	-4,5607	-57,14

Fonte: O autor(2019).

Tabela 5: Parâmetros das componentes da série velocidades de vento 02/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Componente	Parâmetro	Valor	Erro	Estatística t	BIC
C_1 ARIMA(0,1,1)	Constante	-0,00183714	5,23144e-05	-35,1173	n.a
	MA{1}	0,0164453	0,207691	0,0791815	-1913,03
C_2 ARIMA(9,0,0)	Constante	1,70554e-15	2,85475e-15	-1,0881e-05	n.a
	AR{9}	-0,0799357	0,00604171	-13,2306	-1936,52
C_3 ARIMA(13,0,1)	Constante	1,27007e-08	0,0166329	7,63588e-07	n.a
	AR{13}	0,420626	0,17488	-2,40523	-1926,53
	MA{1}	-0,265834	6,64717e-07	-399920	-1910,53
C_4 ARIMA(10,0,1)	Constante	-2,95641e-09	0,0838213	-3,52704e-08	n.a
	AR{10}	0,0758664	3,11207e-05	2437,81	-1932,53
	MA{1}	-0,642437	5,84548e-10	-1,09903e+09	-1928,53
C_5 ARIMA(6,1,6)	Constante	-0,00122225	0,00114128	-1,07095	n.a
	AR{6}	-0,793522	0,0348603	10,34	-631,571
	MA{6}	0,000912399	0,089057	0,0102451	-769,23
X_r ARIMA(6,1,1)	Constante	0,0271371	0,0340655	0,796615	n.a
	AR{6}	0,0874337	0,104881	0,833647	2,448
	MA{1}	0,586266	0,315976	1,85541	-0,0377

Fonte: O autor(2019).

Tabela 6: Parâmetros das componentes da série velocidades de vento 03/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Componente	Parâmetro	Valor	Erro	Estatística t	BIC
	Constante	0,000371773	0,00856287	0,0434169	n.a
C_1 ARIMA(12,0,1)	AR{12}	-0,238971	0,720345	-0,331745	-1352,16
	MA{1}	0,870387	0,977622	0,89031	-1920,94
	Constante	-4,93301e-08	0,0142765	-3,45533e-06	n.a
C_2 ARIMA(14,0,1)	AR{14}	0,132933	0,0685412	1,93947	-1924,53
	MA{1}	-0,343274	4,37238e-07	-785097	-1898,53
	Constante	-2,45775e-08	0,0057766	-4,25466e-06	n.a
C_3 ARIMA(14,0,1)	AR{14 }	0,713745	0,00285024	250,416	-1924,53
	MA{1}	0,900573	1,05146e-07	8,56499e+06	-806,05
	Constante	1,10927e-07	0,00451492	2,45689e-05	n.a
C_4 ARIMA(14,0,1)	AR{ 14 }	-0,52894	0,00969186	-54,5756	-1924,51
	MA{1}	0,520424	5,38104e-07	967143	-1910,53
	Constante	0,000619296	0,00207522	0,298424	n.a
C_5 ARIMA(14,2,1)	AR{14}	0,254044	0,205331	1,23724	-625,89
	MA{1}	0,0245826	0,514938	0,0477389	-739,25
	Constante	-0,0869547	0,0207095	-4,19878	n.a
X_r ARIMA(1,0,1)	AR{1}	0,371043	0,0752147	4,93312	-63,017
	MA{1}	0,247415	0,212728	1,16306	-60,10

Fonte: O autor(2019).

Tabela 7: Parâmetros das componentes da série velocidades de vento 04/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Componente	Parâmetro	Valor	Erro	Estatística t	BIC
C_1 ARIMA(15,0,1)	Constante	0,00123959	0,00590268	0,21004	n.a
	AR{15}	0,061868	0,93437	0,0662137	-1066,66
	MA{1}	0,925099	1,1628	0,795577	-1912,76
C_2 ARIMA(13,0,0)	Constante	4,29128e-08	0,0117284	3,65889e-06	n.a
	AR{13}	-0,048557	0,0723163	-0,671453	-1928,53
C_3 ARIMA(18,0,1)	Constante	2,14474e-08	0,03552	6,03746e-07	n.a
	AR{18}	0,00293793	0,108034	0,0271944	-1916,53
	MA{1}	0,979033	3,44953e-05	28381,6	-1892,45
C_4 ARIMA(12,0,1)	Constante	2,08819e-06	0,007425	0,0002812	n.a
	AR{12}	-0,31163	0,011874	-26,2447	1928,52
	MA{1}	0,745126	1,41033e-06	528336	-898,56
	MA{1}	757,34e-03	6,0269e-07	1,2566e+06	-1912,53
C_5 ARIMA(12,1,1)	Constante	0,00053252	0,0016417	0,32437	n.a
	AR{12})	0,46788	0,057879	8,0837	-1892,45
	MA{1}	0,63468	0,1906	3,33	-1916,53
X_r ARIMA(1,1,1)	Constante	-0,00050796	0,0013984	-0,36324	n.a
	AR{1}	0,47477	0,097125	4,8883	-73,0183
	MA{1}	-0,91348	0,042823	-21,332	-71,0254

Fonte: O autor(2019).

Tabela 8: Parâmetros das componentes da série de velocidades de vento 05/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Componente	Parâmetro	Valor	Erro	Estatística t	BIC
C_1 ARIMA(1,7,1)	Constante	1,36342e-15	2,93051e-15	-0,465251	n.a
	AR{1}	-0,409004	2,54938e-10	-1,60433e+09	-8506,79
	MA{1}	-0,252629	5,53963e-17	-4,56039e+15	-8386,95
C_2 ARIMA(15,0,1)	Constante	9,43885e-08	0,0229057	-4,12074e-06	n.a
	AR{15}	-0,285014	0,174452	-1,63376	-1922,53
	MA{1}	0,936182	5,10534e-05	18337,3	-1369,24
C_3 ARIMA(14,0,1)	Constante	-1,10174e-07	0,589144	-1,87007e-07	n.a
	AR{14}	0,550623	0,527118	1,04459	-1369,24
	MA{1}	0,980056	2,32718e-06	421135	-1026,26
C_4 ARIMA(15,0,8)	Constante	7,14897e-09	1,57784e-14	453087	n.a
	AR{15}	0,169894	1,21596e-17	1,3972e+16	-1026,26
	MA{8}	-0,125584	7,91193e-23	-1,58727e+21	-1908,5
C_5 ARIMA(18,0,9)	Constante	-0,0065917	0,012611	-0,52268	n.a
	AR{18}	-0,19313	0,33217	0,58142	-664,870
	MA{9}	-0,31682	0,77039	-0,41124	-960,23
X_r ARIMA(1,0,6)	Constante	0,039085	0,028408	1,3758	n.a
	AR{1}	0,61104	0,19405	3,1489	-23,10
	MA{6}	0,37482	0,085933	4,3618	-13,572

Fonte: O autor(2019).

Tabela 9: Parâmetros das componentes da série de velocidades de vento para o trimestre 01/nov/2018 a 31/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Componente	Parâmetro	Valor	Erro	Estatística t	BIC
C_1 ARIMA(13,0,1)	Constante	0,0014278	0,028332	0,050395	n.a
	AR{13}	0,18949	0,75757	0,25012	-1183,52
	MA{1}	0,71056	1,1648	0,61004	-1147,07
C_2 ARIMA(18,0,1)	Constante	1,4055e-07	0,026781	5,2481e-06	n.a
	AR{18}	0,38614	0,041825	9,2322	-1210,01
	MA{1}	0,34527	8,1692e-07	4,2265e+05	-535,86
C_3 ARIMA(15,0,1)	Constante	1,2459e-07	0,050589	2,4628e-06	n.a
	AR{15}	- 0,29051	0,053377	5,4427	-1216,01
	MA{1}	0,98784	2,1797e-06	4,532e+05	-479,32
C_4 ARIMA(14,0,1)	Constante	1,676e-07	0,023768	7,0516e-06	n.a
	AR{14}	0,1099	0,061345	1,7916	-123,76
	MA{1}	-0,19913	8,5107e-08	-2,3397e+06	-439,31
C_5 ARIMA(16,0,14)	Constante	0,11293	0,156	0,72392	n.a
	AR{16}	-0,37722	0,23893	-1,5788	142,03
	MA{14}	-0,54621	0,26568	-2,0559	132,66
C_6 ARIMA(1,1,1)	Constante	0,0010029	0,0010143	0,98871	n.a
	AR{1}	-0,056756	0,1095	-0,51831	-113,56
	MA{1}	-0,93315	0,038525	-24,222	-112,336

Fonte: O autor(2019).

Tabela 10: Parâmetros da série de velocidades de vento para os dias 01 a 05 de janeiro de 2019

Distribuição de Probabilidade Condicional: Gaussiana					
Série	Parâmetro	Valor	Erro	Estatística t	BIC
Dia 01 - ARIMA(1,1,16)	Constante	0,015101	0,006636	2,2756	n.a
	AR{1}	-0,50164	0,52279	-0,95954	54,340
	MA{16}	-0,095492	0,12202	-0,78257	56,23
Dia 02 - ARIMA(1,1,4)	Constante	-0,0024524	0,019673	-0,12465	n.a
	AR{1}	0,42542	0,20485	2,0767	71,176
	MA{4}	0,21747	0,11746	1,8515	70,45
Dia 03 - ARIMA(1,1,1)	Constante	-0,024538	0,024737	-0,99197	n.a
	AR{1}	-0,14272	0,42706	-0,33418	50,59
	MA{1}	-0,014895	0,43389	-0,03433	55,28
Dia 04 - ARIMA(1,1,1)	Constante	0,011762	0,0295857	0,39757	n.a
	AR{1}	0,084871	1,2669	0,066993	27,19
	MA{1}	-0,014117	1,274	-0,01108	33,25
Dia 05 - ARIMA(16,1,1)	Constante	0,010782	0,039849	0,27058	n.a
	AR{16}	0,1286	0,14374	0,89465	82,93
	MA{1}	0,58248	0,31255	1,8637	75,337

Fonte: O autor(2019).

Tabela 11: Parâmetros para a série de velocidades de vento do trimestre 01/nov/2018 a 31/jan/2019

Distribuição de Probabilidade Condicional: Gaussiana					
Série	Parâmetro	Valor	Erro	Estatística t	BIC
Trimestre - ARIMA(5,1,1)	Constante	-0,00087747	0,0046476	-0,1888	n.a
	AR{5}	-0,86151	0,077014	-11,186	191,34
	MA{1}	0,12634	0,13061	0,96731	182,29

Fonte: O autor(2019).

Pode-se observar que, ao comprar os modelos ARIMA obtidos para as séries sem a etapa de decomposição nas Tabelas 10 e 11, com os obtidos para as componentes, a ordem dos modelos para as componentes individuais são inferiores, além de possuir menor número de parâmetros. Ou seja, para a série original, o modelo obtido possui maior complexidade quando comparado aos modelos obtidos para cada componente.

4.1.3 PARÂMETROS DA REDE LSTM

Além da técnica paramétrica ARIMA, será utilizado também a rede neural recorrente LSTM (maiores detalhes ver Apêndice D) que foi desenvolvida utilizando o software Matlab™. Esta rede será aplicada para obter previsões utilizando a estratégia recursiva (Apêndice A.1), e tem o conjunto de parâmetros descritos na Tabela 12, que foram definidos a *priori*, seguindo o padrão encontrado na maioria dos trabalhos que aplicam tal técnica (SALAZAR, 2015; LI et al., 2019). Pode-se observar que a rede LSTM utiliza um otimizador Adam para ajustar os pesos da rede, esta sigla é a abreviatura de estimativa adaptativa do momento (do inglês, *adaptive moment estimation*), cujo algoritmo e maiores detalhes é apresentado no Apêndice E.

A rede LSTM será aplicada para obter previsões recorrentes de 1 passo-à-frente até o limite $K = 10$ para a série temporal de velocidades diárias, observadas no trimestre de 01/nov/2018 a 31/jan/2019. Já para as previsões dos intervalos de 10min, a mesma será utilizada para prever o horizonte de 12 passos-à-frente. A configuração de seus parâmetros foi mantida fixa para todas as previsões das componentes.

A divisão da série temporal para a etapa de treinamento e validação da rede neural LSTM considerou 85% das amostras para treinamento e 15% validação. A validação ocorreu considerando sempre a avaliação do resultado do erro da previsão para um passo-à-frente.

Tabela 12: Parâmetros Rede LSTM

Parâmetros Rede LSTM	
Parâmetro	Valor
Otimizador	Adam
Máximo N° de épocas	1500
Limiar do gradiente	1
Taxa inicial de aprendizado	0,009
Período da taxa de queda de aprendizado	125
Fator de queda da taxa de aprendizado	0,2
Tamanho de cada mini lote	125
Total de camadas escondidas	200

Fonte: O autor(2019).

4.2 RESUMO DO CAPÍTULO

Neste capítulo foi apresentado os conjuntos de dados que serão utilizados junto ao modelo proposto para obter a decomposição, e posteriormente as previsões de 1 a K passos-à-frente. Foi abordado também a escolha e definição dos parâmetros para cada técnica que compõe o modelo híbrido, além de discutir os parâmetros adotados para as técnicas de previsão. Neste trabalho os modelos de previsão adotados foram o modelo paramétrico ARIMA de Box & Jenkins (cujo conceitos estão detalhados no APÊNDICE B), e também um modelo não paramétrico, representado pela rede neural LSTM. Os modelos ARIMA obtidos para as séries sem a etapa de decomposição, ou seja, para a série original, possuem maior complexidade quando comparados aos modelos obtidos para cada componente resultante da decomposição. Os resultados da aplicação da técnica proposta serão discutidos no próximo capítulo.

5 AVALIAÇÃO DO DESEMPENHO DO MODELO HÍBRIDO DE DECOMPOSIÇÃO

Este capítulo faz uma síntese dos resultados obtidos ao aplicar o modelo híbrido proposto na decomposição de séries temporais de velocidades de vento. Cada conjunto de modos obtidos como resultado da decomposição, é apresentado juntamente com a matriz de correlação entre esses modos, facilitando a análise das correlações entre cada componente obtida após a aplicação do modelo híbrido de decomposição proposto.

5.1 RESULTADOS DA DECOMPOSIÇÃO

Ao aplicar o VMD-SSA para obter a decomposição das séries de velocidades de vento descritas anteriormente, obtemos para cada série seis componentes. Estas componentes bem como as séries originais e a matriz de correlação é apresentada nas Figuras 39 à 88. Para evitar tornar extensa essa seção, não serão apresentados todo os resultados da decomposição para os 31 ou 30 dias de cada mês. Os resultados da decomposição são apresentados para 5 dias consecutivos de cada mês, além do resultado para os dias 10, 20 e 30 de cada mês; e por fim o resultado da decomposição para a série do trimestre, que tem seus dados amostrados em base diária e que contém 92 dias.

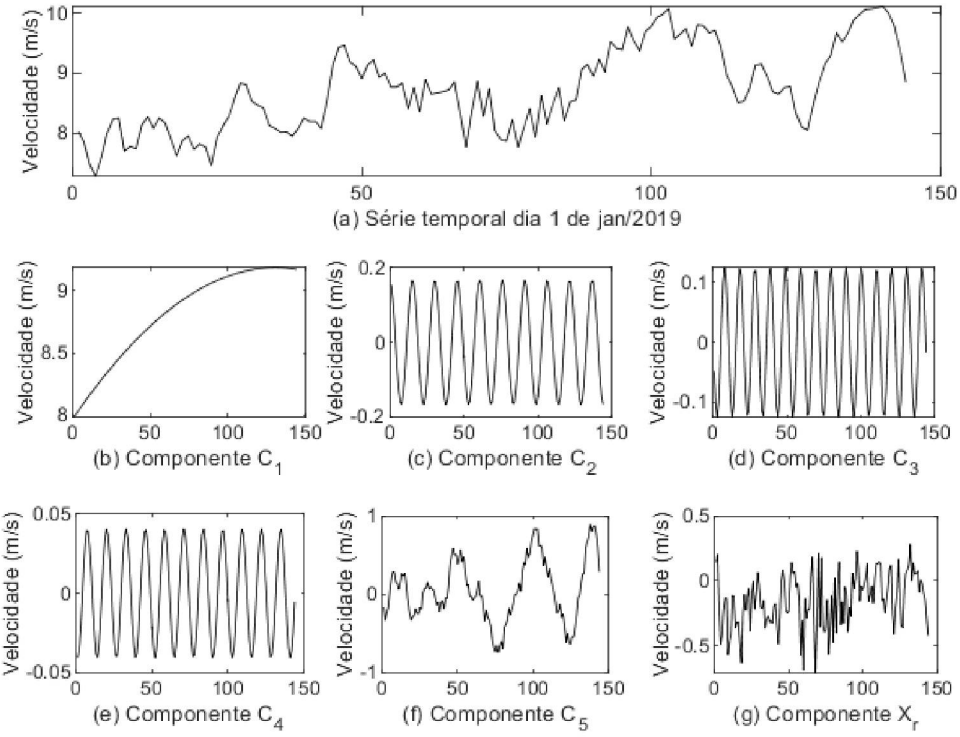
Desta forma, o conjunto de resultados será composto pelos dias 1 a 5 de cada mês; além dos múltiplos de 10 e os resultados para os 92 dias do trimestre que compreende os meses de novembro de 2018 a janeiro de 2019. Com essa base de dados, onde esta presente alta sazonalidade e variabilidade, espera-se poder comprovar a eficiência do modelo proposto, ao aplicá-lo para decomposição da série de velocidades de vento. A decomposição para ser útil, deve representar a dinâmica do mecanismo gerador da série, ou seja, deve após a etapa de reconstrução do sinal, garantir à preservação do comportamento da série original, que estará representado de forma isolado em cada uma das componentes, e que

a priori, devem estar descorrelacionadas ou possuírem fraca correlação, como poderá ser observado no desenvolvimento deste capitulo.

Os modos reconstruídos no domínio do tempo são apresentados nas Figuras 39 à 87. Pode-se perceber que o modo C_1 representado na Figura 39(b) é a componente de tendência da série de velocidades de vento; Os modos C_2 à C_4 representados nas Figuras 39(c) à 39(e) são as componente oscilatórias (ou harmônicas) da série; Já os modos C_5 e X_r possuem ainda resíduos de outros modos de alta frequência. Relembrando que nesta etapa do algoritmo não é descartada nenhuma componente, as mesmas serão então processadas pelo SSA, e só então na etapa de agrupamento do SSA que uma subcomponente da componente de alta frequência X_r é descartada como ruído. Essa interpretação se estende a todas as outras figuras.

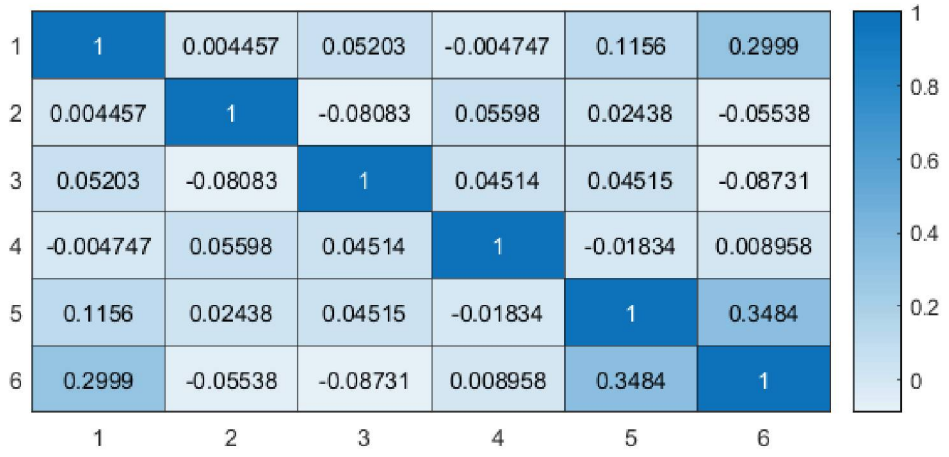
Já a Figura 40 apresenta a matriz de correlações das componentes. Analisando cada matriz, pode-se inferir que o modelo híbrido consegue ser eficiente, obtendo componentes praticamente descorrelacionadas. Ainda analisando cada componente obtida, pode-se notar que se torna mais simples o ajuste de um modelo para obter as previsões, dado que o comportamento de cada componente, na maiores das vezes se mantém inalterado ao longo do desenvolvimento da série. Exceções são as componentes C_5 e X_r, pois estas possuem um nível maior ruído incorporado.

Figura 39: Decomposição VMD-SSA da série temporal velocidades de vento (10min) de 01/jan/2019



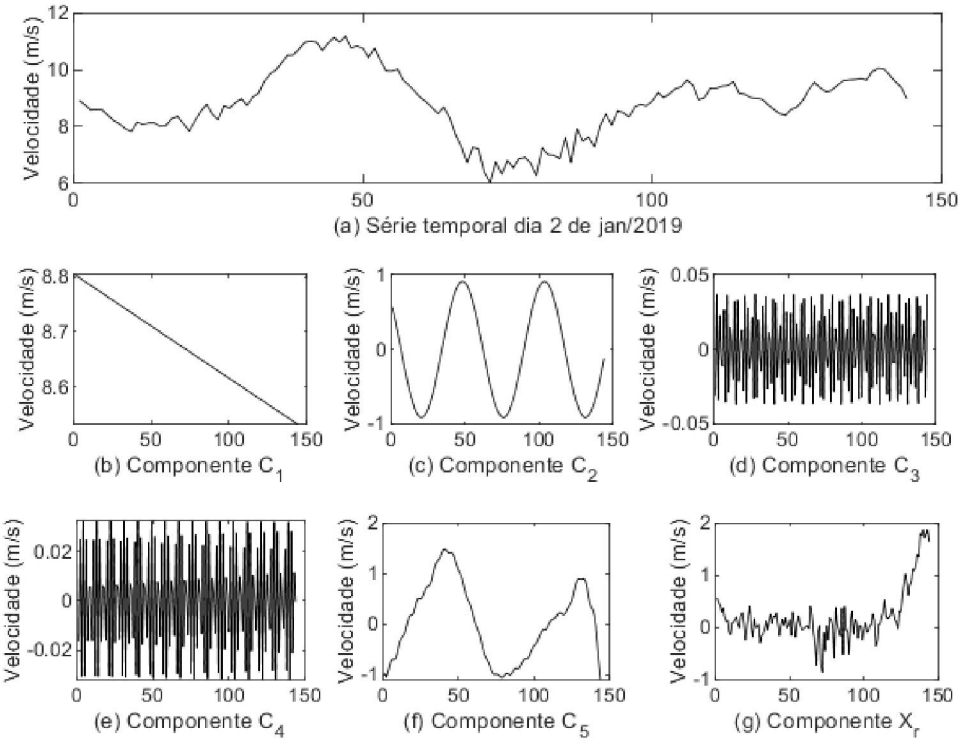
Fonte: O autor(2019).

Figura 40: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/jan/2019



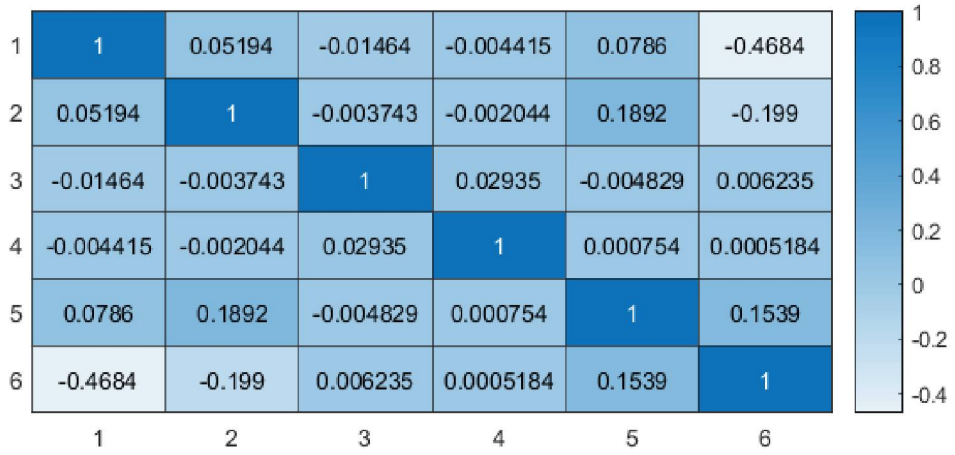
Fonte: O autor(2019).

Figura 41: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/jan/2019



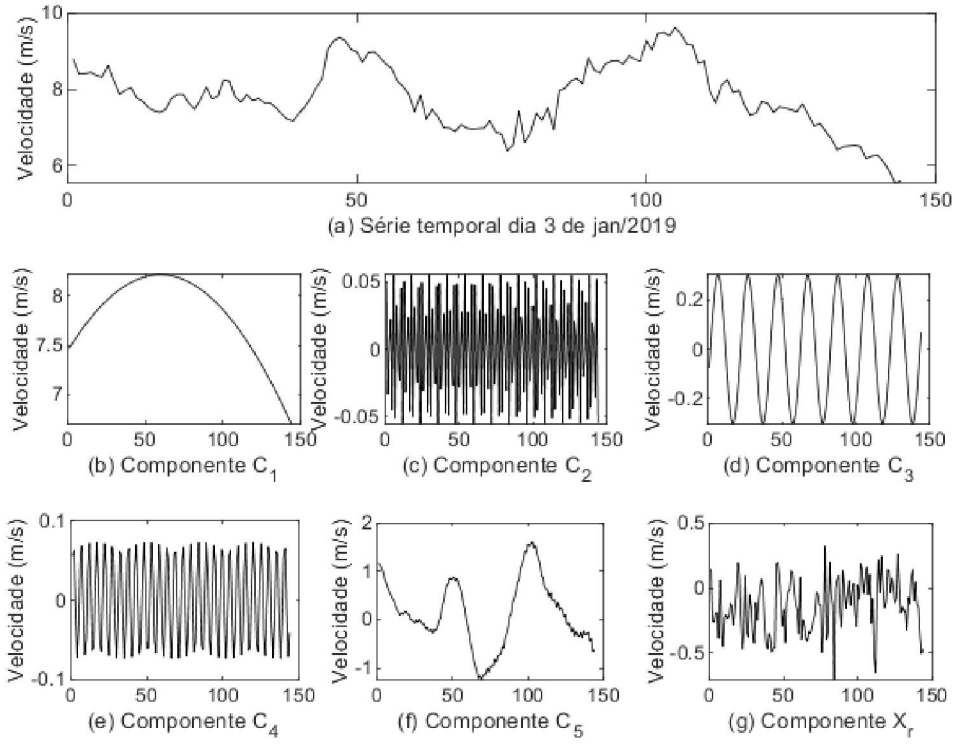
Fonte: O autor(2019).

Figura 42: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/jan/2019



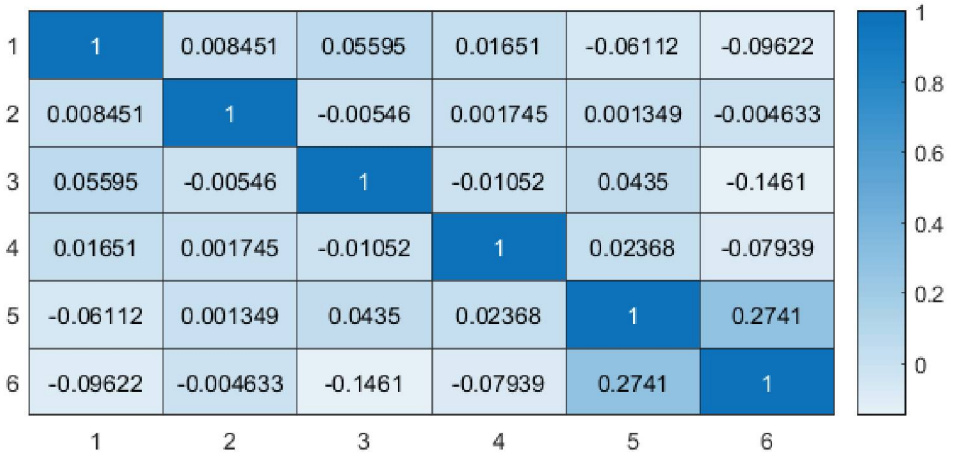
Fonte: O autor(2019).

Figura 43: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/jan/2019



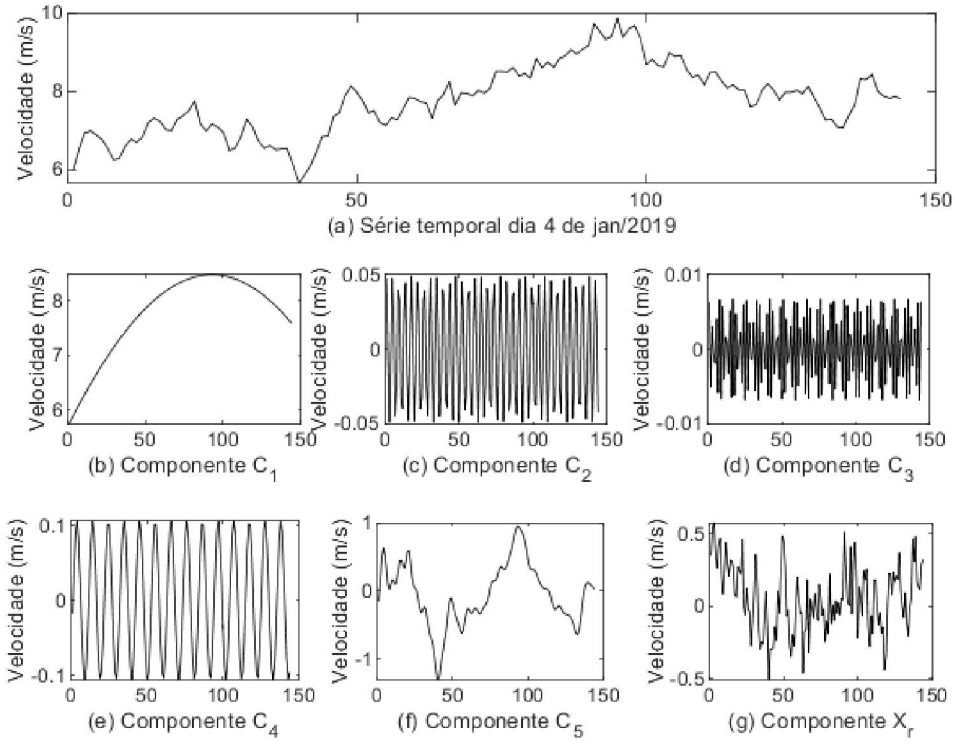
Fonte: O autor(2019).

Figura 44: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/jan/2019



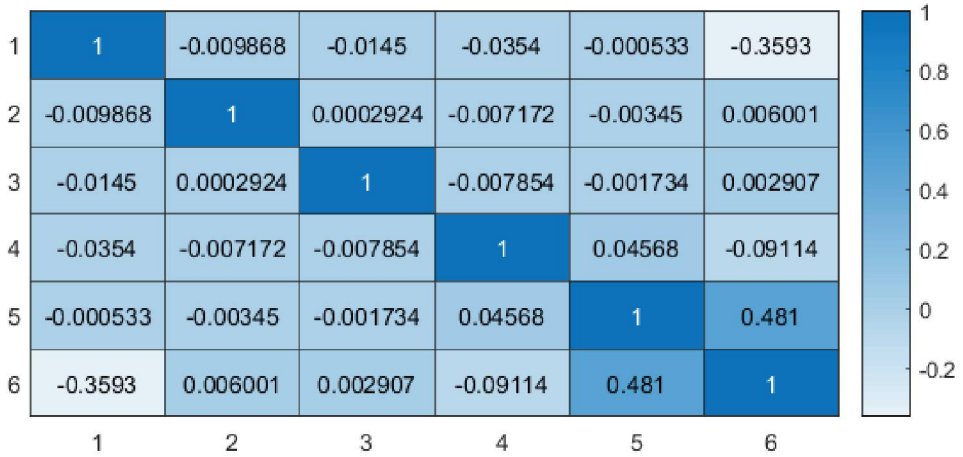
Fonte: O autor(2019).

Figura 45: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/jan/2019



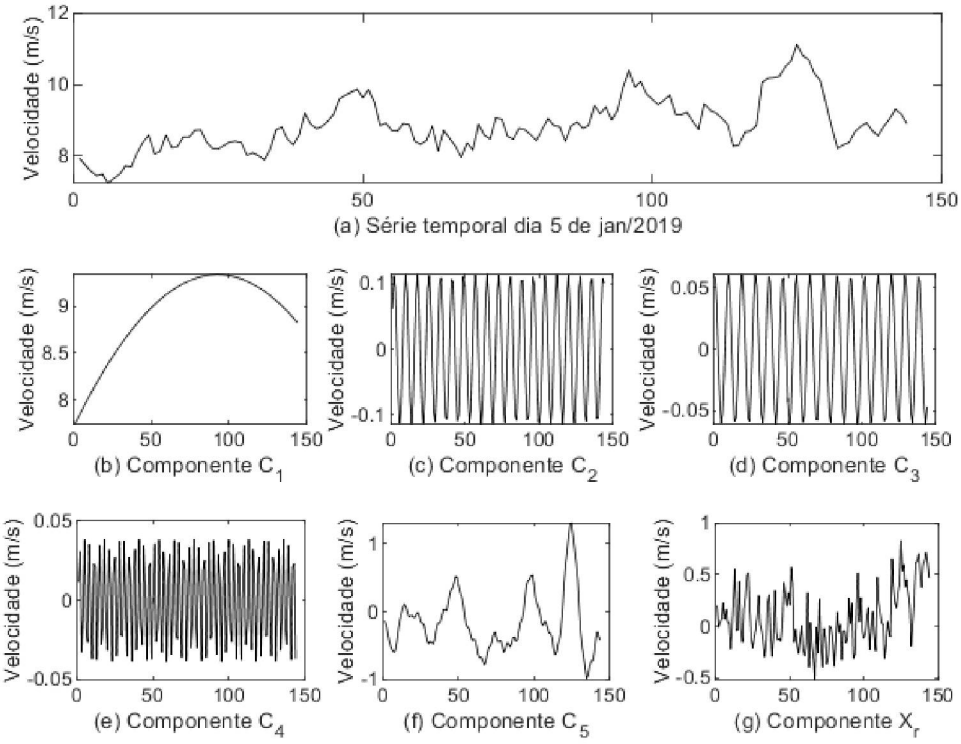
Fonte: O autor(2019).

Figura 46: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/jan/2019



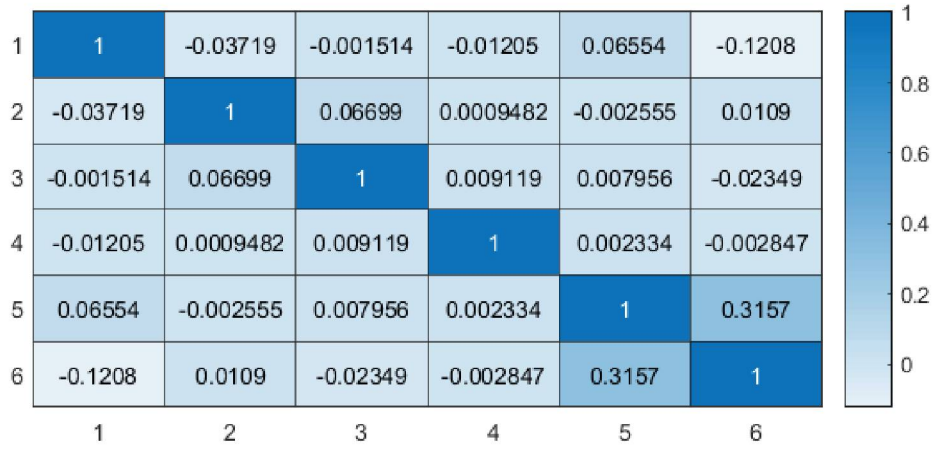
Fonte: O autor(2019).

Figura 47: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/jan/2019



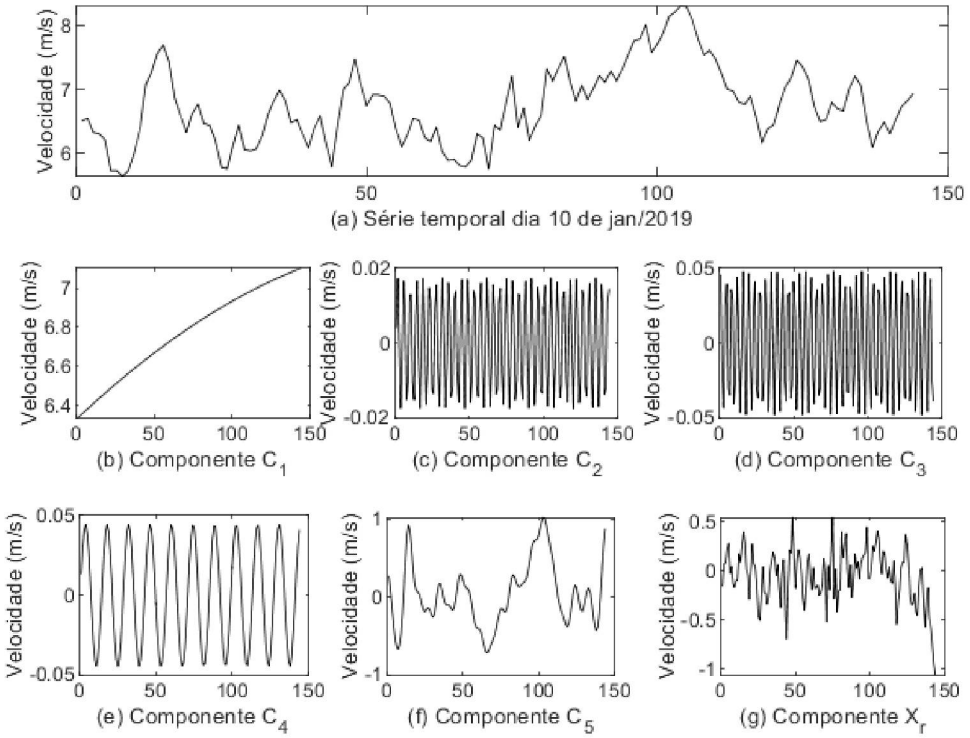
Fonte: O autor(2019).

Figura 48: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/jan/2019



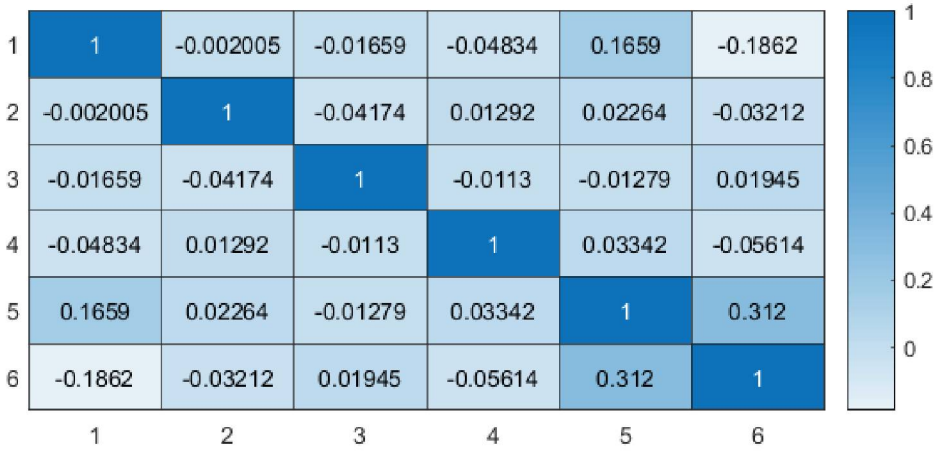
Fonte: O autor(2019).

Figura 49: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/jan/2019



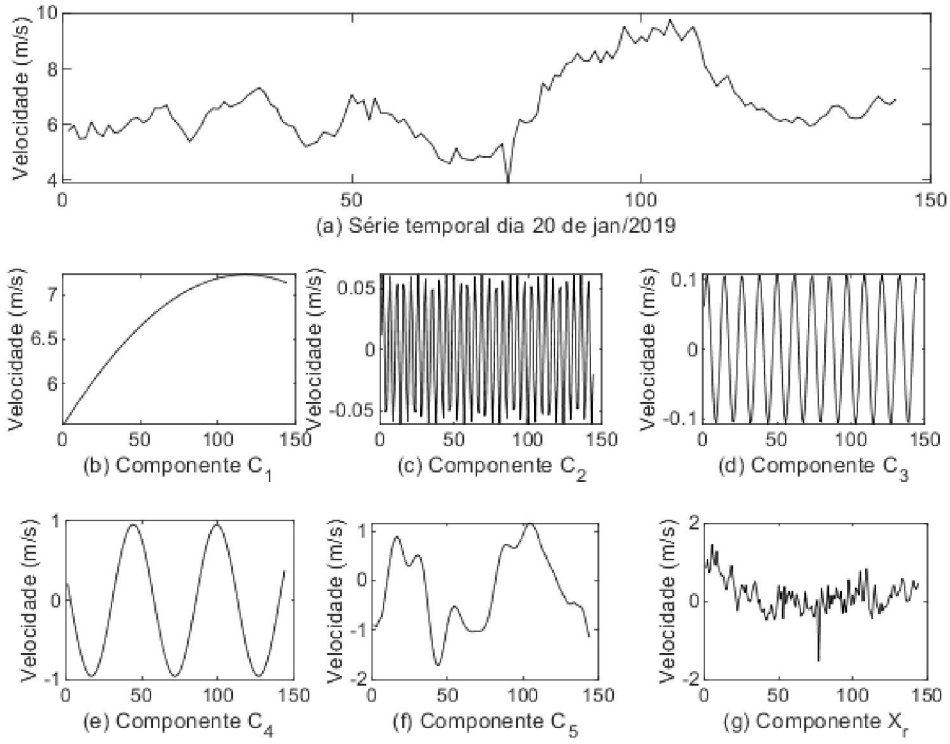
Fonte: O autor(2019).

Figura 50: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/jan/2019



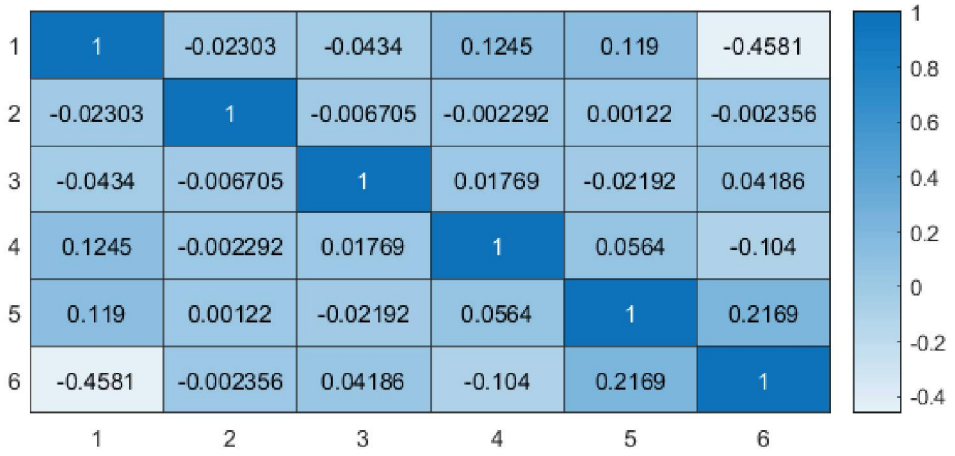
Fonte: O autor(2019).

Figura 51: Decomposição VMD-SSA da série temporal velocidades de vento (10min) de 20/jan/2019



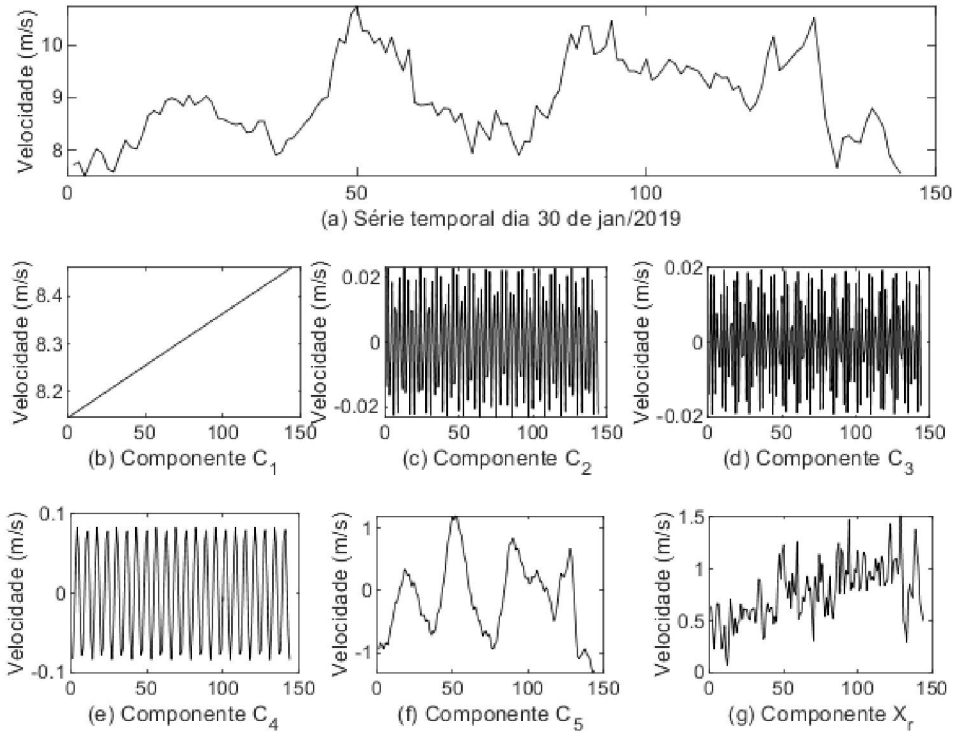
Fonte: O autor(2019).

Figura 52: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/jan/2019



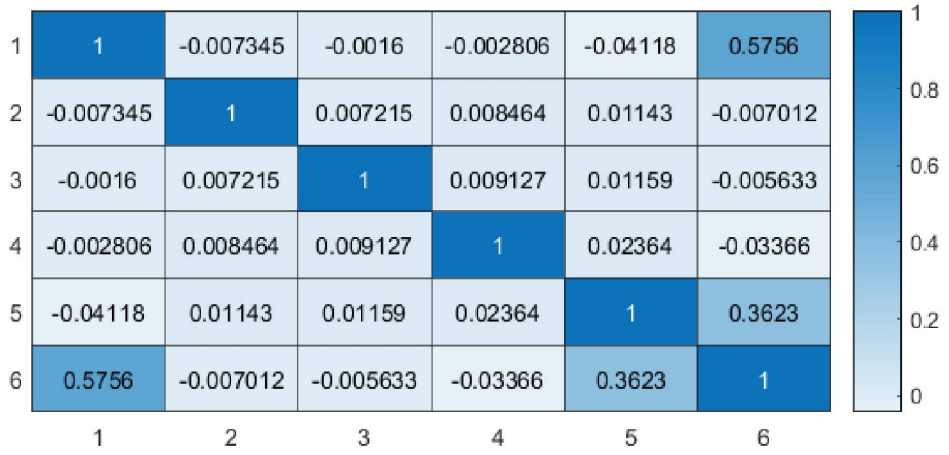
Fonte: O autor(2019).

Figura 53: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/jan/2019



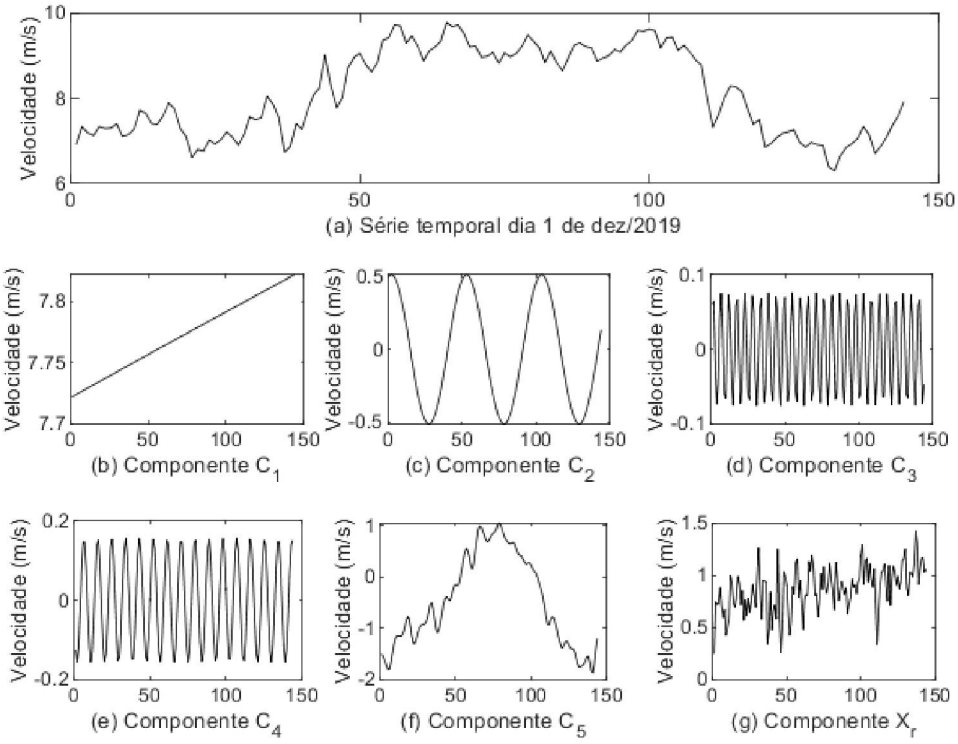
Fonte: O autor(2019).

Figura 54: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/jan/2019



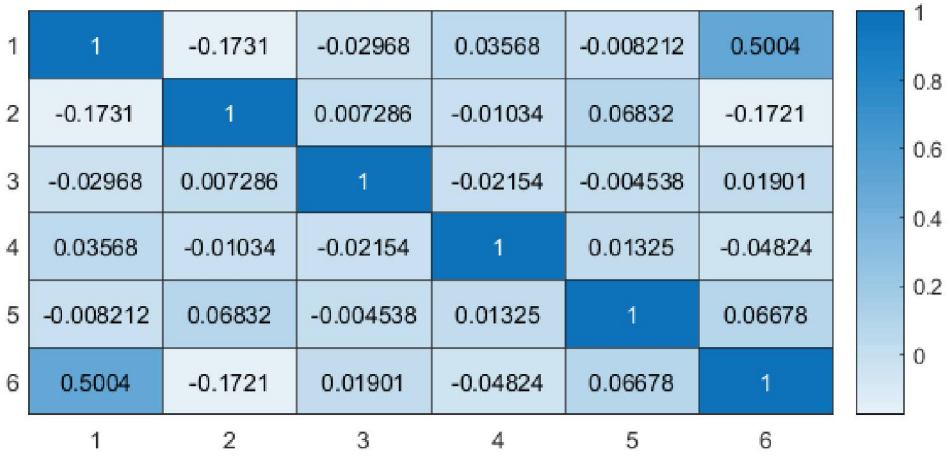
Fonte: O autor(2019).

Figura 55: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/dez/2018



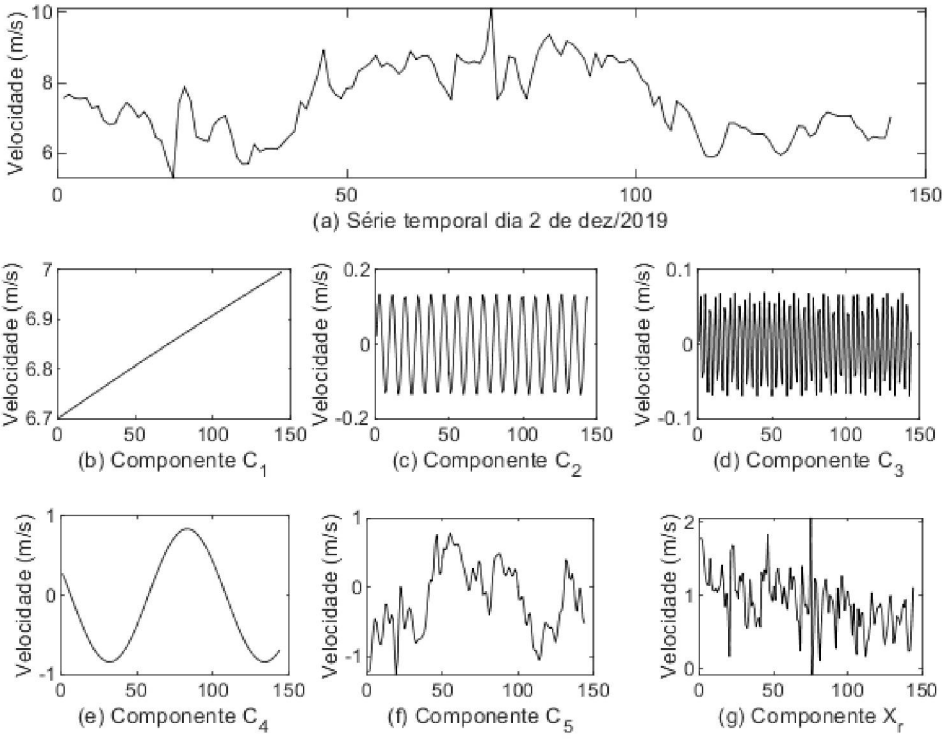
Fonte: O autor(2019).

Figura 56: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/dez/2018



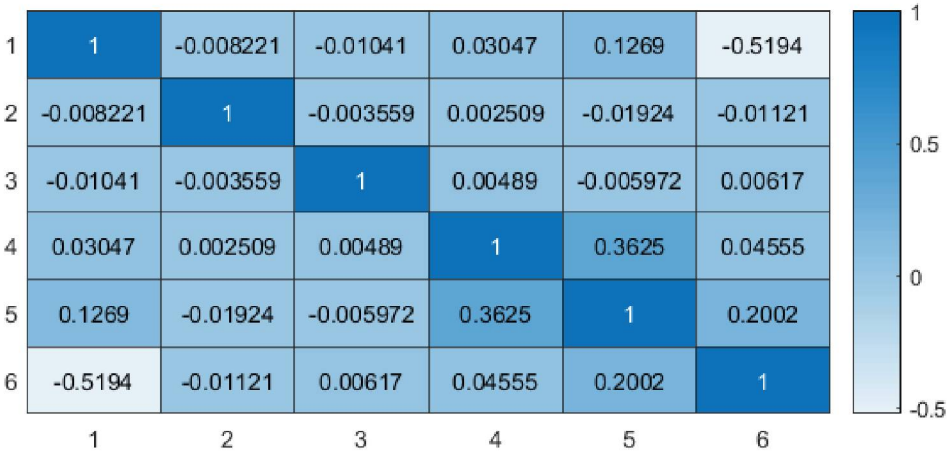
Fonte: O autor(2019).

Figura 57: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/dez/2018



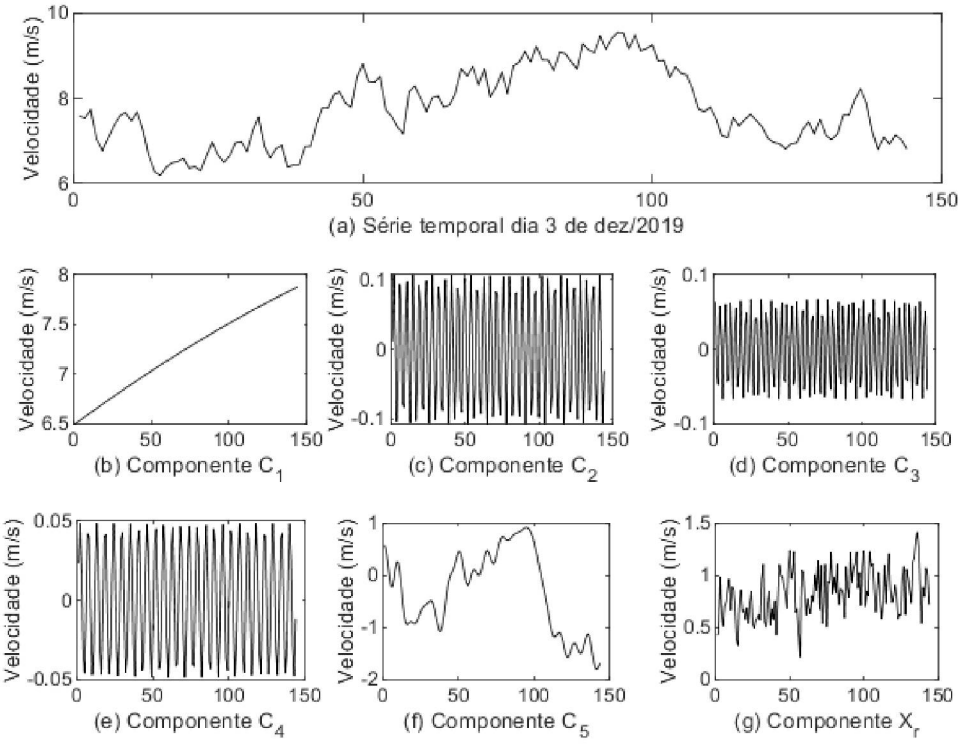
Fonte: O autor(2019).

Figura 58: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/dez/2018



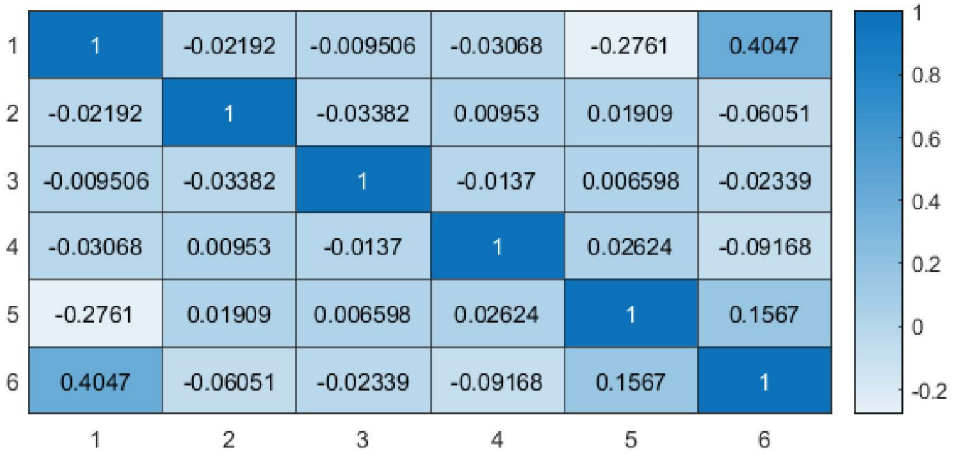
Fonte: O autor(2019).

Figura 59: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/dez/2018



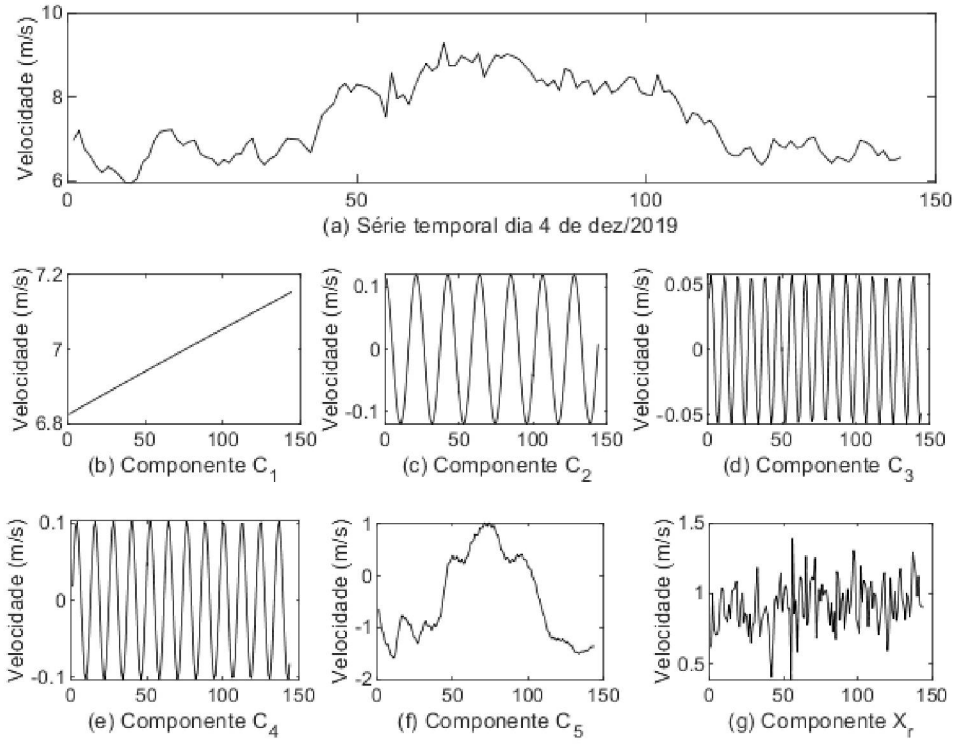
Fonte: O autor(2019).

Figura 60: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/dez/2018



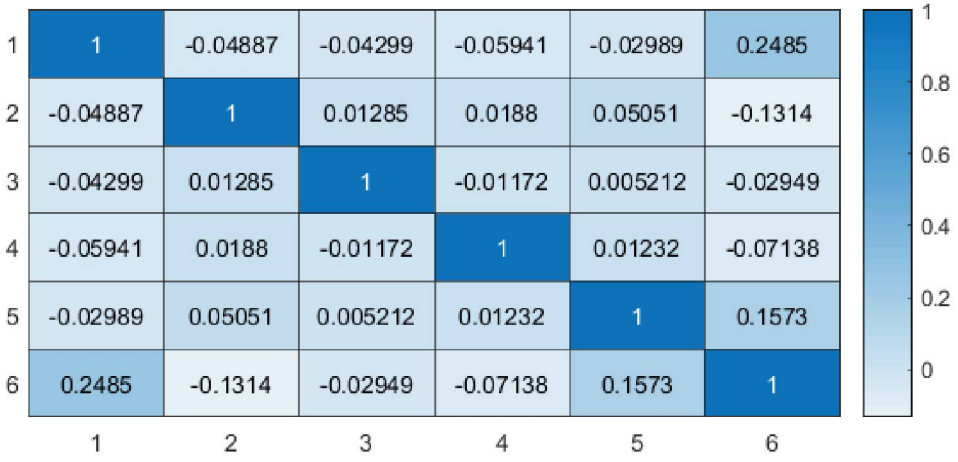
Fonte: O autor(2019).

Figura 61: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/dez/2018



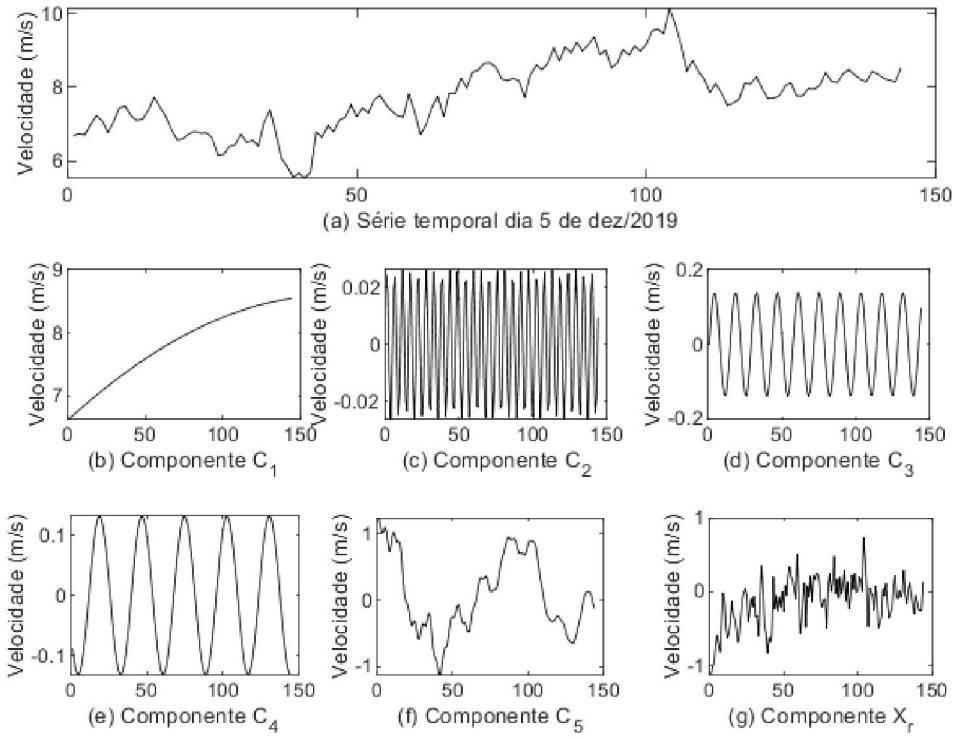
Fonte: O autor(2019).

Figura 62: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/dez/2018



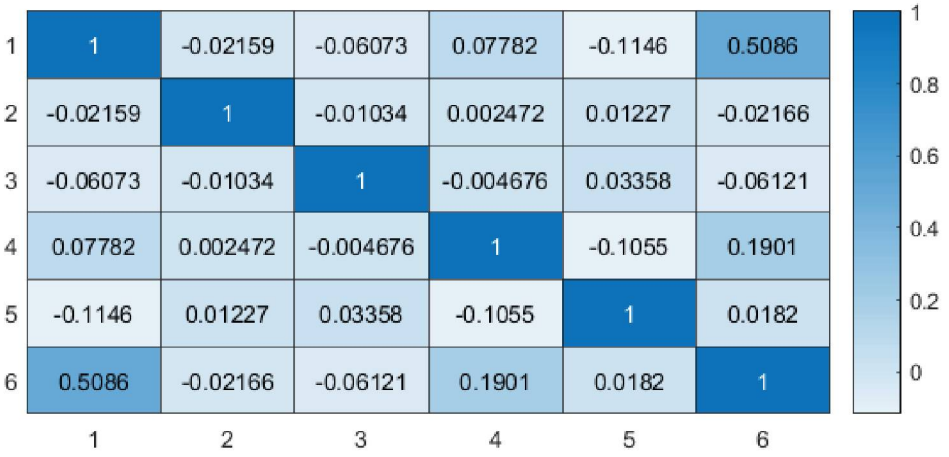
Fonte: O autor(2019).

Figura 63: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/dez/2018



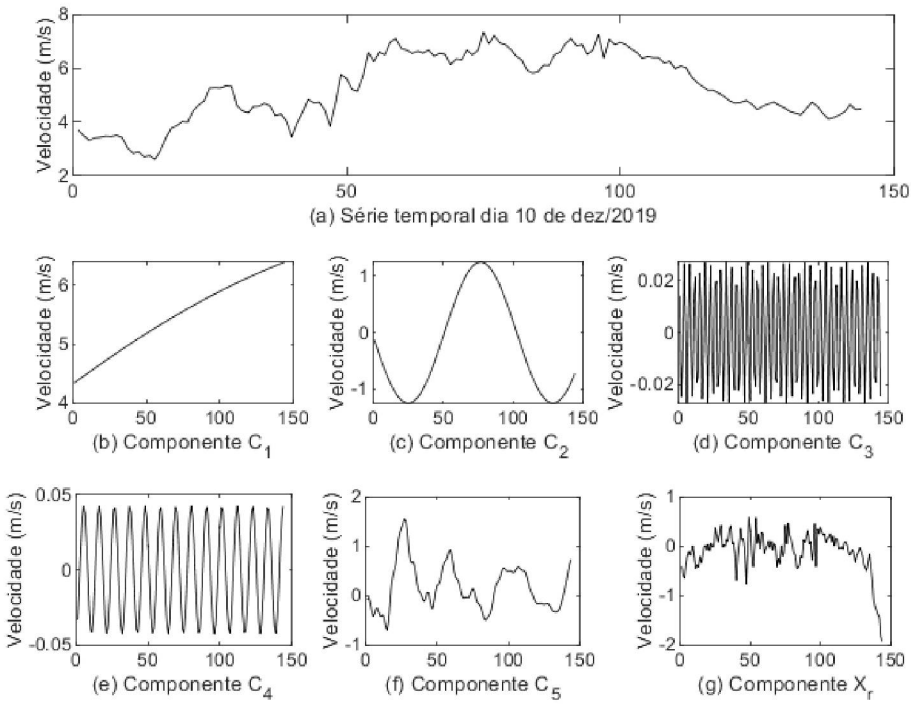
Fonte: O autor(2019).

Figura 64: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/dez/2018



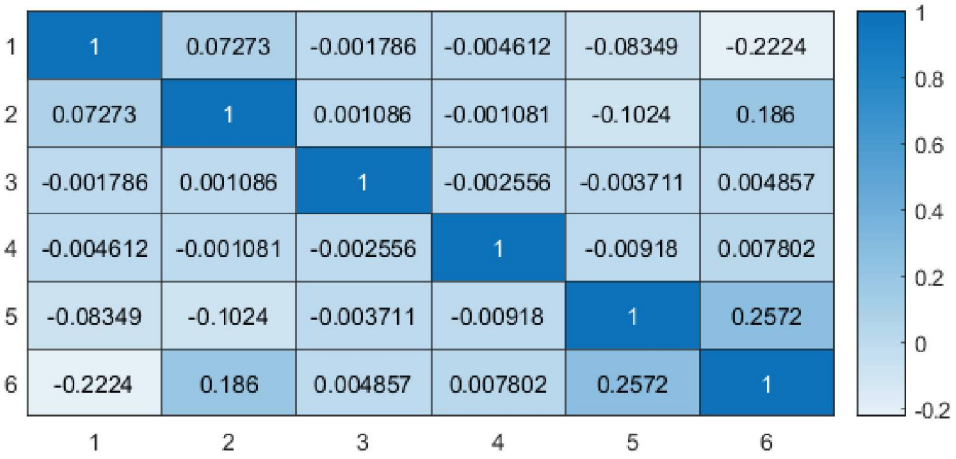
Fonte: O autor(2019).

Figura 65: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/dez/2018



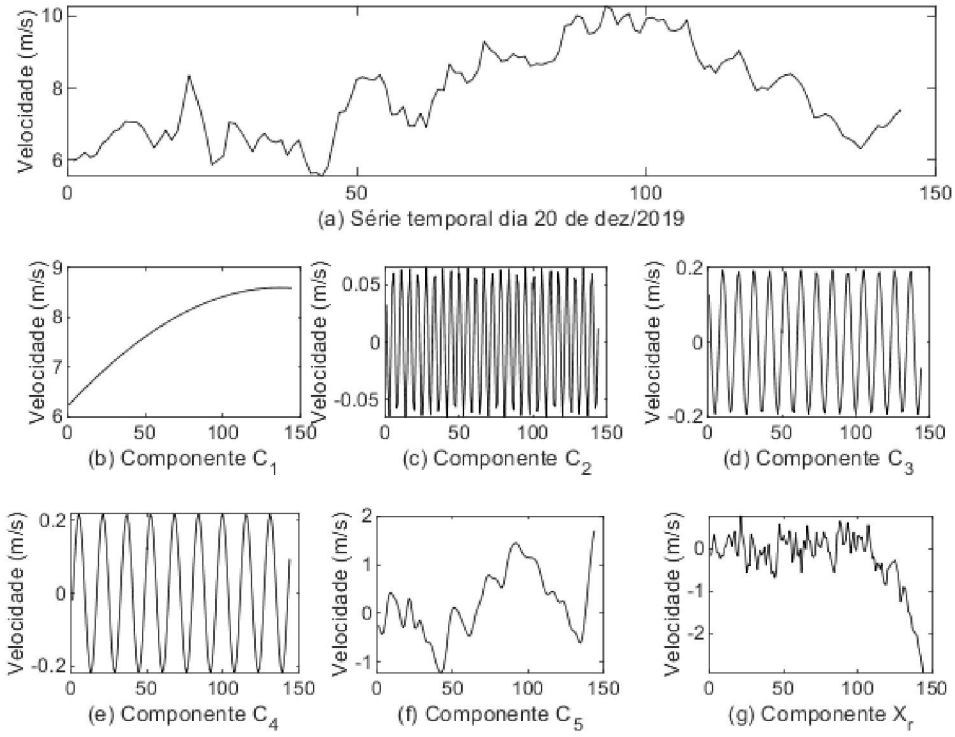
Fonte: O autor(2019).

Figura 66: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/dez/2018



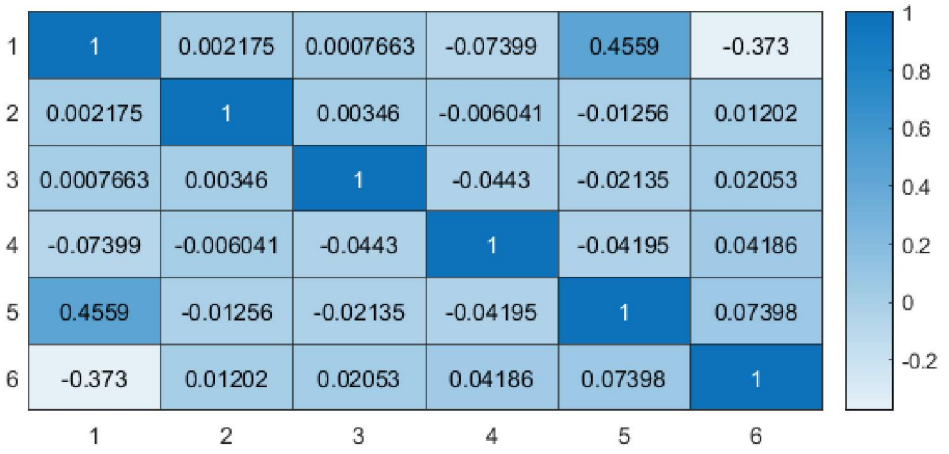
Fonte: O autor(2019).

Figura 67: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/dez/2018



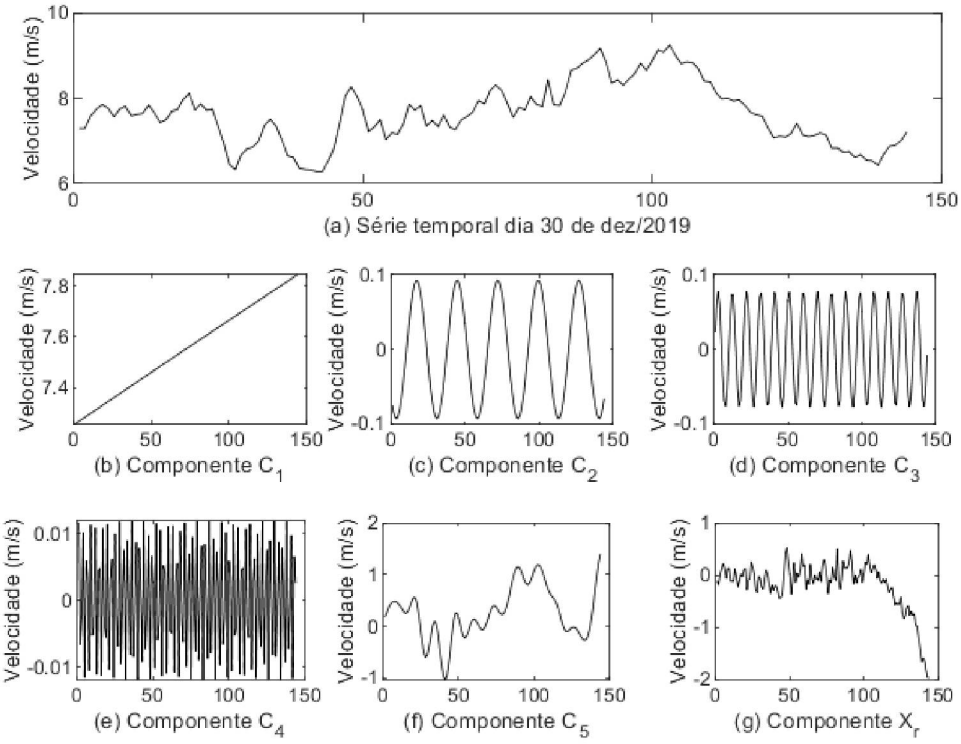
Fonte: O autor(2019).

Figura 68: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/dez/2018



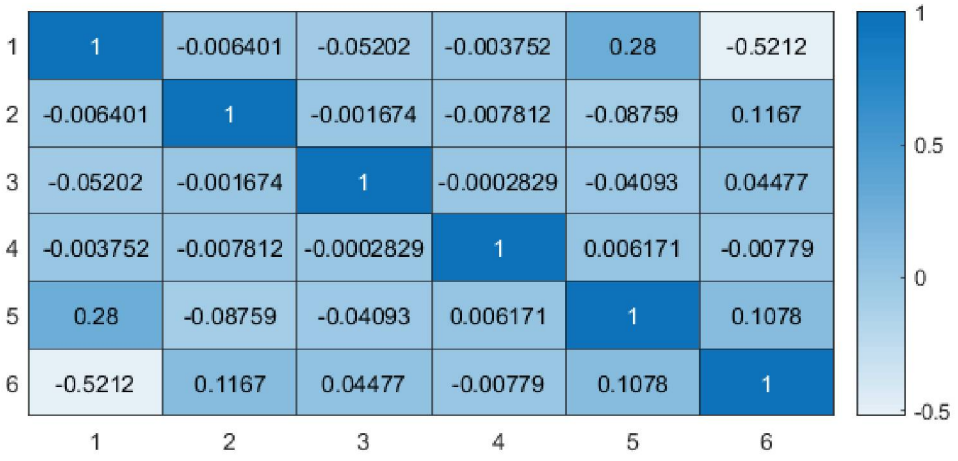
Fonte: O autor(2019).

Figura 69: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/dez/2018



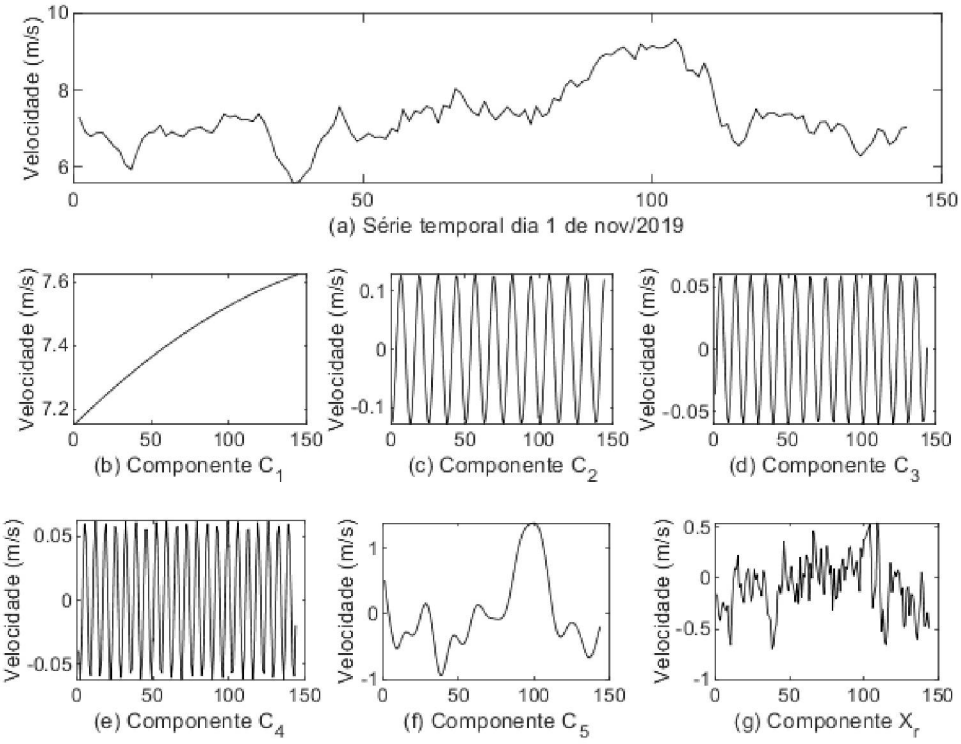
Fonte: O autor(2019).

Figura 70: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/dez/2018



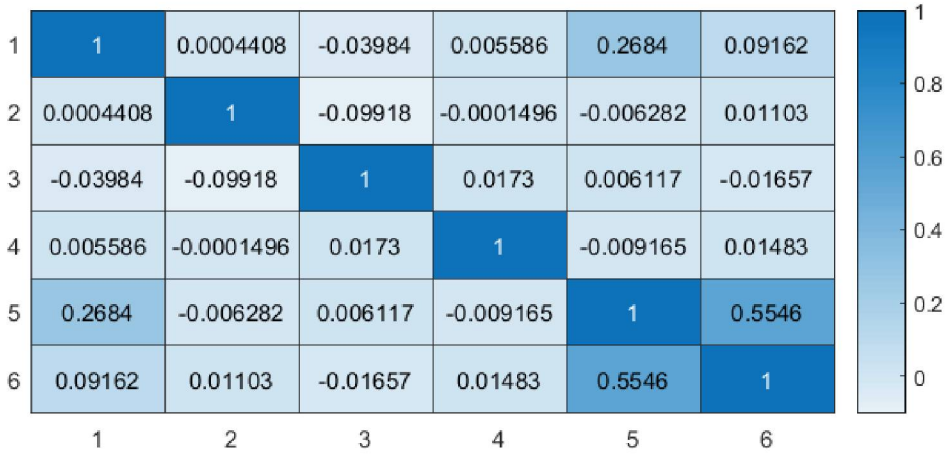
Fonte: O autor(2019).

Figura 71: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/nov/2018



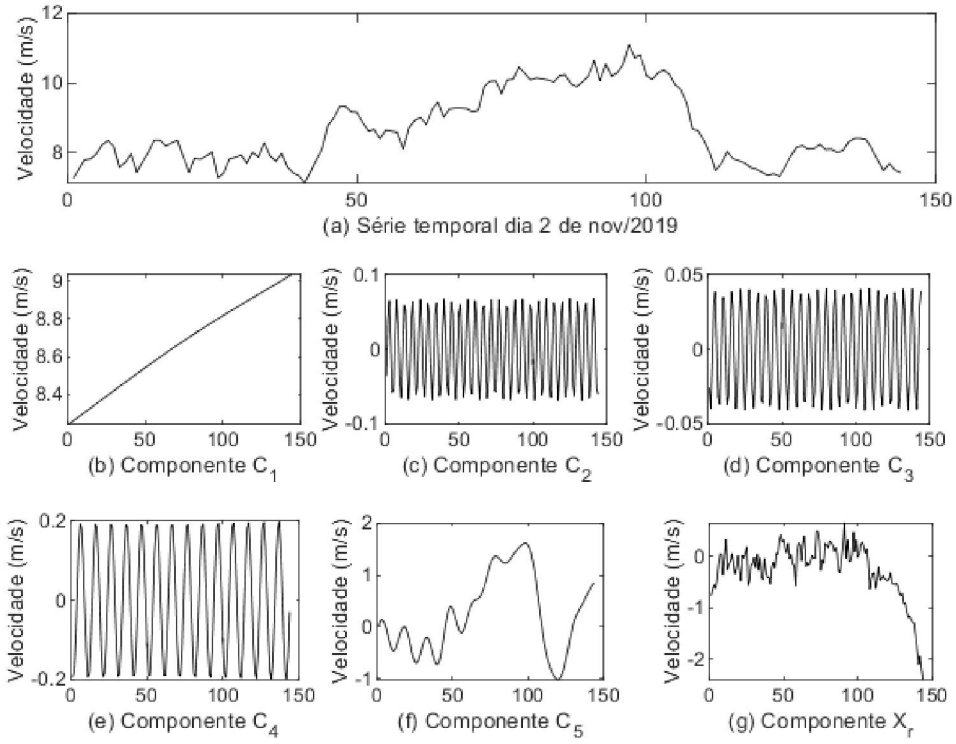
Fonte: O autor(2019).

Figura 72: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 01/nov/2018



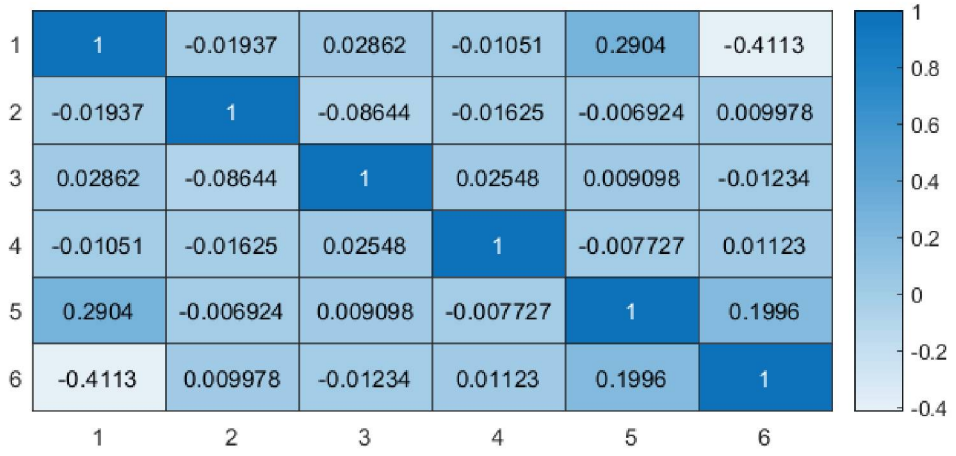
Fonte: O autor(2019).

Figura 73: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/nov/2018



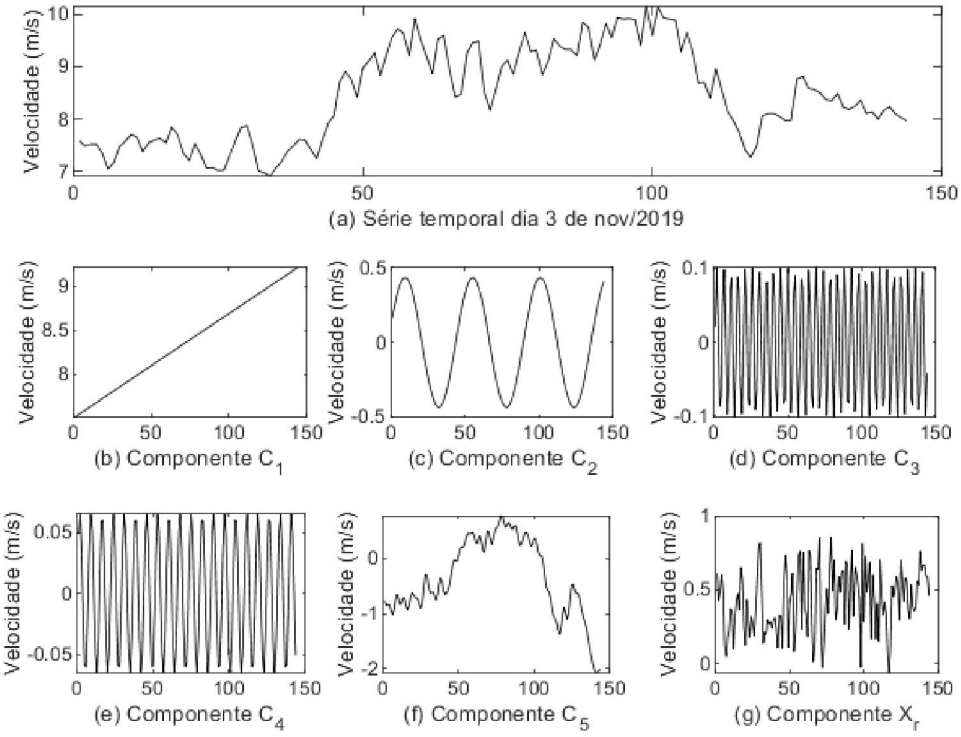
Fonte: O autor(2019).

Figura 74: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 02/nov/2018



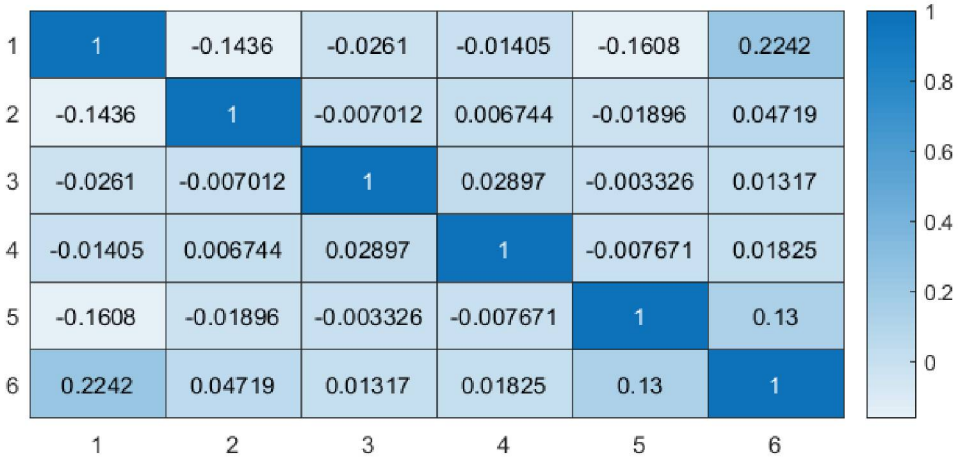
Fonte: O autor(2019).

Figura 75: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/nov/2018



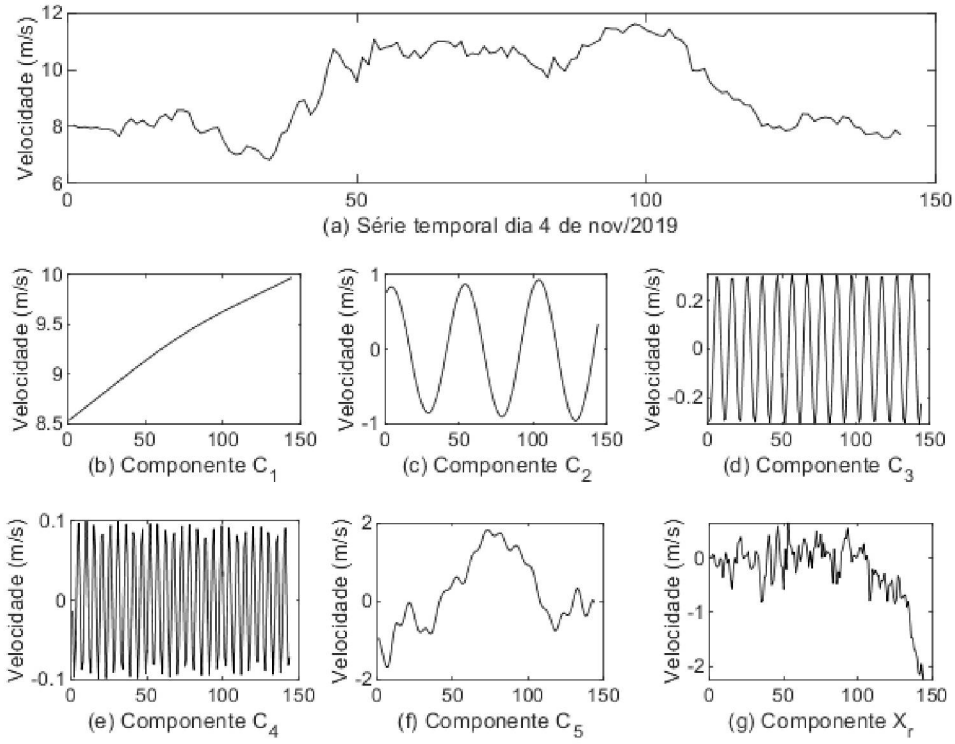
Fonte: O autor(2019).

Figura 76: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 03/nov/2018



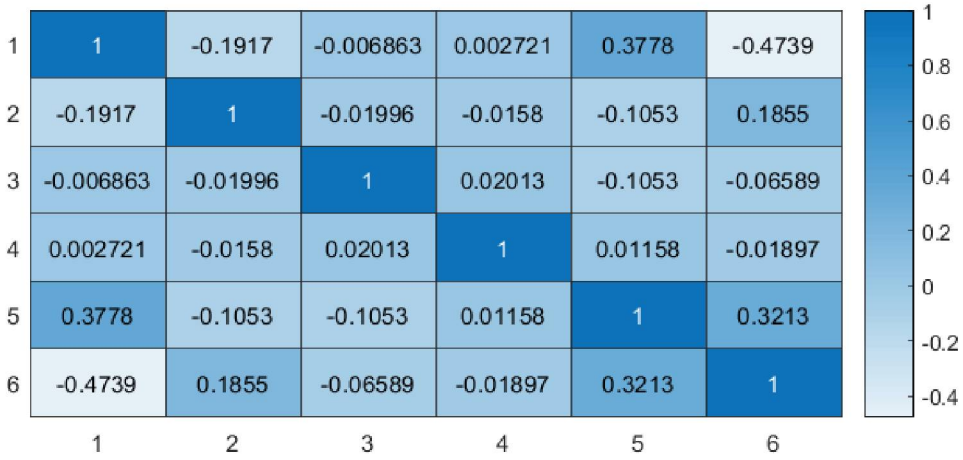
Fonte: O autor(2019).

Figura 77: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/nov/2018



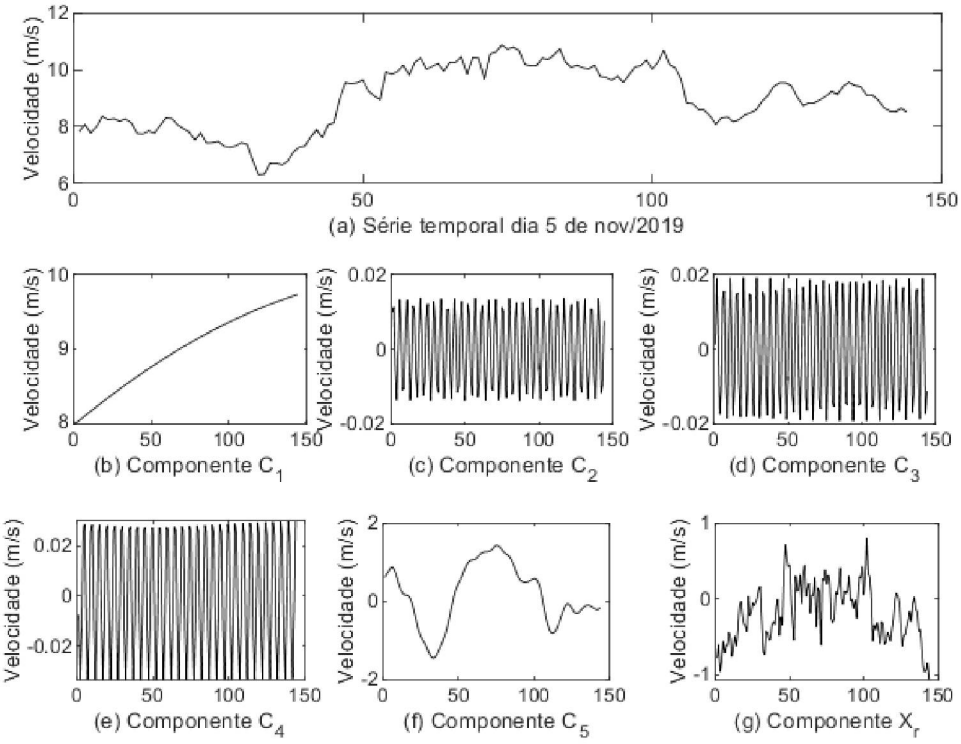
Fonte: O autor(2019).

Figura 78: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 04/nov/2018



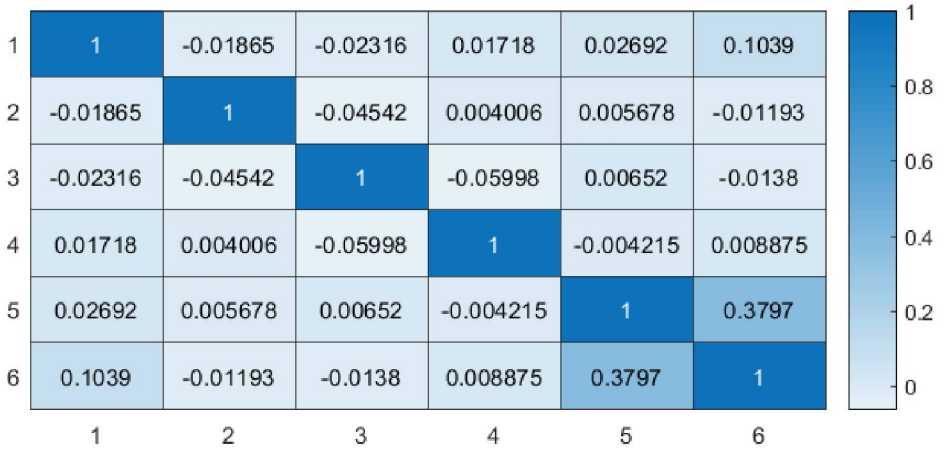
Fonte: O autor(2019).

Figura 79: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/nov/2018



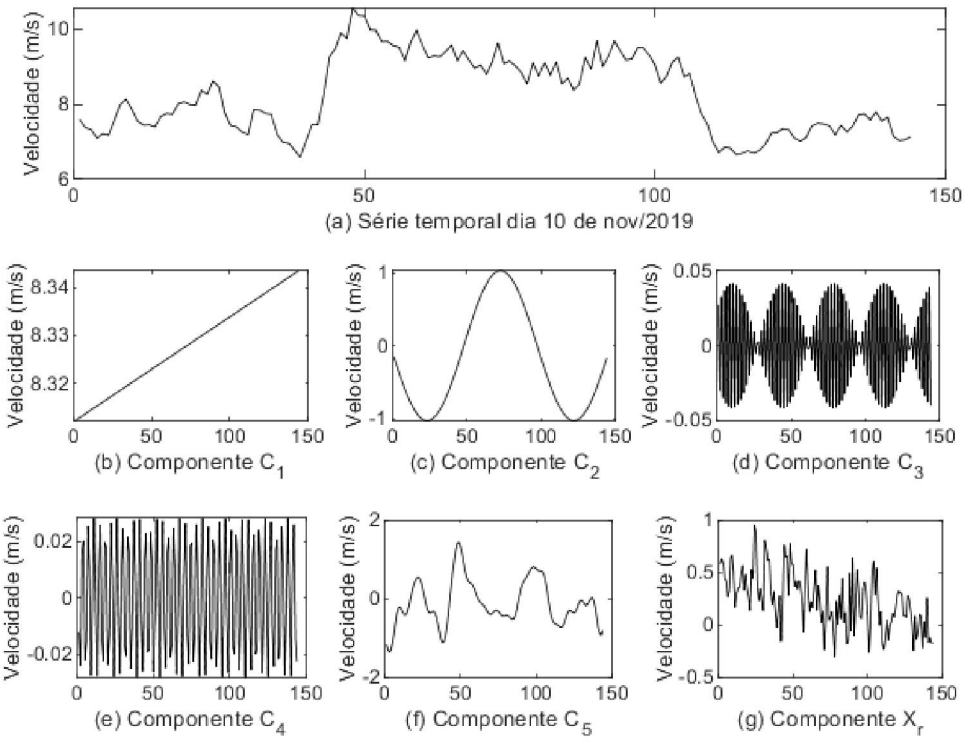
Fonte: O autor(2019).

Figura 80: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 05/nov/2018



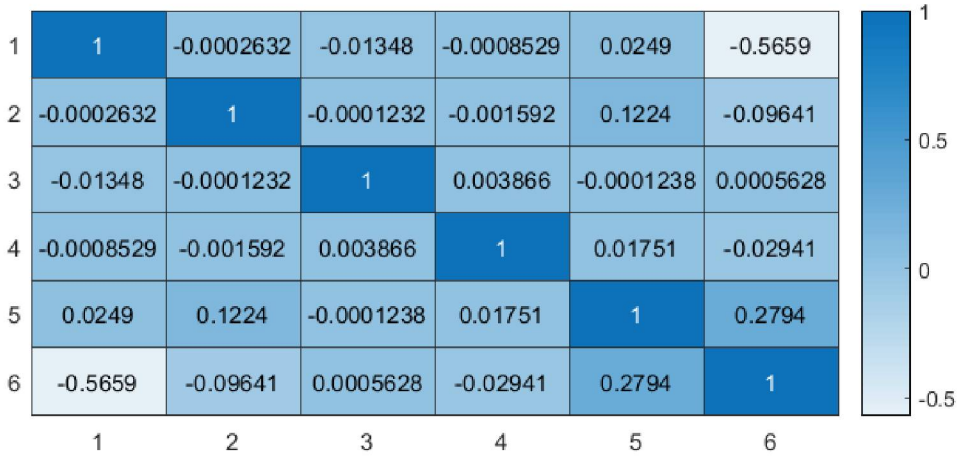
Fonte: O autor(2019).

Figura 81: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/nov/2018



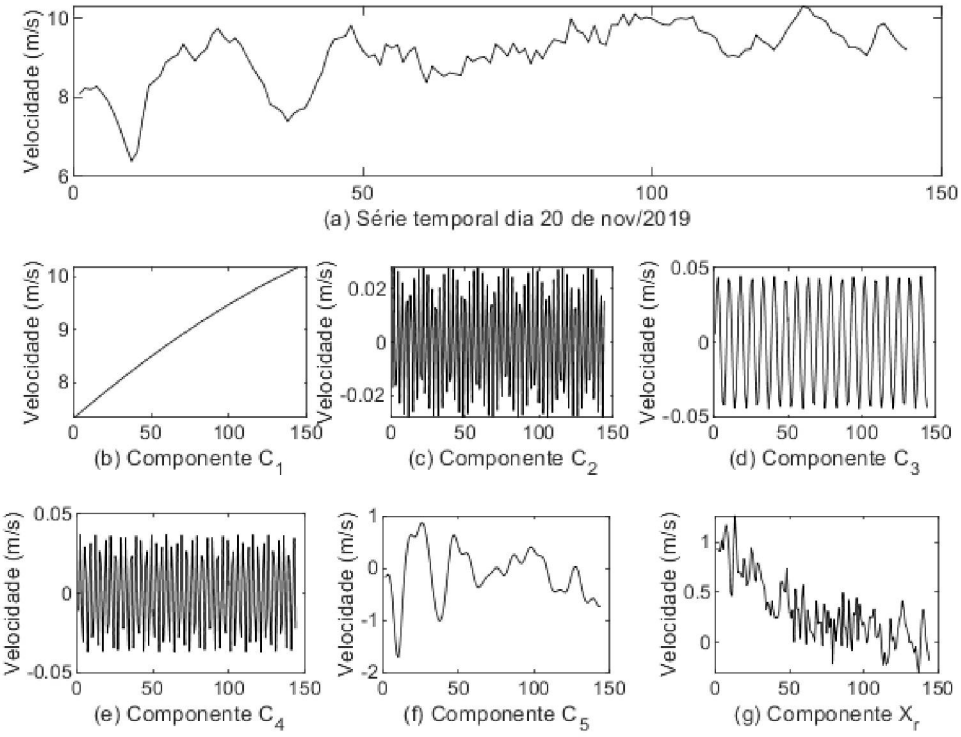
Fonte: O autor(2019).

Figura 82: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 10/nov/2018



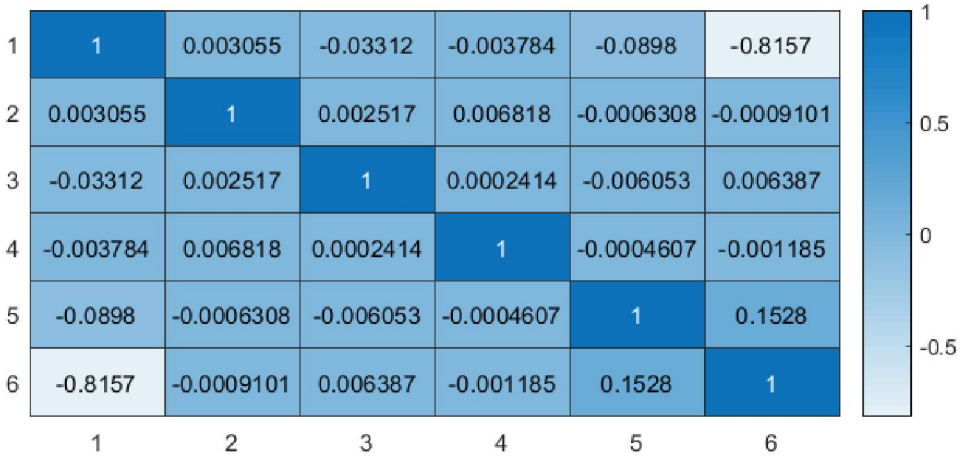
Fonte: O autor(2019).

Figura 83: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/nov/2018



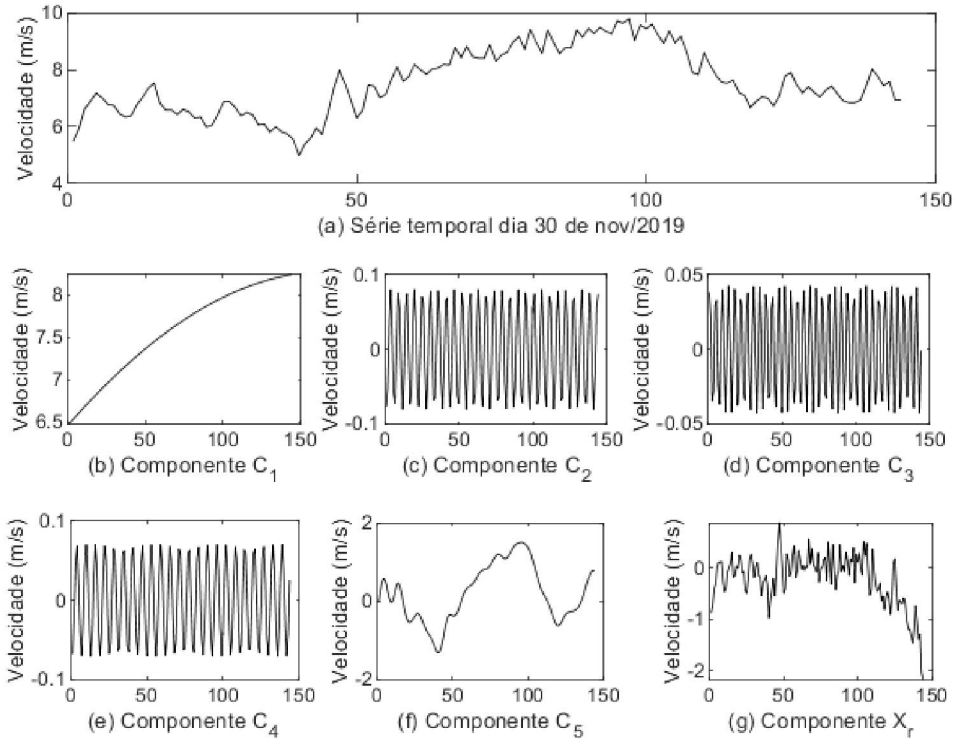
Fonte: O autor(2019).

Figura 84: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 20/nov/2018



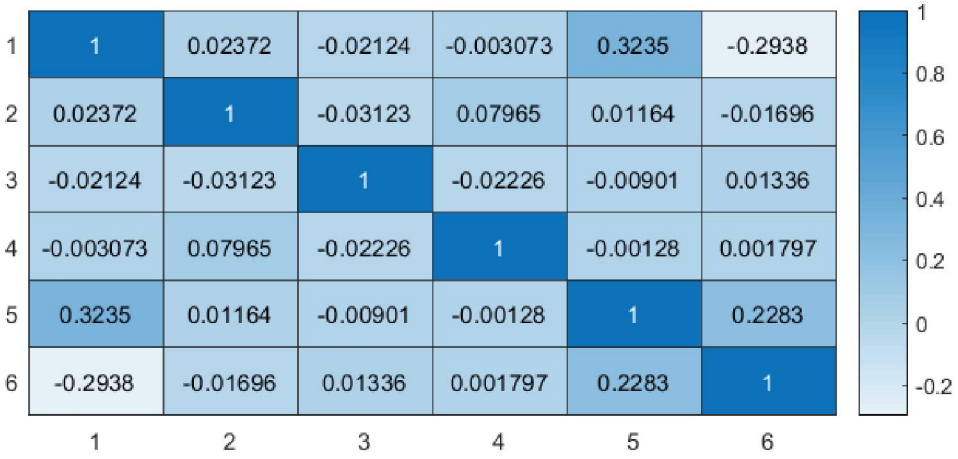
Fonte: O autor(2019).

Figura 85: Decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/nov/2018



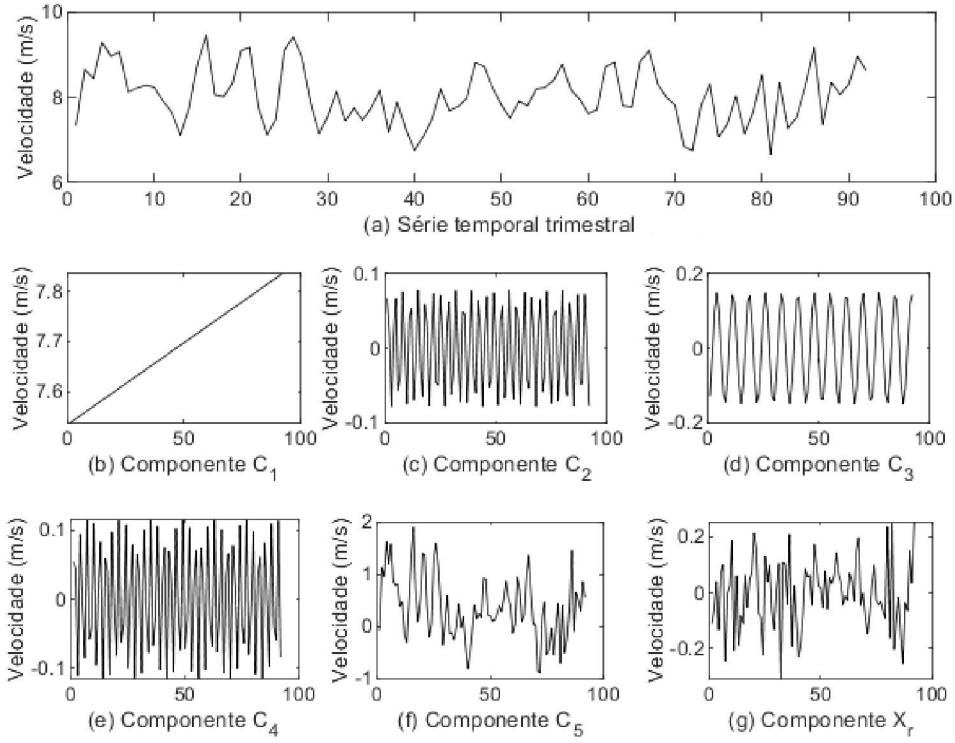
Fonte: O autor(2019).

Figura 86: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (10min) de 30/nov/2018



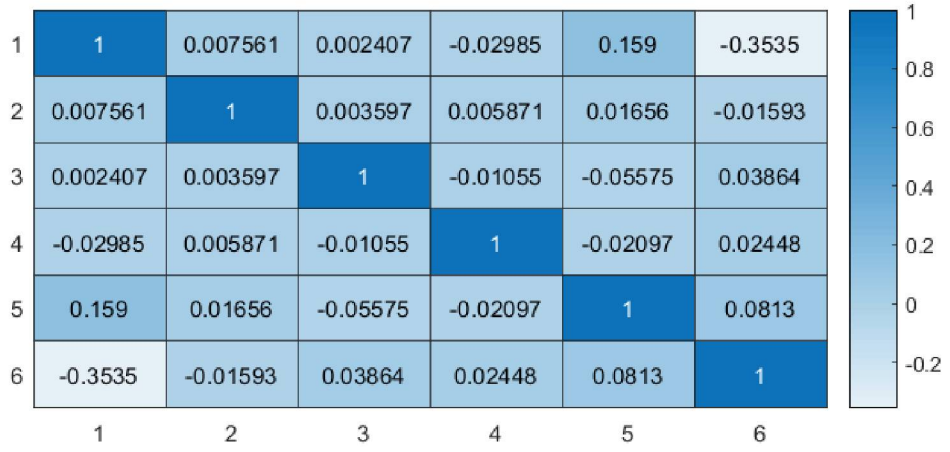
Fonte: O autor(2019).

Figura 87: Decomposição VMD-SSA da série temporal velocidades de vento (média diária) trimestre 01/nov/2018 à 31/jan/2019



Fonte: O autor(2019).

Figura 88: Correlação entre componentes obtidas pela decomposição VMD-SSA para série temporal velocidades de vento (média diária) trimestre 01/nov/2018 à 31/jan/2019



Fonte: O autor(2019).

5.2 RESUMO DO CAPÍTULO

Os resultados da decomposição foram apresentados para um conjunto de séries de velocidades de vento amostradas a cada 10min, totalizando 144 amostras/dia. O

modelo proposto se mostrou eficiente e capaz de separar as componentes da série temporal, de forma que explicitamente se tenha a componente de tendência e as componentes harmônicas ou sazonais, além das componentes com características mais ruidosas. A eficiência da separação das componentes foi medida analisando as correlações entre cada componente. Pode-se avaliar que para a maior parte do número de séries temporais utilizadas, as componentes resultantes estão descorrelacionadas ou apresentam fraca correlação.

A importância de se obter componentes descorrelacionadas ou com fraca correlação esta relacionada a acurácia da etapa de previsão, dado que componentes sazonais bem separadas são facilmente modeladas por senos e cossenos, tornando os erros das previsões mais estáveis. O mesmo ocorre por exemplo, com a componente de tendência, que normalmente possui comportamento linear ou quase-linear, cujo modelo de previsão pode ser facilmente representado por uma regressão, reduzindo desta forma a complexidade do modelo previsor. Os resultados das previsões são apresentadas no próximo capítulo, onde poderá ser avaliado o desempenho do modelo de decomposição, através da análise das métricas de desempenho, apresentadas também no próximo capítulo.

6 AVALIAÇÃO DO DESEMPENHO DO MODELO HÍBRIDO DE PREVISÃO

As previsões de 1 à K-passos-à-frente são realizadas utilizando o modelo paramétrico ARIMA e a rede neural LSTM, esta última, pertencente a classe de modelos não paramétricos. Os resultados são apresentados neste capítulo, utilizando como entrada as séries temporais com comportamentos e comprimentos bastante distintos, além das componentes distintas obtidas na etapa de decomposição com o modelo híbrido proposto. Além de ser apresentado os resultados das previsões para os modelos de Box & Jenkins e da técnica LSTM, as medidas de qualidade da previsão são apresentadas e discutidas, além de outras considerações sobre os resultados que são feitas no capítulo.

6.1 RESULTADOS DAS PREVISÕES

As previsão de 1 até 12 passos-à-frente foram obtidas utilizando inicialmente os modelos de Box & Jenkins, conforme detalhado na Seção 4.1.2. Vale destacar que ao analisarmos o comportamento de cada componente, fica evidente a forma de aplicar os modelos de Box & Jenkins, pois cada componente possui um comportamento específico e que, na maioria das vezes, é representado por um modelo de estrutura bem simples, apenas com o componente AR(p) ou o MA(q). Salvo para as componentes C_5 e X_r, onde é necessário alguma transformação para estabilizar o modelo. Vale recordar que a etapa de previsão é aplicada à cada componente, sendo ajustado um modelo que as representam, e por fim, se obtém as previsões para cada modo. Após a obtenção das mesmas, essas componentes são agrupadas novamente, reconstruindo a série original e a parcela prevista.

Para não tornar tediosa e extenuante a leitura desta seção, limitamos o conjunto de dados das séries que serão utilizadas na previsão. Desta forma, nem todas as séries diárias apresentadas na etapa anterior serão utilizadas na previsão. Na seção anterior fez sentido avaliarmos cada uma das séries, pois queríamos comprovar a fraca correlação obtida entre as componentes, para tal um número razoável de amostras é necessário. Feito

esse aparte, o conjunto de dados utilizado para a previsão esta detalhado na Tabela 13.

Tabela 13: Conjuntos de dados para a etapa de previsão

Série Ref.	Num. de Amostras	Interv. t_s	Horizonte K_{step}	Tipo de Previsão	Técnica
01/Jan/19	144	10 min.	12	*out-of-sample*	VMD-SSA-ARIMA, ARIMA
02/Jan/19	144	10 min.	12	*out-of-sample*	VMD-SSA-ARIMA, ARIMA
03/Jan/19	144	10 min.	12	*out-of-sample*	VMD-SSA-ARIMA, ARIMA
04/Jan/19	144	10 min.	12	*out-of-sample*	VMD-SSA-ARIMA, ARIMA
05/Jan/19	144	10 min.	12	*out-of-sample*	VMD-SSA-ARIMA, ARIMA
01/Nov/18 a 31/Jan/19	92	1 dia	12	*out-of-sample*	VMD-SSA-ARIMA, ARIMA
01/Jan/19	132	10 min.	12	*out-of-sample*	LSTM, VMD-SSA-LSTM
02/Jan/19	132	10 min.	12	*out-of-sample*	LSTM, VMD-SSA-LSTM
03/Jan/19	132	10 min.	12	*out-of-sample*	LSTM, VMD-SSA-LSTM
04/Jan/19	132	10 min.	12	*out-of-sample*	LSTM, VMD-SSA-LSTM
05/Jan/19	132	10 min.	12	*out-of-sample*	LSTM, VMD-SSA-LSTM
01/Nov/18 a 31/Jan/19	82	1 dia	10	*out-of-sample*	LSTM, VMD-SSA-LSTM

Fonte: O autor(2019).

É importante ressaltar que as séries amostradas a cada 10min, formam 144 amostras em um dia que compreende o intervalo de início às 00h:10min e fim 00h00min, ou seja, o dia começa às 00h:10min do primeiro intervalo amostrado e finaliza uma amostra após a virada do dia. Este esclarecimento de faz necessário para evitar problemas de interpretação dos resultados. Cabe ainda ressaltar que este método de *time stamp* é muito comum no Setor Elétrico.

Os resultados das previsões para cada componente da série de velocidades para os dias 01 à 05 de janeiro de 2019 e para o trimestre, são apresentados nas próximas seções. Cada série de velocidades utilizada para a etapa de previsão possui 144 amostras (1 a cada 10min), conforme já detalhado, desta forma, a série sempre termina às 00h00 do dia corrente D, e as previsões *out-of-sample* são para as duas primeiras horas do dia seguinte, que correspondem aos 12 valores relativos ao intervalos de 00h:10min à 02h:00min do dia $D+1$ para o modelo ARIMA. Já as previsões para à rede neural LSTM, a série temporal de velocidades de vento termina às 22h00min e as previsões são relativas as duas horas para concluir o dia D, ou seja, 12 intervalos de 10min.

Para facilitar a comparação de resultados, foi também realizado as previsões para a série original sem a etapa de decomposição VMD-SSA, porém utilizando a mesma técnica de previsão, visando destacar os benefícios introduzidos pelo modelo híbrido de decomposição VMD-SSA. As previsões obtidas com a rede LSTM, seja na forma híbrida ou não, consideram a série temporal de entrada para o modelo com apenas 132 amostras, onde se deseja prever as 12 amostras finais do dia, totalizando desta forma 144 amostras.

Tais benefícios ficam explícitos quando analisamos o erro das previsões e os indicadores de desempenho do modelo para cada passo-à-frente de previsão até o horizonte final. Os indicadores do índice de desempenho são apresentados imediatamente após os resultados da previsão para facilitar a análise do resultados. Porém como os mesmos não foram definidos ainda, cabe apresentá-los já próxima seção.

6.2 MÉTRICAS DE DESEMPENHO

Para avaliar o desempenho do modelo proposto, as medidas de erro utilizadas são as normalmente encontradas na literatura (DESPOTOVIC et al., 2016), a saber: erro médio quadrático (MSE); raiz do erro médio quadrático (RMSE); erro médio absoluto (MAE); erro médio percentual absoluto (MAPE) e raiz do erro médio quadrático relativo (RRMSE). As Equações 6.2 à 6.5 relativas a cada indicador são apresentadas na sequência.

O MAE representa a soma dos valores absolutos dos erros divididos pelo número de observações. Essa quantidade é frequentemente usada em estatística como medida de dispersão entre os valores previstos $\hat{y}_i$ e os observados y_i. Segundo o trabalho de Despotovic et al. (2016), o mesmo aponta que alguns autores identificaram algumas vantagens do MAE em relação RMSE, quando realizavam avaliações dimensionais e inter-comparações do erro médio de desempenho do modelo. Na prática o RMSE possui forte relação com o número de amostras, isso implica que quando o número de amostras de teste (y_i) cresce, o valor do RMSE será muito maior que o valor do MAE.

O RRMSE segundo Despotovic et al. (2016) fornece uma indicação da qualidade do ajuste do modelo, tendo a escala de classificação, conforme Equação 6.1,

$$\text{Acuracia do Modelo} = \begin{cases} \text{excelente}, & \text{se RRMSE} < 10\% \\ \text{bom}, & \text{se } 10\% < \text{RRMSE} < 20\% \\ \text{regular}, & \text{se } 20\% < \text{RRMSE} < 30\% \\ \text{fraco}, & \text{se RRMSE} > 30\% \end{cases} \tag{6.1}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{6.2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{6.3}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{6.4}$$

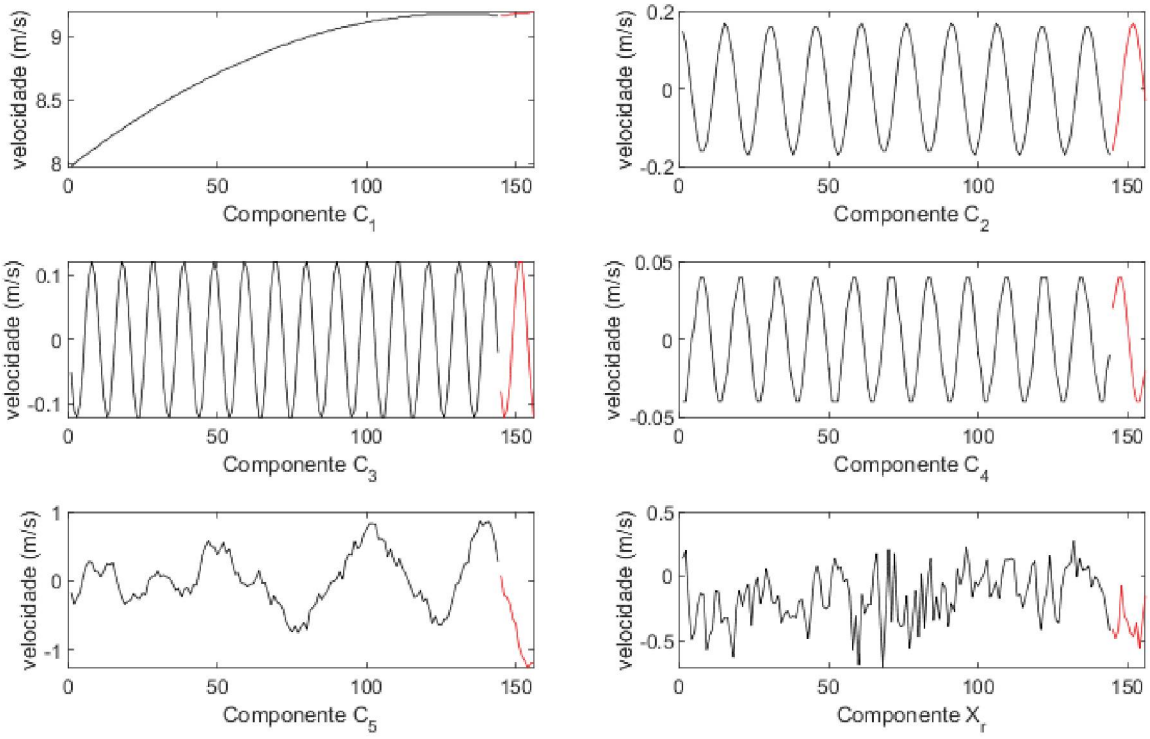
$$MAPE[\%] = \frac{100}{N}\sum_{i=1}^{N}\frac{|\hat{y}_i - y_i|}{y_i} \tag{6.5}$$

$$RRMSE = \frac{RMSE}{\frac{1}{N}\sum_{i=1}^{N}\left(y^{real}\right)} \tag{6.6}$$

6.2.1 RESULTADOS DAS PREVISÕES PARA OS DADOS OBSERVADOS EM 01/JAN/2019

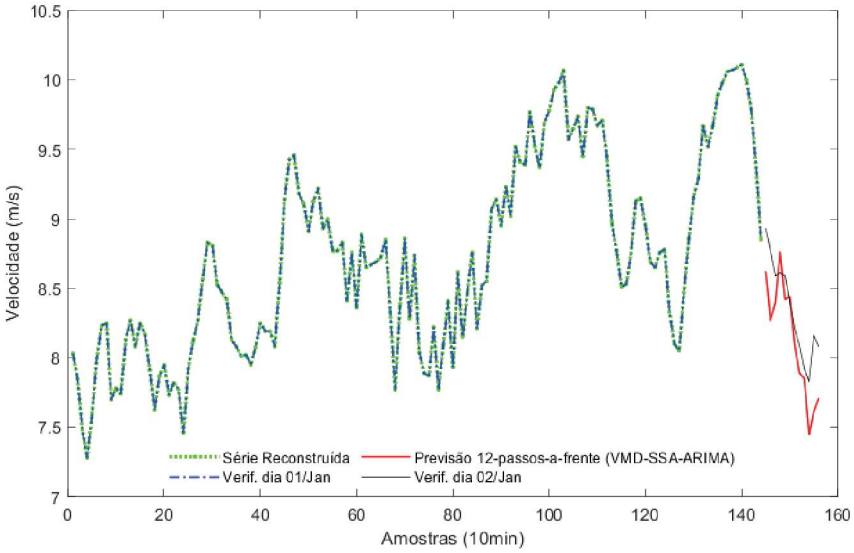
Como já destacado, para cada componente obtida na decomposição foi ajustado um modelo ARIMA(p,d,q) para representar o comportamento individual das mesmas. Os resultados para as previsões relativas à continuação da série verificada no dia 01/jan estão destacados nas Figuras 89 à 91. Vale frisar que os resultados apresentados nas figuras são para um horizonte de 12 intervalos à frente, porém para estabelecer um parâmetro de qualidade dos resultados obtido pelo modelo, os erros são analisados para cada passo de previsão, ou seja, $MSE_{\hat{y}_{N+k}}$ para $k = 1, 2$ é dado por $\hat{y}_{N+1}$ e $\hat{y}_{N+2}$, e assim sucessivamente, para todos os $\hat{y}_{N+k}$ com $k = 1, 2, \ldots, K$ passos-à-frente, até o horizonte $K = 12$.

Figura 89: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 01/jan/2019 através do método de Box & Jenkins.



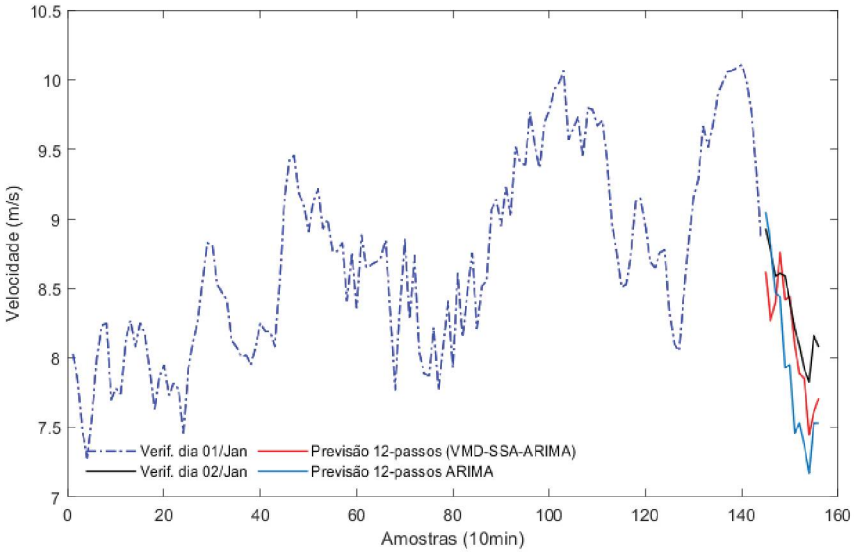
Fonte: O autor(2019).

Figura 90: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 01 e 02/Jan/19.



Fonte: O autor(2019).

Figura 91: Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 01 e 02/Jan/19.



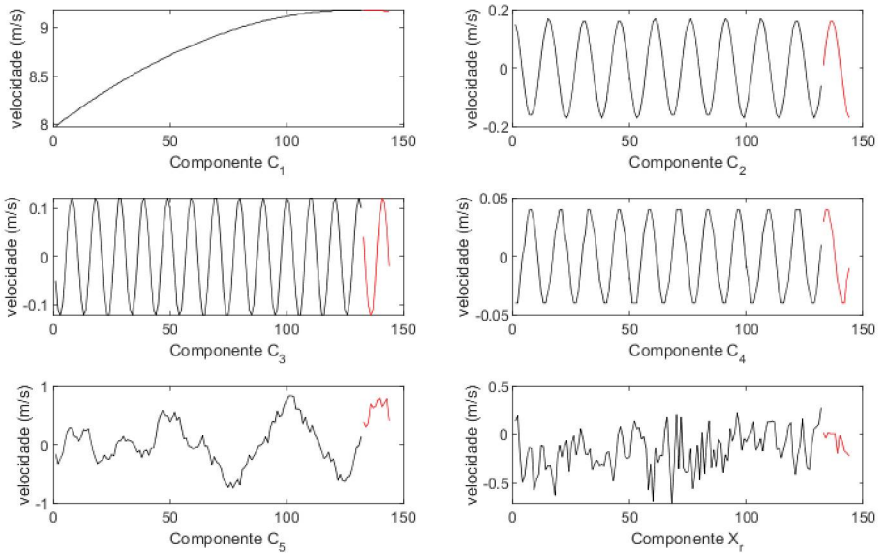
Fonte: O autor(2019).

É possível observar que as previsões para às componentes C_1 à C_4 são as que inserem menor erro na previsão final ao reconstruir a série, dado que o comportamento dessas componentes são cíclicos e de tendência. Já as componentes C_5 e X_r possuem a maior fonte de erro, dado ao comportamento aleatório. Porém ao observamos os gráficos das mesmas, pode-se concluir que as amplitudes dessas componentes são limitadas em -1 e 1, desta forma, o erro de previsão estará limitado à esse intervalo.

Ao avaliar à previsão para a série reconstruída, lembrando que esta previsão é a soma das previsões de cada componente. Pode-se perceber que, o modelo ARIMA(p,d,q) conseguiu capturar o comportamento da série distribuído em cada componente, e realizar previsões com menores erros a partir do quinto intervalo previsto até o final do horizonte de previsão. Já a previsão que utiliza apenas o modelo ARIMA(p,d,q) sem a etapa de decomposição, resultou em erros maiores na previsão de 5 a 12 passos-à-frente.

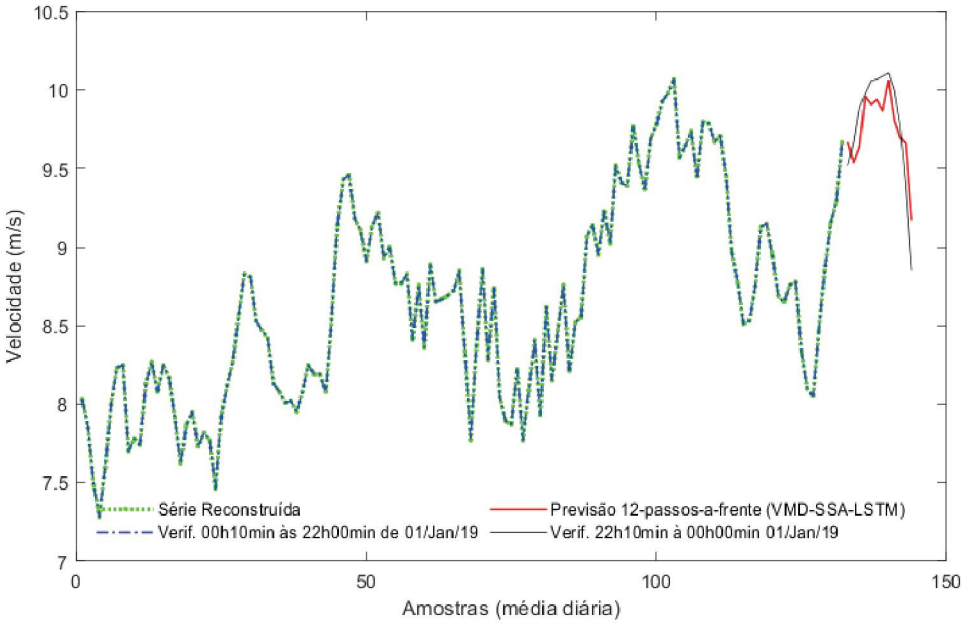
As previsões utilizando a rede neural LSTM são apresentadas nas Figuras 92 à 94, da mesma forma, o resultado do modelo híbrido é comparado ao resultado sem a influência da etapa da decomposição. Pelos resultados apresentados na Figura 94, percebe-se que a previsão resultante do modelo híbrido é superior aquela que utiliza apenas a rede neural LSTM.

Figura 92: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 01/jan/2019 através da LSTM



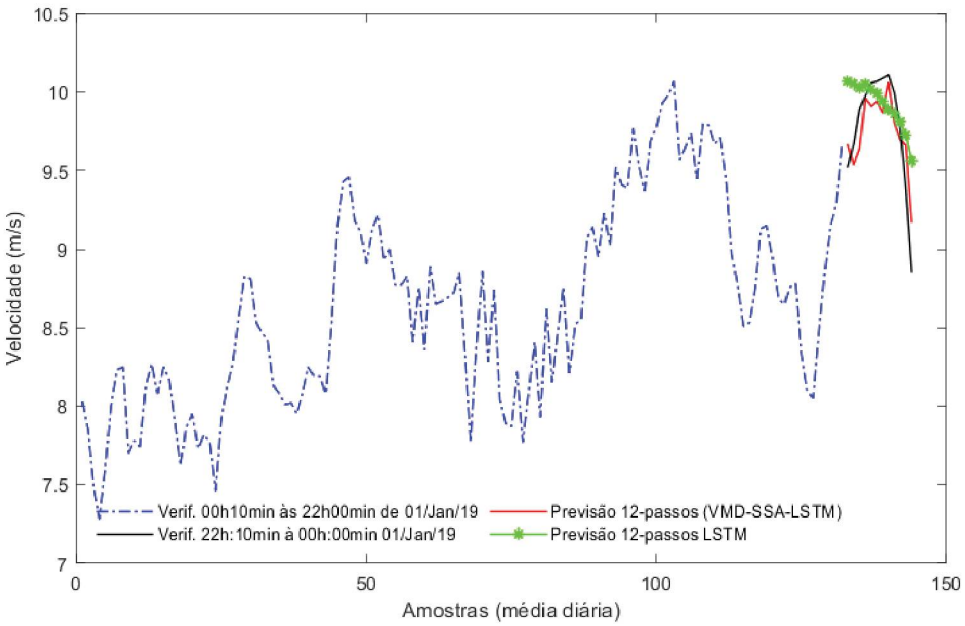
Fonte: O autor(2019).

Figura 93: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 01/jan/19.



Fonte: O autor(2019).

Figura 94: Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 01/jan/19.



Fonte: O autor(2019).

Os erros do modelo relativo às previsões de 1 à 12-passos-à-frente estão destacados na Tabela 14 para os modelos VMD-SSA-ARIMA comparado ao ARIMA, e na Tabela 15 para os modelos VMD-SSA-LSTM comparado à rede neural LSTM. Pode-se observar na Tabela 14 que os erros das previsões para 1 a 4-passos-à-frente foram menores para o modelo ARIMA sem a etapa de decomposição, porém após o horizonte $\hat{y}_{t+5}$, as previsões começam a degradar gradativamente aumentando o erro, enquanto que para o modelo híbrido, os erros permanecem estabilizados, sem apresentarem uma tendência de aumento dos valores, mantendo as previsões mais coerentes.

Tabela 14: Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 02/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-ARIMA	$\hat{y}_{t+1}$	-0,310	0,096	0,310	0,310	3,47	3,47
	$\hat{y}_{t+2}$	-0,499	0,172	0,415	0,404	4,58	4,69
	$\hat{y}_{t+3}$	-0,193	0,127	0,357	0,334	3,80	4,07
	$\hat{y}_{t+4}$	0,152	0,101	0,318	0,288	3,29	3,65
	$\hat{y}_{t+5}$	**-0,165**	**0,086**	**0,294**	**0,264**	**3,02**	**3,38**
	$\hat{y}_{t+6}$	**0,028**	**0,072**	**0,269**	**0,225**	**2,57**	**3,11**
	$\hat{y}_{t+7}$	**-0,123**	**0,064**	**0,253**	**0,210**	**2,42**	**2,95**
	$\hat{y}_{t+8}$	**-0,196**	**0,061**	**0,247**	**0,208**	**2,42**	**2,89**
	$\hat{y}_{t+9}$	**-0,065**	**0,055**	**0,234**	**0,192**	**2,24**	**2,76**
	$\hat{y}_{t+10}$	**-0,382**	**0,064**	**0,252**	**0,211**	**2,51**	**3,01**
	$\hat{y}_{t+11}$	**-0,552**	**0,086**	**0,293**	**0,242**	**2,89**	**3,49**
	$\hat{y}_{t+12}$	**-0,373**	**0,090**	**0,300**	**0,253**	**3,04**	**3,59**
ARIMA	$\hat{y}_{t+1}$	**0,120**	**0,015**	**0,120**	**0,120**	**1,35**	**1,35**
	$\hat{y}_{t+2}$	**0,055**	**0,009**	**0,093**	**0,087**	**0,99**	**1,06**
	$\hat{y}_{t+3}$	**-0,116**	**0,010**	**0,102**	**0,097**	**1,11**	**1,16**
	$\hat{y}_{t+4}$	**-0,171**	**0,015**	**0,123**	**0,115**	**1,33**	**1,41**
	$\hat{y}_{t+5}$	-0,662	0,100	0,316	0,225	2,60	3,63
	$\hat{y}_{t+6}$	-0,460	0,118	0,344	0,264	3,08	3,98
	$\hat{y}_{t+7}$	-0,745	0,181	0,425	0,333	3,94	4,95
	$\hat{y}_{t+8}$	-0,560	0,197	0,444	0,361	4,31	5,21
	$\hat{y}_{t+9}$	-0,545	0,209	0,457	0,382	4,60	5,40
	$\hat{y}_{t+10}$	-0,662	0,231	0,481	0,410	4,98	5,73
	$\hat{y}_{t+11}$	-0,634	0,247	0,497	0,430	5,24	5,93
	$\hat{y}_{t+12}$	-0,546	0,251	0,501	0,440	5,36	6,00

Fonte: O autor(2019).

Os resultados das previsões para o modelo híbrido VMD-SSA-LSTM foram superiores aos obtidos pela rede neural LSTM, como pode ser observado pelos índices de qualidade da Tabela 15, para todos os horizontes de previsão. Outro ponto importante a ser destacado, conforme já apresentado nas discussões dos resultados dos erros do modelo ARIMA, é estabilização do erro provocado pelo modelo de decomposição.

Tabela 15: Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 01/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-LSTM	$\hat{y}_{t+1}$	**0,152**	**0,023**	**0,152**	**0,152**	**1,60**	**1,60**
	$\hat{y}_{t+2}$	**-0,125**	**0,019**	**0,140**	**0,139**	**1,45**	**1,45**
	$\hat{y}_{t+3}$	**-0,247**	**0,033**	**0,182**	**0,175**	**1,80**	**1,88**
	$\hat{y}_{t+4}$	**-0,016**	**0,025**	**0,158**	**0,135**	**1,39**	**1,62**
	$\hat{y}_{t+5}$	**-0,152**	**0,025**	**0,157**	**0,138**	**1,41**	**1,60**
	$\hat{y}_{t+6}$	**-0,129**	**0,023**	**0,153**	**0,137**	**1,39**	**1,55**
	$\hat{y}_{t+7}$	**-0,225**	**0,027**	**0,165**	**0,149**	**1,51**	**1,67**
	$\hat{y}_{t+8}$	**-0,054**	**0,024**	**0,155**	**0,138**	**1,39**	**1,57**
	$\hat{y}_{t+9}$	**-0,195**	**0,026**	**0,160**	**0,144**	**1,45**	**1,61**
	$\hat{y}_{t+10}$	**-0,080**	**0,024**	**0,154**	**0,138**	**1,39**	**1,55**
	$\hat{y}_{t+11}$	**0,295**	**0,029**	**0,172**	**0,152**	**1,55**	**1,74**
	$\hat{y}_{t+12}$	**0,318**	**0,035**	**0,188**	**0,166**	**1,72**	**1,93**
LSTM	$\hat{y}_{t+1}$	0,555	0,308	0,555	0,555	5,83	5,83
	$\hat{y}_{t+2}$	0,377	0,225	0,474	0,466	4,86	4,94
	$\hat{y}_{t+3}$	0,142	0,157	0,396	0,358	3,72	4,08
	$\hat{y}_{t+4}$	0,071	0,119	0,345	0,286	2,97	3,53
	$\hat{y}_{t+5}$	-0,037	0,095	0,309	0,236	2,45	3,14
	$\hat{y}_{t+6}$	-0,078	0,080	0,284	0,210	2,17	2,87
	$\hat{y}_{t+7}$	-0,147	0,072	0,268	0,201	2,07	2,71
	$\hat{y}_{t+8}$	-0,221	0,069	0,263	0,204	2,08	2,65
	$\hat{y}_{t+9}$	-0,143	0,064	0,252	0,197	2,01	2,54
	$\hat{y}_{t+10}$	0,031	0,057	0,240	0,180	1,84	2,42
	$\hat{y}_{t+11}$	0,358	0,064	0,253	0,196	2,02	2,56
	$\hat{y}_{t+12}$	0,706	0,100	0,316	0,239	2,52	3,23

Fonte: O autor(2019).

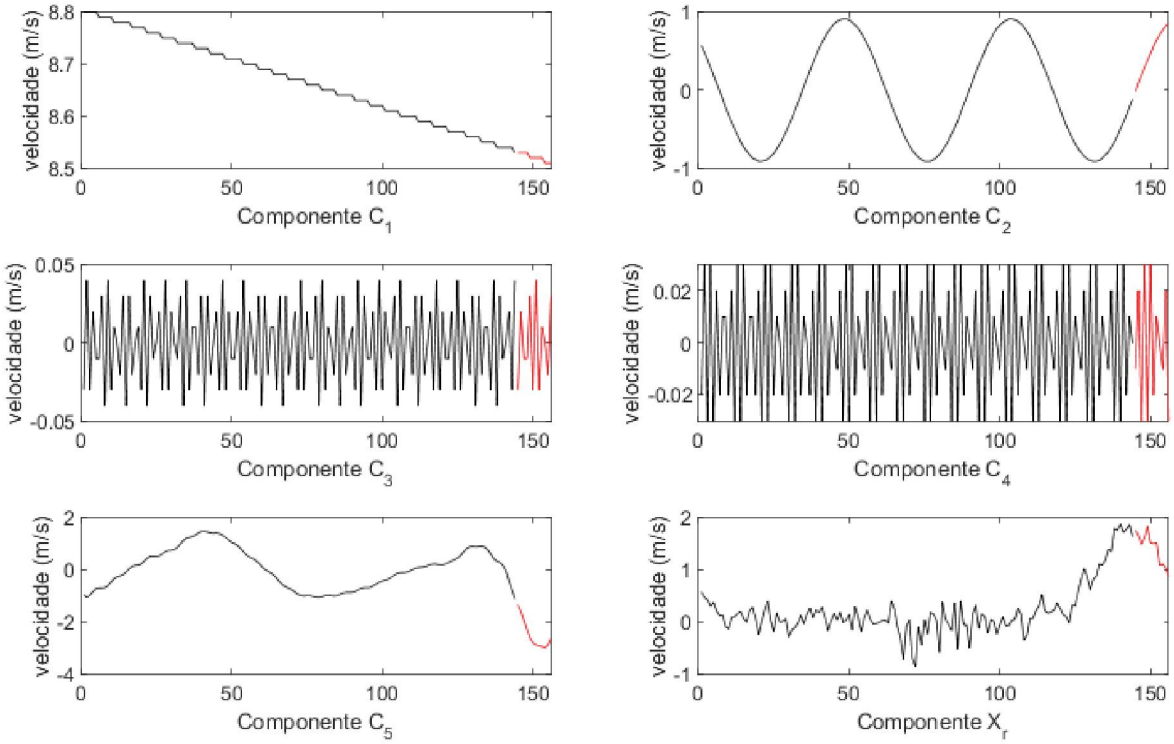
6.2.2 PREVISÕES PARA OS DADOS OBSERVADOS EM 02/JAN/2019

Da mesma forma que a seção anterior, os resultados para as previsões $k = 1, 2, \ldots, K$ com $K = 12$, são apresentados nas Figuras 95 à 100 para as quatro técnicas de previsão. Os índices de qualidade de cada modelo estão em destaque nas Tabelas 16 e 17.

As previsões resultantes do modelo ARIMA, devido a série original estar em uma rampa de descida após uma elevação abrupta, o modelo não conseguiu determinar o comportamento da série de forma adequada, dado que o modelo de forma geral tem a necessidade de capturar e representar os mecanismos formador da série, para que as pre-

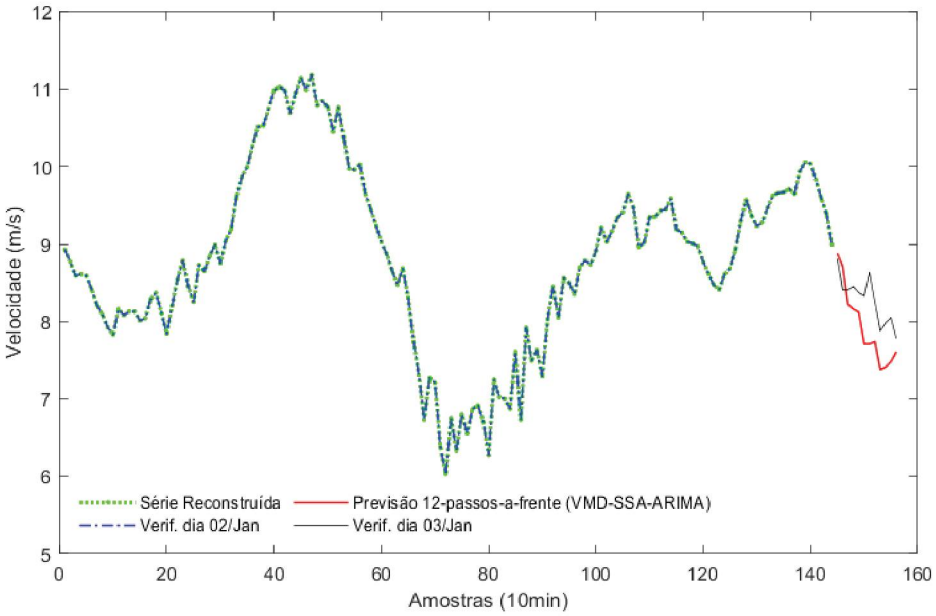
visões serem aderentes. No entanto, as previsões resultantes do modelo ARIMA acoplado ao modelo de decomposição são aderentes a tendência do sinal real, com um pequeno deslocamento de amplitude, conforme pode ser visto na Figura 97.

Figura 95: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 02/Jan/2019 através do método de Box & Jenkins



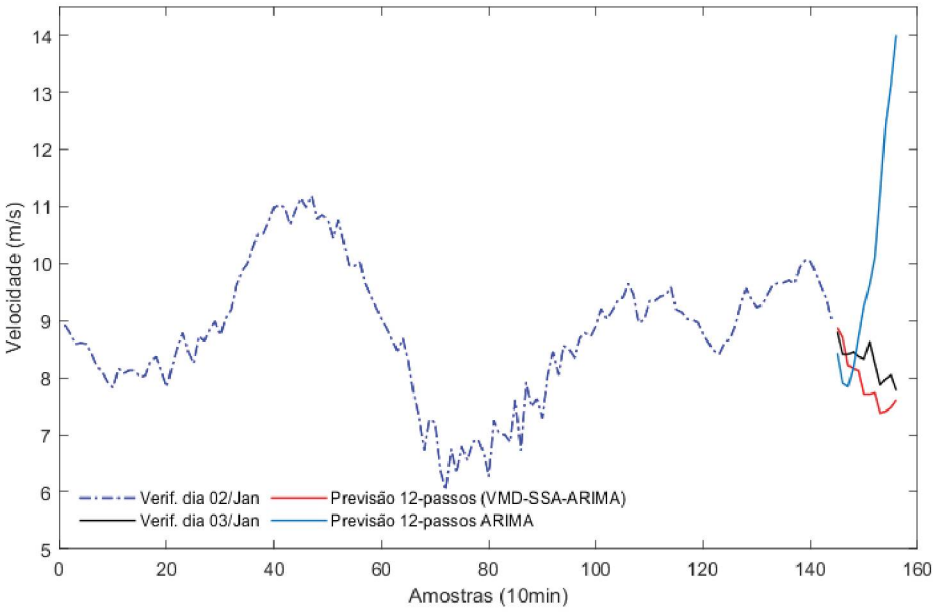
Fonte: O autor(2019).

Figura 96: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 02 e 03/jan/19.



Fonte: O autor(2019).

Figura 97: Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 02 e 03/jan/19.

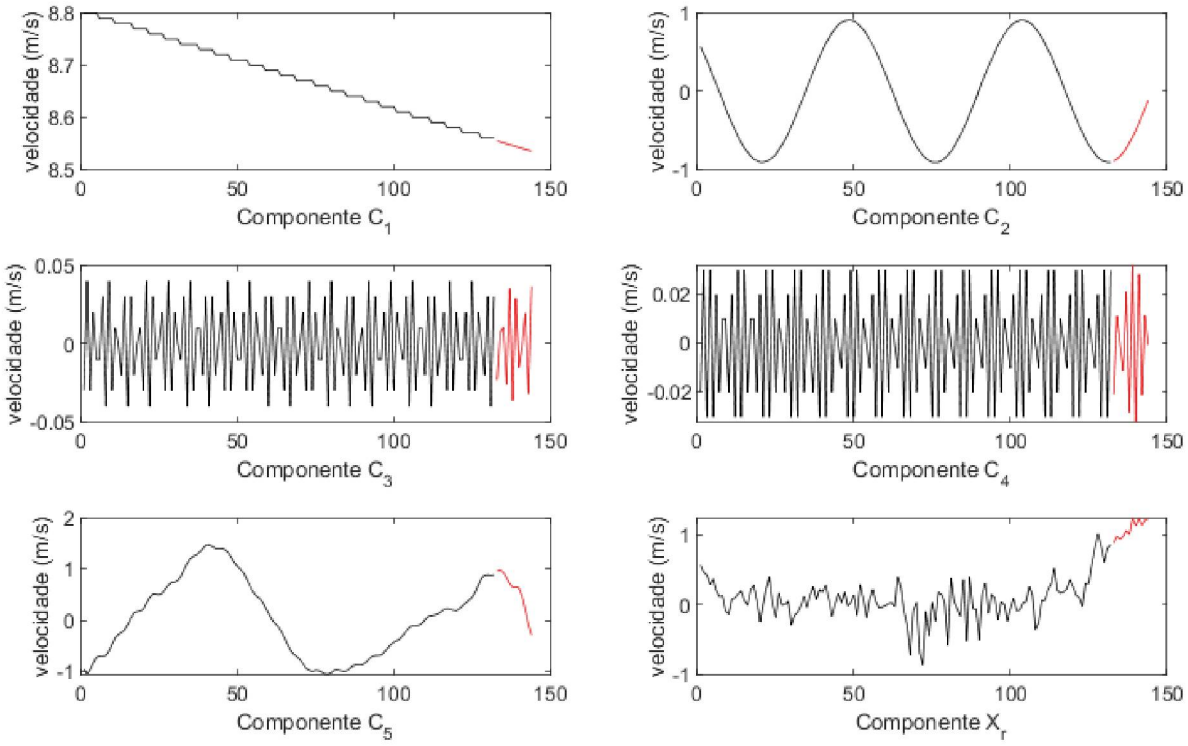


Fonte: O autor(2019).

Os resultados para a técnica VMD-SSA-LSTM comparado aos resultados da rede neural LSTM são apresentados nas Figuras 98 à 100. Os resultados para a série original confirmam a estabilidade do erro, ao analisarmos os resultados do modelo VMD-SSA-

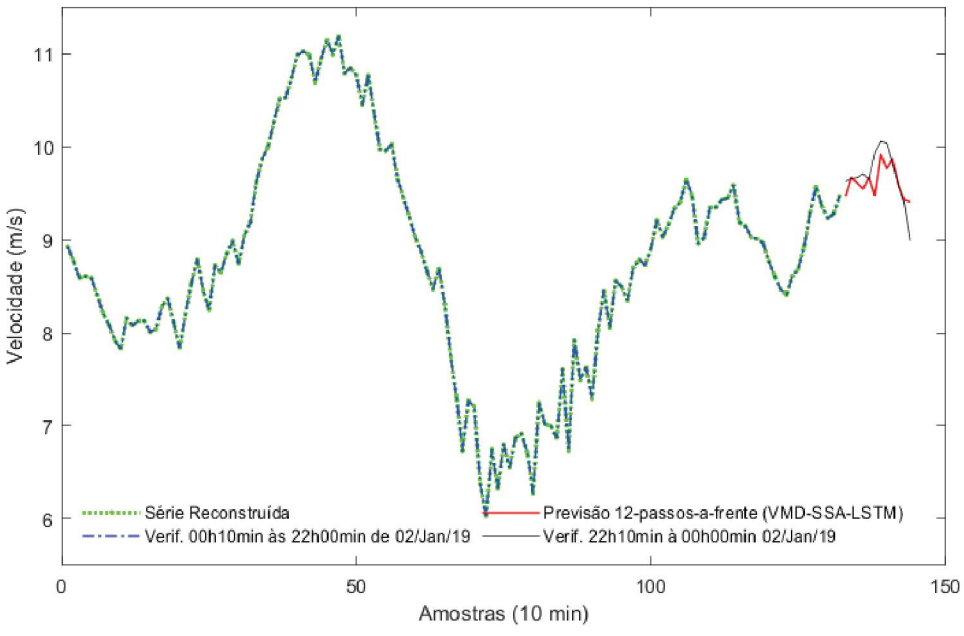
LSTM e compararmos aos resultado da rede neural LSTM.

Figura 98: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 02/jan/2019 através da LSTM



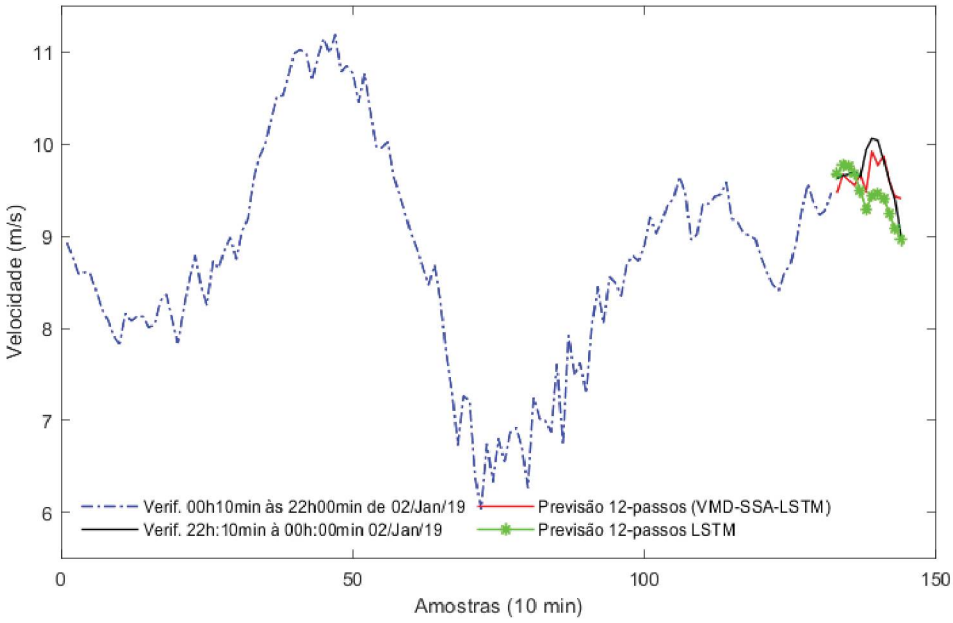
Fonte: O autor(2019).

Figura 99: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 02/jan/19.



Fonte: O autor(2019).

Figura 100: Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 02/jan/19.



Fonte: O autor(2019).

Os indicadores de desempenho do modelo para cada intervalo de previsão utilizando o método VMD-SSA-ARIMA comparado ao ARIMA estão em destaque na Tabela 16. Os resultados para o modelo híbrido VMD-SSA-LSTM *versus* LSTM podem ser encontrados ma Tabela 17.

Conforme resultados da Tabela 16, a qualidade das previsões para o horizonte de 1 a 12-passos-à-frente do modelo VMD-SSA-ARIMA foi superior aos resultados do modelo ARIMA para os dados reais do dia 02/jan/2019, utilizados para prever as duas primeiras horas do dia 03/jan/2019. Mais uma vez, é notório a estabilização da erro proporcionada pela etapa de decomposição.

Tabela 16: Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 03/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-ARIMA	$\hat{y}_{t+1}$	**0,067**	**0,004**	**0,067**	**0,067**	**0,76**	**0,76**
	$\hat{y}_{t+2}$	**0,299**	**0,047**	**0,216**	**0,183**	**2,15**	**2,51**
	$\hat{y}_{t+3}$	**-0,185**	**0,043**	**0,206**	**0,183**	**2,17**	**2,42**
	$\hat{y}_{t+4}$	**-0,293**	**0,053**	**0,231**	**0,211**	**2,49**	**2,71**
	$\hat{y}_{t+5}$	**-0,250**	**0,055**	**0,235**	**0,219**	**2,59**	**2,77**
	$\hat{y}_{t+6}$	**-0,620**	**0,110**	**0,332**	**0,286**	**3,40**	**3,92**
	$\hat{y}_{t+7}$	**-0,919**	**0,215**	**0,464**	**0,376**	**4,44**	**5,46**
	$\hat{y}_{t+8}$	**-0,506**	**0,220**	**0,469**	**0,392**	**4,65**	**5,55**
	$\hat{y}_{t+9}$	**-0,505**	**0,224**	**0,473**	**0,405**	**4,84**	**5,64**
	$\hat{y}_{t+10}$	**-0,567**	**0,234**	**0,484**	**0,421**	**5,07**	**5,79**
	$\hat{y}_{t+11}$	**-0,571**	**0,242**	**0,492**	**0,435**	**5,25**	**5,91**
	$\hat{y}_{t+12}$	**-0,173**	**0,224**	**0,474**	**0,413**	**5,00**	**5,72**
ARIMA	$\hat{y}_{t+1}$	-0,378	0,143	0,378	0,378	4,30	4,30
	$\hat{y}_{t+2}$	-0,499	0,249	0,499	0,439	5,11	5,80
	$\hat{y}_{t+3}$	-0,557	0,310	0,557	0,478	5,62	6,52
	$\hat{y}_{t+4}$	-0,268	0,072	0,268	0,426	5,00	3,14
	$\hat{y}_{t+5}$	0,360	0,130	0,360	0,412	4,86	4,24
	$\hat{y}_{t+6}$	0,940	0,884	0,940	0,500	5,93	11,13
	$\hat{y}_{t+7}$	0,998	0,995	0,998	0,571	6,74	11,76
	$\hat{y}_{t+8}$	1,864	3,473	1,864	0,733	8,72	22,03
	$\hat{y}_{t+9}$	3,343	11,172	3,343	1,023	12,46	39,82
	$\hat{y}_{t+10}$	4,467	19,955	4,467	1,367	16,82	53,49
	$\hat{y}_{t+11}$	5,081	25,816	5,081	1,705	21,02	61,04
	$\hat{y}_{t+12}$	6,232	38,835	6,232	2,082	25,95	75,27

Fonte: O autor(2019).

Os erros de previsão para o modelo VMD-SSA-LSTM foram ligeiramente superiores aos da rede neural LSTM para os 4 primeiros períodos do horizonte de previsão, conforme dados da Tabela 17, porém a característica de estabilização do erro apontada como uma das principais contribuições da técnica proposta, fez com que a partir do 5 intervalo de previsão, os resultados do modelo híbrido superassem os do modelo da rede neural LSTM. É interessante comparar a ordem de grande do MSE para ambas as técnicas, fica claro o poder da técnica hibrida proposta em manter o erro controlado.

Tabela 17: Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 02/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-LSTM	$\hat{y}_{t+1}$	-0,163	0,027	0,163	0,163	1,70	1,70
	$\hat{y}_{t+2}$	0,013	0,013	0,116	0,088	0,92	1,20
	$\hat{y}_{t+3}$	-0,060	0,010	0,101	0,079	0,82	1,04
	$\hat{y}_{t+4}$	-0,157	0,014	0,117	0,098	1,02	1,21
	$\hat{y}_{t+5}$	**0,017**	0,011	0,105	**0,082**	**0,85**	1,09
	$\hat{y}_{t+6}$	**-0,457**	**0,044**	**0,210**	**0,145**	**1,47**	**2,16**
	$\hat{y}_{t+7}$	**-0,142**	**0,041**	**0,202**	**0,144**	**1,47**	**2,07**
	$\hat{y}_{t+8}$	**-0,274**	**0,045**	**0,212**	**0,160**	**1,62**	**2,17**
	$\hat{y}_{t+9}$	**0,045**	**0,040**	**0,200**	**0,148**	**1,49**	**2,05**
	$\hat{y}_{t+10}$	**-0,001**	**0,036**	**0,190**	**0,133**	**1,34**	**1,95**
	$\hat{y}_{t+11}$	**0,040**	**0,033**	**0,182**	**0,124**	**1,26**	**1,87**
	$\hat{y}_{t+12}$	**0,418**	**0,045**	**0,212**	**0,149**	**1,54**	**2,19**
LSTM	$\hat{y}_{t+1}$	**0,050**	**0,002**	**0,050**	**0,050**	**0,52**	**0,52**
	$\hat{y}_{t+2}$	**0,111**	**0,007**	**0,086**	**0,080**	**0,83**	**0,89**
	$\hat{y}_{t+3}$	**0,092**	**0,008**	**0,088**	**0,084**	**0,87**	**0,91**
	$\hat{y}_{t+4}$	**-0,025**	**0,006**	**0,077**	**0,069**	**0,72**	**0,80**
	$\hat{y}_{t+5}$	-0,164	**0,010**	**0,101**	0,088	0,91	**1,04**
	$\hat{y}_{t+6}$	-0,649	0,079	0,281	0,182	1,85	2,89
	$\hat{y}_{t+7}$	-0,620	0,122	0,350	0,244	2,47	3,58
	$\hat{y}_{t+8}$	-0,570	0,148	0,384	0,285	2,87	3,92
	$\hat{y}_{t+9}$	-0,418	0,151	0,388	0,300	3,02	3,96
	$\hat{y}_{t+10}$	-0,346	0,148	0,384	0,304	3,08	3,93
	$\hat{y}_{t+11}$	-0,309	0,143	0,378	0,305	3,10	3,88
	$\hat{y}_{t+12}$	-0,033	0,131	0,362	0,282	2,87	3,74

Fonte: O autor(2019).

6.2.3 PREVISÕES PARA OS DADOS OBSERVADOS EM 03/JAN/2019

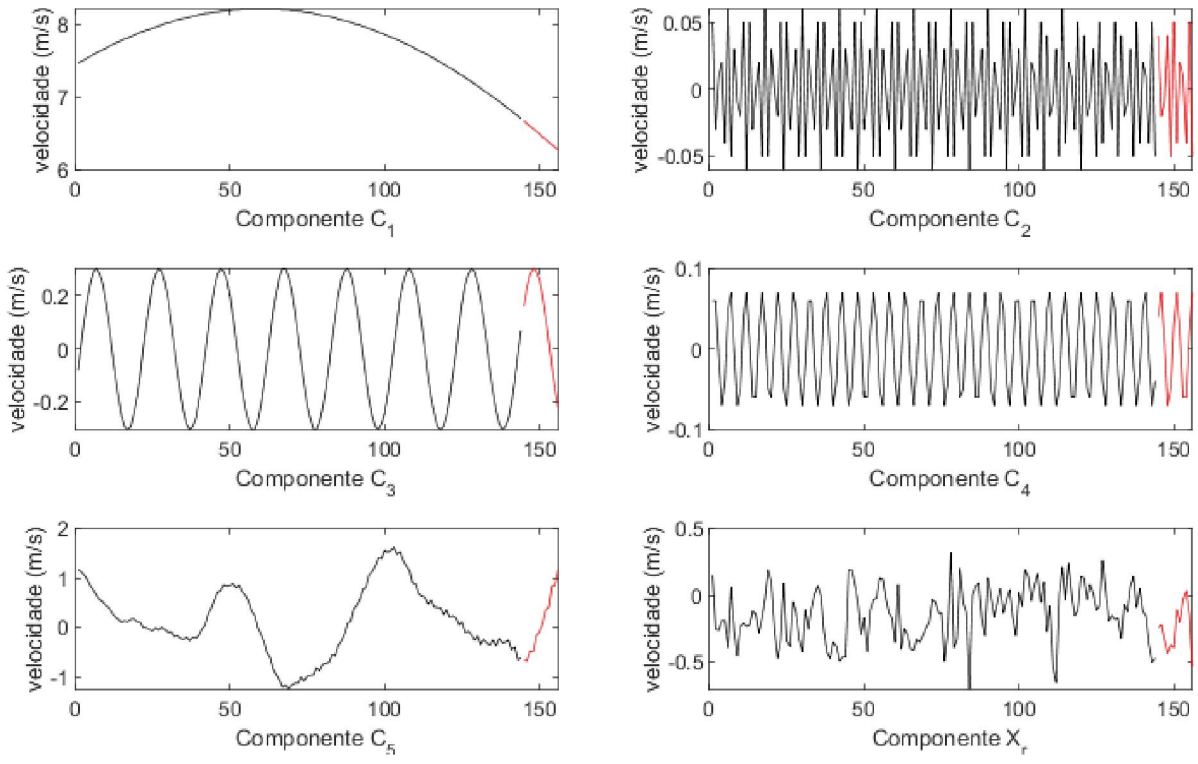
A série verificada no dia 03 de janeiro de 2019 possui dois picos e em sequência um afundamento vertiginoso, características essas relativa a rajadas de vento. Vale aqui ressaltar a importância de um modelo de previsão ser capaz de identificar corretamente essas variações, pois essas rajadas intermitentes podem provocar o colapso do aerogerador (danos estruturais), ou caso a amplitude não supere valores limites de resistência da estrutura do aerogerador, tem-se ainda o impacto no sistema interligado de distribuição. Esse impacto é hoje uma das grandes preocupações do ONS, pois a rampa de aumento de produção de um conjunto de parques eólicos sob efeito dessa rajada pode levar ao colapso

do Sistema de Transmissão, com desligamentos de linhas e apagões. Estes problemas estão associados ao excesso de produção em um curto espaço de tempo (o consumo do sistema já está em equilíbrio) e a impossibilidade do sistema elétrico absorver essa variação tão brusca.

Feito esse breve comentário, apenas para ressaltar a importância da previsibilidade de séries de velocidades de vento, os resultados relativos a série verificada no dia em questão são apresentadas nas Figuras 101 à 103 para as técnicas VMD-SSA-ARIMA e ARIMA, e nas Figuras 104 à 106 para as variantes VMD-SSA-LSTM e LSTM. Os índices de qualidade do modelo estão em destaque nas Tabelas 18 e 19.

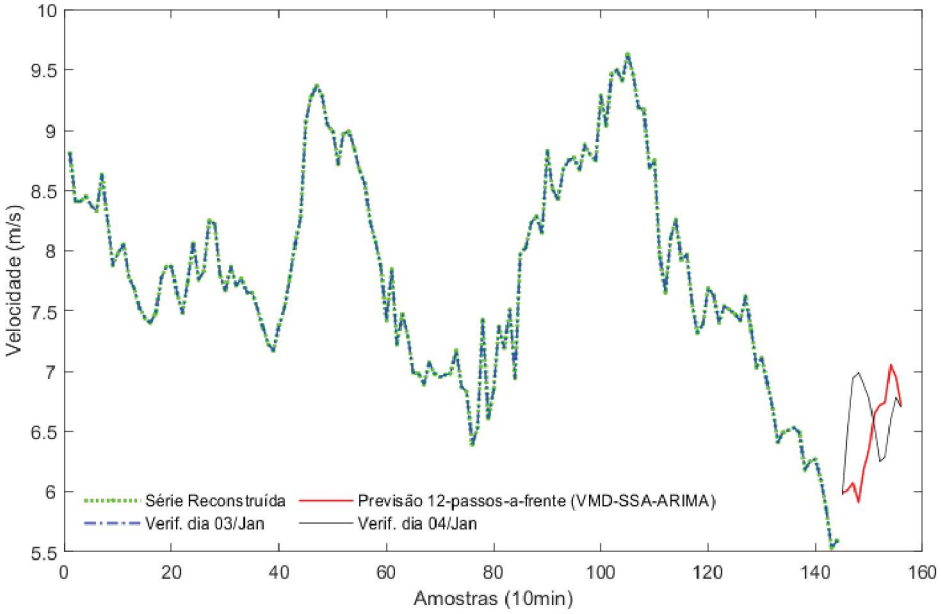
Ao comparar o resultado da previsão na Figura 103 é evidente que o modelo ARIMA não foi capaz de representar o comportamento da série de velocidades de vento, após a rampa de descida. Observe que o modelo apenas continuou com a tendência de queda pelos próximos 12 passos-à-frente, enquanto que o modelo híbrido VMD-SSA-ARIMA identificou corretamente a nova tendência da série, prevendo de forma satisfatória a continuação da série, com desvios apenas na amplitude do sinal previsto.

Figura 101: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 03/jan/2019 através do método de Box & Jenkins



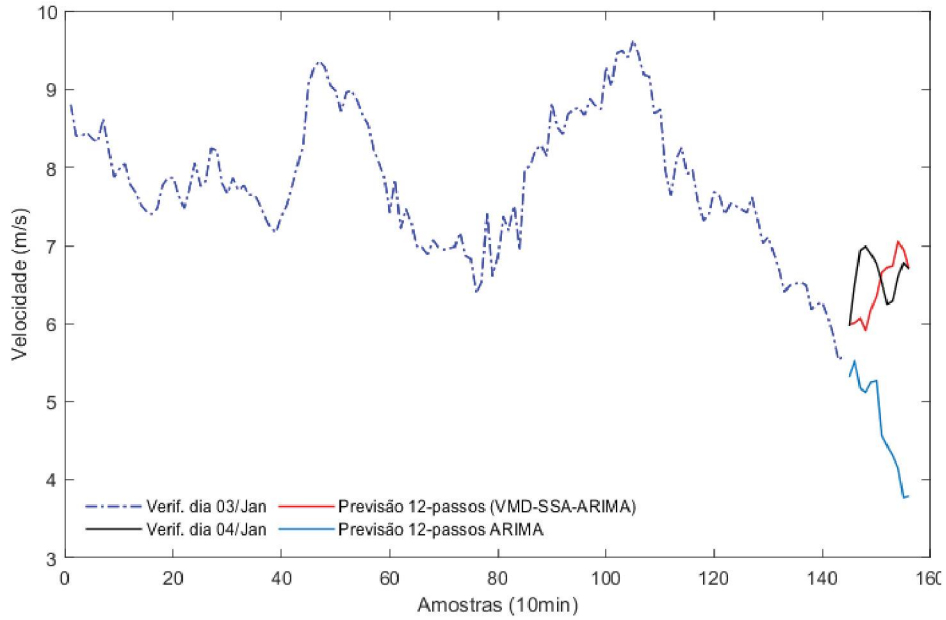
Fonte: O autor(2019).

Figura 102: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 03 e 04/jan/19.



Fonte: O autor(2019).

Figura 103: Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 03 e 04/jan/19.

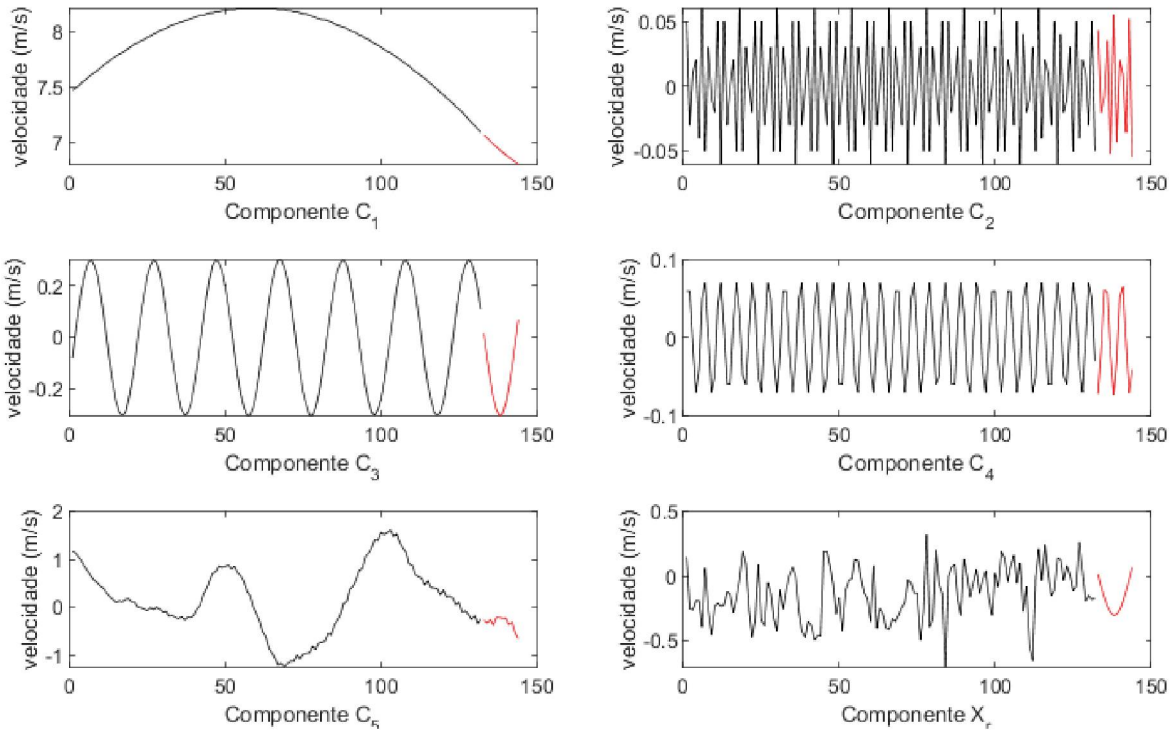


Fonte: O autor(2019).

As previsões resultantes do modelo VMD-SSA-LSTM para cada componente são apresentadas na Figura 104, enquanto que a comparação do resultado final é apresentado na Figura 106. O modelo híbrido superou os resultados da rede neural LSTM para todos

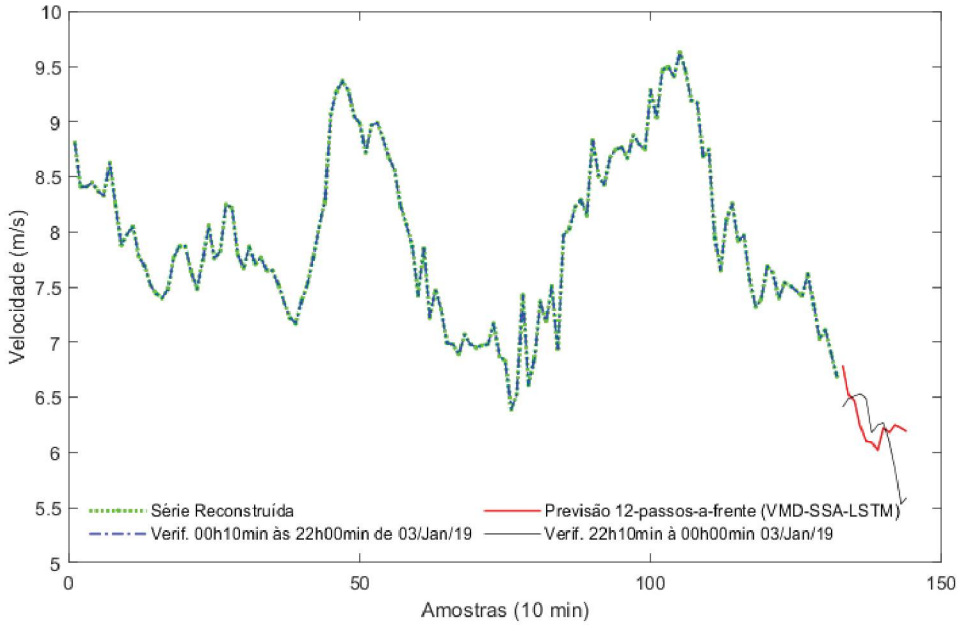
os passos de previsão a frente, fato observado ao comparar os valores de todas as métricas utilizadas para validar o modelo.

Figura 104: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 03/jan/2019 através da LSTM



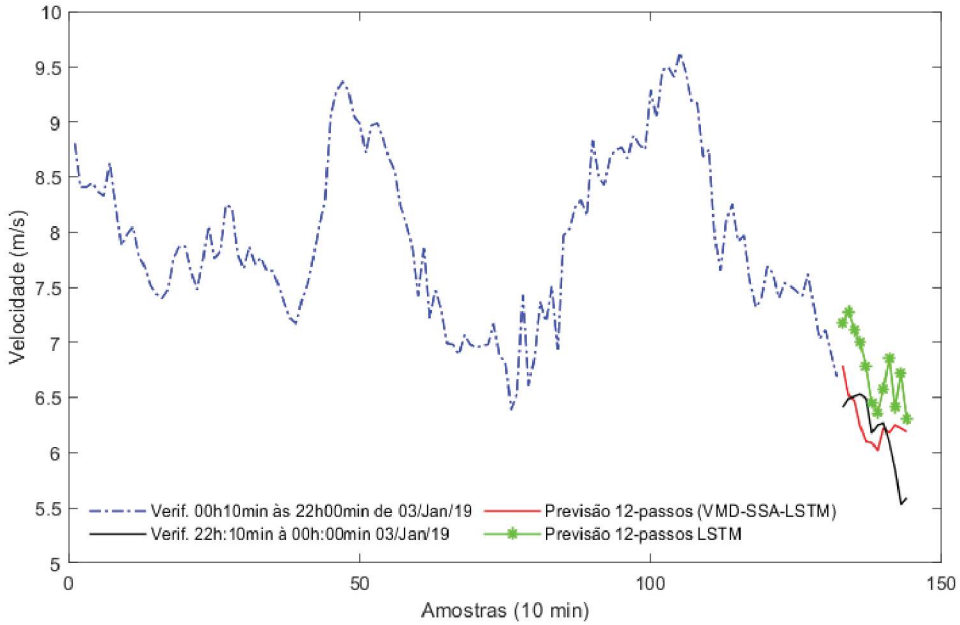
Fonte: O autor(2019).

Figura 105: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 03/jan/19.



Fonte: O autor(2019).

Figura 106: Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 03/jan/19.



Fonte: O autor(2019).

A Tabela 18 apresenta a comparação entre as medidas de precisão utilizada para avaliar a qualidade do modelo em termo de minimização do erro. Para todos os horizontes de previsão o modelo VMD-SSA-ARIMA superou os resultados do modelo ARIMA.

Tabela 18: Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 04/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-ARIMA	$\hat{y}_{t+1}$	**0,024**	**0,001**	**0,024**	**0,024**	**0,40**	**0,40**
	$\hat{y}_{t+2}$	**-0,504**	**0,127**	**0,357**	**0,264**	**4,07**	**5,72**
	$\hat{y}_{t+3}$	**-0,872**	**0,338**	**0,582**	**0,467**	**6,90**	**8,99**
	$\hat{y}_{t+4}$	**-1,077**	**0,544**	**0,737**	**0,619**	**9,03**	**11,17**
	$\hat{y}_{t+5}$	**-0,707**	**0,535**	**0,731**	**0,637**	**9,28**	**10,98**
	$\hat{y}_{t+6}$	**-0,422**	**0,475**	**0,690**	**0,601**	**8,77**	**10,32**
	$\hat{y}_{t+7}$	**0,121**	**0,410**	**0,640**	**0,532**	**7,78**	**9,61**
	$\hat{y}_{t+8}$	**0,470**	**0,386**	**0,621**	**0,525**	**7,75**	**9,40**
	$\hat{y}_{t+9}$	**0,450**	**0,366**	**0,605**	**0,516**	**7,68**	**9,20**
	$\hat{y}_{t+10}$	**0,449**	**0,349**	**0,591**	**0,510**	**7,59**	**8,99**
	$\hat{y}_{t+11}$	**0,166**	**0,320**	**0,566**	**0,478**	**7,13**	**8,58**
	$\hat{y}_{t+12}$	**0,011**	**0,293**	**0,542**	**0,439**	**6,55**	**8,20**
ARIMA	$\hat{y}_{t+1}$	-0,646	0,417	0,646	0,646	10,82	10,82
	$\hat{y}_{t+2}$	-0,995	0,703	0,839	0,820	13,05	13,44
	$\hat{y}_{t+3}$	-1,770	1,513	1,230	1,137	17,20	19,00
	$\hat{y}_{t+4}$	-1,865	2,004	1,416	1,319	19,57	21,44
	$\hat{y}_{t+5}$	-1,637	2,140	1,463	1,383	20,41	21,96
	$\hat{y}_{t+6}$	-1,501	2,158	1,469	1,402	20,70	22,00
	$\hat{y}_{t+7}$	-1,958	2,397	1,548	1,482	22,03	23,26
	$\hat{y}_{t+8}$	-1,819	2,511	1,585	1,524	22,91	23,99
	$\hat{y}_{t+9}$	-1,977	2,667	1,633	1,574	23,86	24,85
	$\hat{y}_{t+10}$	-2,450	3,000	1,732	1,662	25,18	26,35
	$\hat{y}_{t+11}$	-3,009	3,551	1,884	1,784	26,93	28,58
	$\hat{y}_{t+12}$	-2,912	3,961	1,990	1,878	28,31	30,15

Fonte: O autor(2019).

Os resultados de avaliação dos erros para o modelo VMD-SSA-LSTM *versus* os da rede neural LSTM estão apresentados na Tabela 19. O modelo proposto apresentou melhor desempenho nas previsões quando comparado ao LSTM puro, para todos os intervalos do horizonte de previsão da série temporal de velocidades de vento observada no início do dia 04/janeiro.

Tabela 19: Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 03/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-LSTM	$\hat{y}_{t+1}$	**0,379**	**0,143**	**0,379**	**0,379**	**5,91**	**5,91**
	$\hat{y}_{t+2}$	**0,026**	**0,072**	**0,268**	**0,202**	**3,15**	**4,16**
	$\hat{y}_{t+3}$	**-0,040**	**0,049**	**0,220**	**0,148**	**2,31**	**3,41**
	$\hat{y}_{t+4}$	**-0,293**	**0,058**	**0,240**	**0,184**	**2,85**	**3,71**
	$\hat{y}_{t+5}$	**-0,385**	**0,076**	**0,276**	**0,224**	**3,47**	**4,25**
	$\hat{y}_{t+6}$	**-0,088**	**0,065**	**0,254**	**0,202**	**3,13**	**3,95**
	$\hat{y}_{t+7}$	**-0,233**	**0,063**	**0,251**	**0,206**	**3,21**	**3,92**
	$\hat{y}_{t+8}$	**-0,053**	**0,056**	**0,236**	**0,187**	**2,92**	**3,69**
	$\hat{y}_{t+9}$	**0,092**	**0,050**	**0,224**	**0,177**	**2,76**	**3,53**
	$\hat{y}_{t+10}$	**0,399**	**0,061**	**0,248**	**0,199**	**3,17**	**3,93**
	$\hat{y}_{t+11}$	**0,688**	**0,099**	**0,314**	**0,243**	**4,01**	**5,04**
	$\hat{y}_{t+12}$	**0,600**	**0,121**	**0,347**	**0,273**	**4,57**	**5,62**
LSTM	$\hat{y}_{t+}$	0,769	0,591	0,769	0,769	12,00	12,00
	$\hat{y}_{t+2}$	0,794	0,611	0,782	0,782	12,12	12,12
	$\hat{y}_{t+3}$	0,599	0,527	0,726	0,721	11,14	11,22
	$\hat{y}_{t+4}$	0,467	0,449	0,670	0,657	10,14	10,34
	$\hat{y}_{t+5}$	0,292	0,377	0,614	0,584	9,01	9,46
	$\hat{y}_{t+6}$	0,274	0,326	0,571	0,532	8,25	8,88
	$\hat{y}_{t+7}$	0,110	0,282	0,531	0,472	7,33	8,28
	$\hat{y}_{t+8}$	0,312	0,258	0,508	0,452	7,03	7,95
	$\hat{y}_{t+9}$	0,767	0,295	0,543	0,487	7,65	8,55
	$\hat{y}_{t+10}$	0,560	0,297	0,545	0,494	7,84	8,64
	$\hat{y}_{t+12}$	1,195	0,400	0,632	0,558	9,09	10,14
	$\hat{y}_{t+12}$	0,724	0,410	0,640	0,572	9,42	10,36

Fonte: O autor(2019).

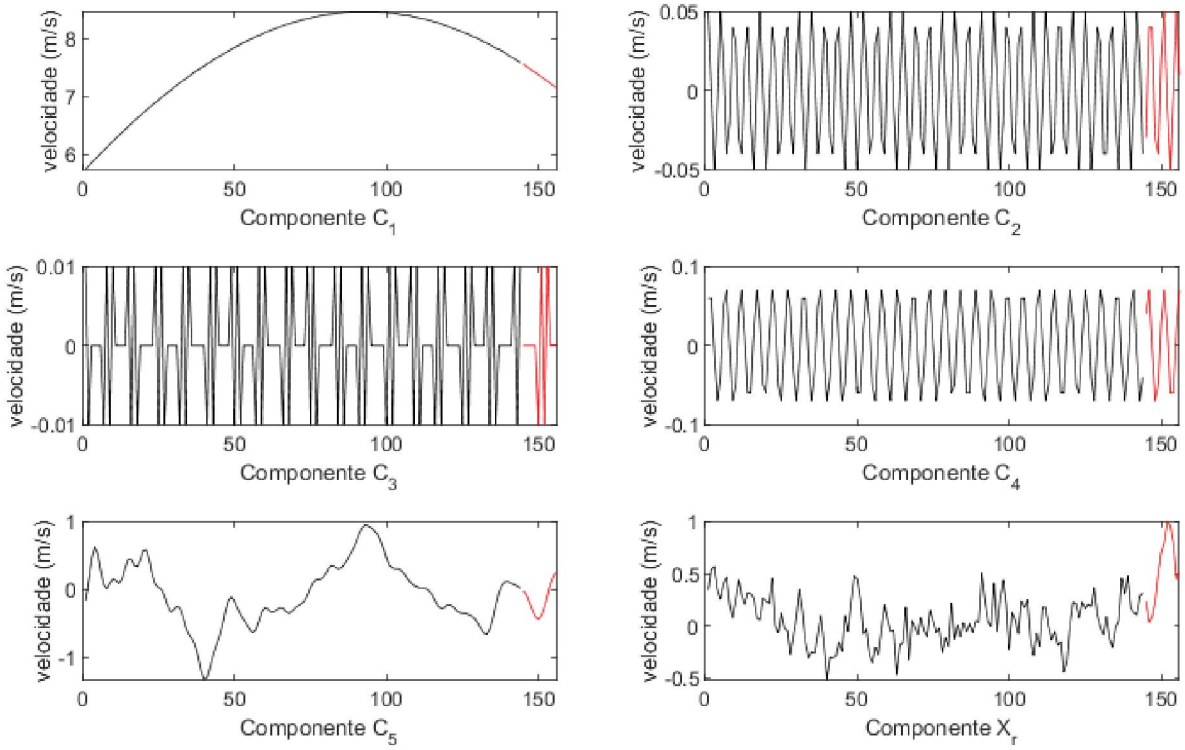
6.2.4 PREVISÕES PARA OS DADOS OBSERVADOS EM 04/JAN/2019

Os resultados para cada série são apresentados nas Figuras 107 à 112. As Tabelas 20 e 21 apresentam as comparações dos índices de desempenho para cada horizonte de previsão.

As previsões obtidas como o modelo proposto superou as obtidas pelo modelo ARIMA em todos os horizontes de previsão, conforme resultados da Figura 109. Pode-se observar que os resultados das previsões do modelo ARIMA manteve a tendência de queda observada na série original. Já o modelo VMD-SSA-ARIMA detectou a mudança da tendência de forma correta, de forma que as previsões dos 12-passos-à-frente estão bem

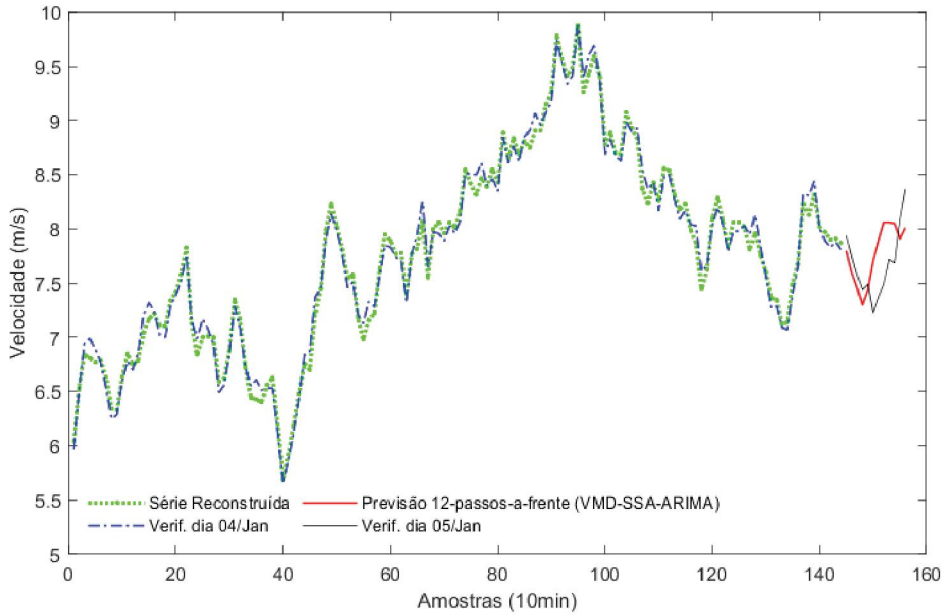
alinhadas com os valores reais da série.

Figura 107: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 04/jan/2019 através do método de Box & Jenkins



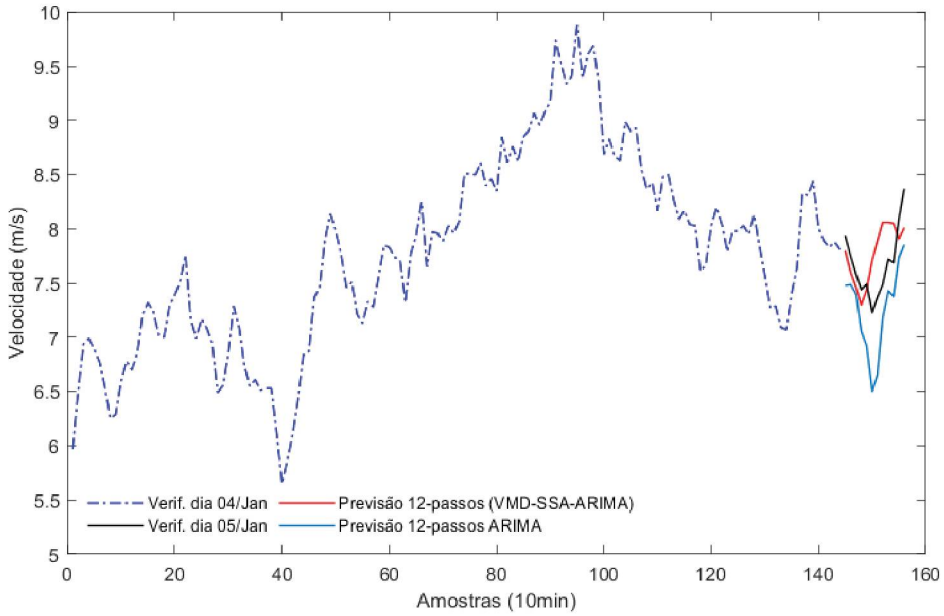
Fonte: O autor(2019).

Figura 108: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 04 e 05/jan/19.



Fonte: O autor(2019).

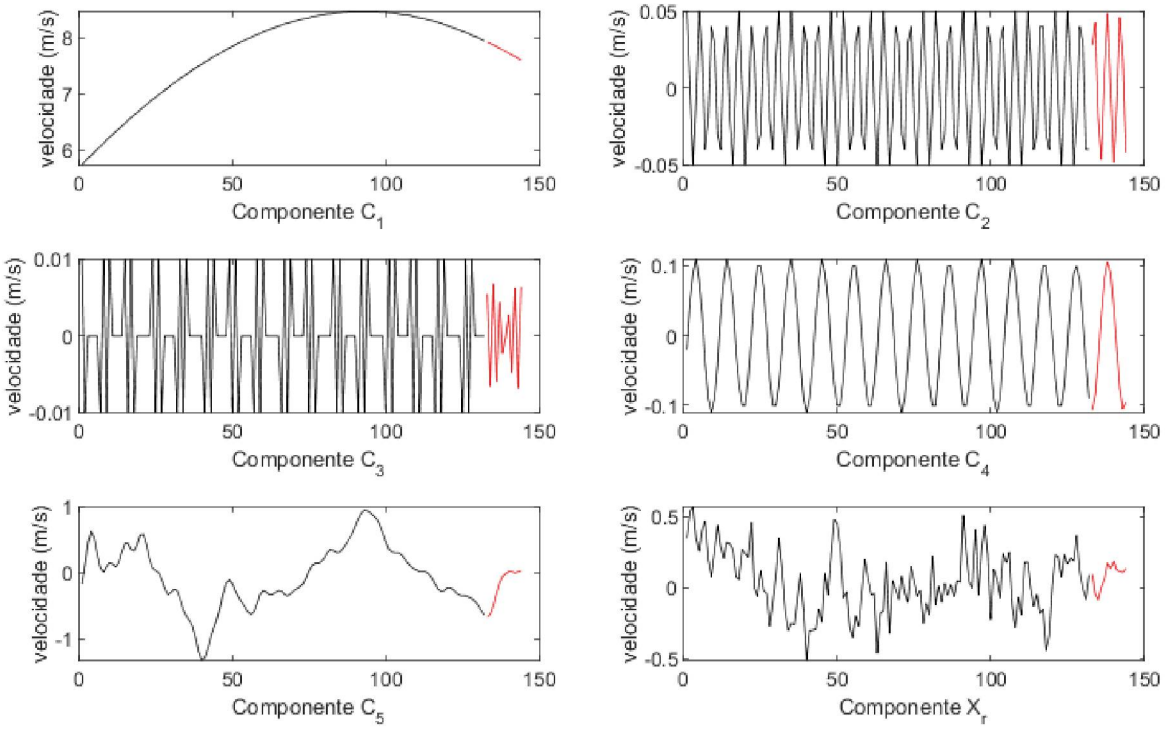
Figura 109: Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 04 e 05/jan/19.



Fonte: O autor(2019).

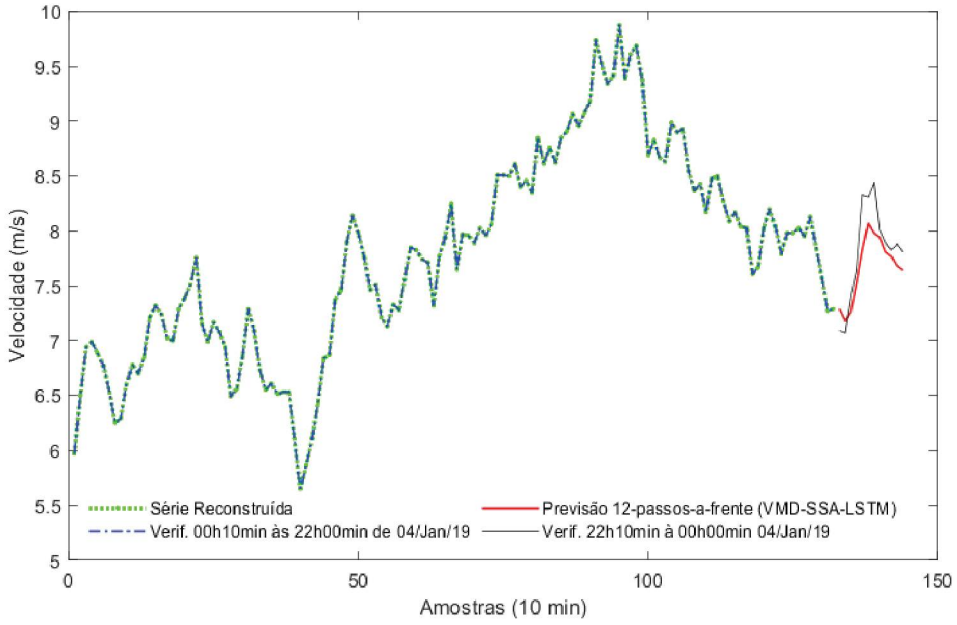
Ao avaliar o resultado das previsões para cada componente na Figura 110, pode-se notar que as previsões para a componente C_3 não foram adequadas, ao contrário por exemplo, dos resultados obtidos para a mesma componente com o modelo VMD-SSA-ARIMA destacados na Figura 107. Porém o impacto dessa componente no resultado final foi pouco sentido, devido a pequena amplitude desta componente, esse fato é comprovado pelos valores dos índices de desempenho do modelo avaliado foram superiores aos obtidos pela rede neural LSTM.

Figura 110: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 04/jan/2019 através da LSTM



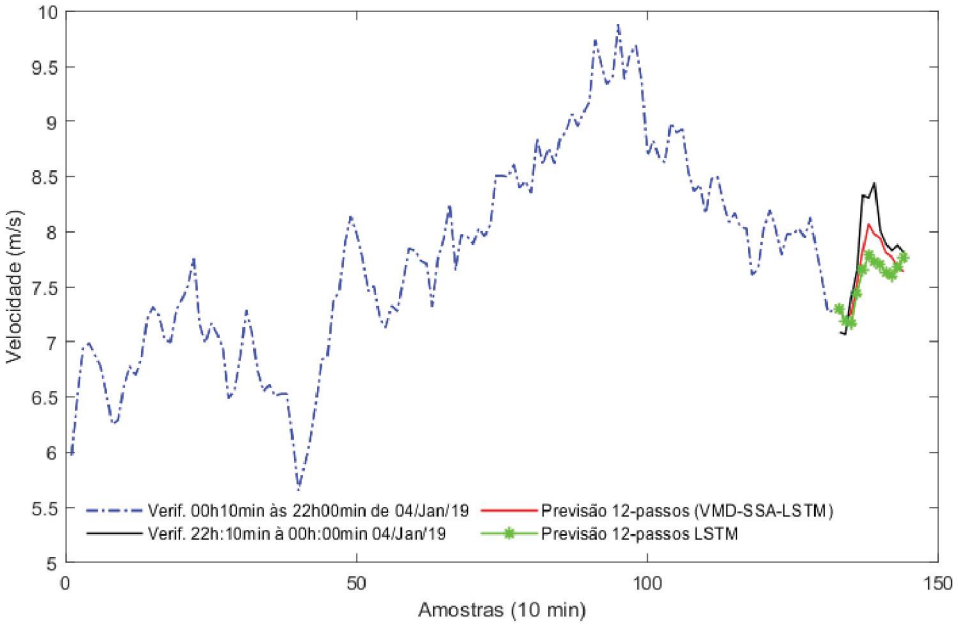
Fonte: O autor(2019).

Figura 111: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 04/jan/19.



Fonte: O autor(2019).

Figura 112: Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 04/jan/19.



Fonte: O autor(2019).

Os índices de desempenho para os 12 passos de previsão estão destacados na Tabela 20 para os modelos VMD-SSA-ARIMA comparado ao ARIMA. Importante observar que os valores de MAE e MAPE obtidos com o modelo proposto VMD-SSA-ARIMA representam para cada horizonte de previsão aproximadamente 50% do valor da ARIMA.

Tabela 20: Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 05/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-ARIMA	$\hat{y}_{t+1}$	**-0,139**	**0,019**	**0,139**	**0,139**	**1,75**	**1,75**
	$\hat{y}_{t+2}$	**-0,147**	**0,020**	**0,143**	**0,143**	**1,82**	**1,82**
	$\hat{y}_{t+3}$	**-0,122**	**0,019**	**0,136**	**0,136**	**1,75**	**1,76**
	$\hat{y}_{t+4}$	**-0,144**	**0,019**	**0,138**	**0,138**	**1,80**	**1,80**
	$\hat{y}_{t+5}$	**-0,047**	**0,016**	**0,125**	**0,120**	**1,56**	**1,64**
	$\hat{y}_{t+6}$	**0,478**	**0,051**	**0,226**	**0,179**	**2,40**	**2,99**
	$\hat{y}_{t+7}$	**0,517**	**0,082**	**0,286**	**0,228**	**3,06**	**3,80**
	$\hat{y}_{t+8}$	**0,573**	**0,113**	**0,336**	**0,271**	**3,64**	**4,46**
	$\hat{y}_{t+9}$	**0,337**	**0,113**	**0,336**	**0,278**	**3,72**	**4,45**
	$\hat{y}_{t+10}$	**0,360**	**0,115**	**0,339**	**0,286**	**3,81**	**4,47**
	$\hat{y}_{t+11}$	**-0,175**	**0,107**	**0,327**	**0,276**	**3,66**	**4,29**
	$\hat{y}_{t+12}$	**-0,364**	**0,109**	**0,330**	**0,284**	**3,72**	**4,30**
ARIMA	$\hat{y}_{t+1}$	-0,463	0,214	0,463	0,463	5,83	5,83
	$\hat{y}_{t+2}$	-0,263	0,142	0,376	0,363	4,61	4,80
	$\hat{y}_{t+3}$	-0,183	0,106	0,325	0,303	3,88	4,19
	$\hat{y}_{t+4}$	-0,385	0,116	0,341	0,323	4,20	4,44
	$\hat{y}_{t+5}$	-0,571	0,158	0,398	0,373	4,89	5,21
	$\hat{y}_{t+6}$	-0,728	0,220	0,469	0,432	5,75	6,20
	$\hat{y}_{t+7}$	-0,715	0,262	0,512	0,473	6,31	6,78
	$\hat{y}_{t+8}$	-0,310	0,241	0,491	0,452	6,04	6,52
	$\hat{y}_{t+9}$	-0,293	0,224	0,473	0,435	5,79	6,26
	$\hat{y}_{t+10}$	-0,312	0,211	0,460	0,422	5,62	6,07
	$\hat{y}_{t+11}$	-0,354	0,203	0,451	0,416	5,51	5,92
	$\hat{y}_{t+12}$	-0,516	0,209	0,457	0,424	5,56	5,95

Fonte: O autor(2019).

Da mesma forma que os resultados obtidos pelo modelo VMD-SSA-ARIMA, os resultados para o modelo híbrido com a rede neural LSTM também superaram os resultados do modelo de previsão que utiliza apenas a rede neural LSTM. Ao comparar os valores do MSE para ambos os modelos, a característica de estabilidade dos erros é notória para o modelo que utiliza a decomposição da série temporal proposta nesta tese.

Tabela 21: Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas últimas horas do dia 04/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-LSTM	$\hat{y}_{t+1}$	**0,200**	**0,040**	**0,200**	**0,200**	**2,82**	**2,82**
	$\hat{y}_{t+2}$	**0,114**	**0,027**	**0,163**	**0,157**	**2,22**	**2,30**
	$\hat{y}_{t+3}$	**-0,142**	**0,024**	**0,156**	**0,152**	**2,12**	**2,17**
	$\hat{y}_{t+4}$	**-0,126**	**0,022**	**0,149**	**0,146**	**2,00**	**2,05**
	$\hat{y}_{t+5}$	**-0,498**	**0,067**	**0,260**	**0,216**	**2,80**	**3,46**
	$\hat{y}_{t+6}$	**-0,242**	**0,066**	**0,257**	**0,220**	**2,82**	**3,36**
	$\hat{y}_{t+7}$	**-0,461**	**0,087**	**0,295**	**0,255**	**3,19**	**3,80**
	$\hat{y}_{t+8}$	**-0,073**	**0,077**	**0,277**	**0,232**	**2,91**	**3,56**
	$\hat{y}_{t+9}$	**-0,076**	**0,069**	**0,262**	**0,215**	**2,69**	**3,36**
	$\hat{y}_{t+10}$	**-0,062**	**0,062**	**0,250**	**0,199**	**2,50**	**3,20**
	$\hat{y}_{t+11}$	**-0,202**	**0,060**	**0,246**	**0,200**	**2,51**	**3,15**
	$\hat{y}_{t+12}$	**-0,172**	**0,058**	**0,240**	**0,197**	**2,48**	**3,08**
LSTM	$\hat{y}_{t+1}$	0,206	0,043	0,206	0,206	2,91	2,91
	$\hat{y}_{t+2}$	0,124	0,029	0,170	0,165	2,34	2,41
	$\hat{y}_{t+3}$	-0,228	0,037	0,192	0,186	2,58	2,67
	$\hat{y}_{t+4}$	-0,213	0,039	0,197	0,193	2,63	2,70
	$\hat{y}_{t+5}$	-0,676	0,123	0,350	0,290	3,73	4,66
	$\hat{y}_{t+6}$	-0,522	0,148	0,384	0,328	4,16	5,03
	$\hat{y}_{t+7}$	-0,710	0,199	0,446	0,383	4,76	5,75
	$\hat{y}_{t+8}$	-0,311	0,186	0,431	0,374	4,65	5,54
	$\hat{y}_{t+9}$	-0,262	0,173	0,416	0,361	4,51	5,33
	$\hat{y}_{t+10}$	-0,226	0,161	0,401	0,348	4,34	5,14
	$\hat{y}_{t+11}$	-0,200	0,150	0,387	0,334	4,18	4,95
	$\hat{y}_{t+12}$	-0,042	0,137	0,371	0,310	3,88	4,75

Fonte: O autor(2019).

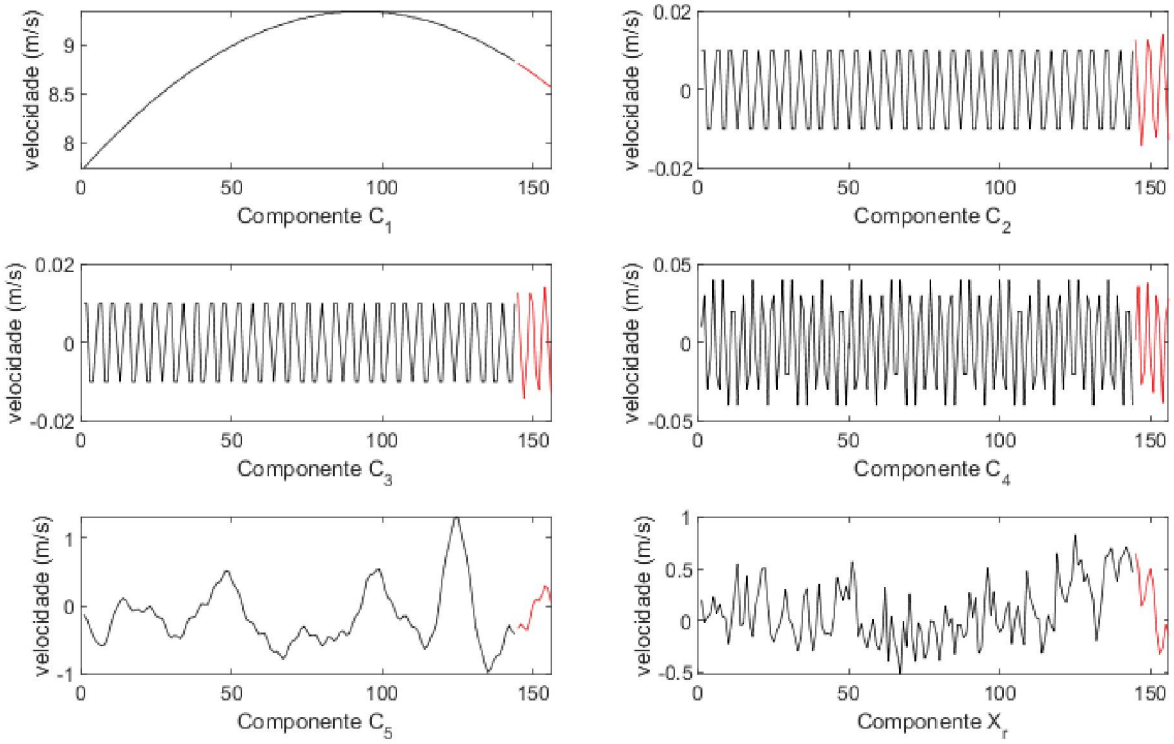
6.2.5 PREVISÕES PARA OS DADOS OBSERVADOS EM 05/JAN/2019

Nesta seção apresentamos os resultados das previsões a cada 10 min da série de velocidades de vento para o dia 05/jan/2019. Os modelos aplicados para a previsão seguem o padrão dos dias anteriores. Os resultados das previsões são apresentados nas Figuras 113 à 118. As Tabelas 22 e 23 sintetizam o desempenho dos modelos de previsão em continuar a série, baseado apenas no comportamento passado.

O resultado obtido com o modelo VMD-SSA-ARIMA se aproximou dos valores reais tanto em amplitude dos valores, como na representação da tendência e da sazonalidade entre períodos. O mesmo não ocorreu, por exemplo, ao aplicar o modelo ARIMA

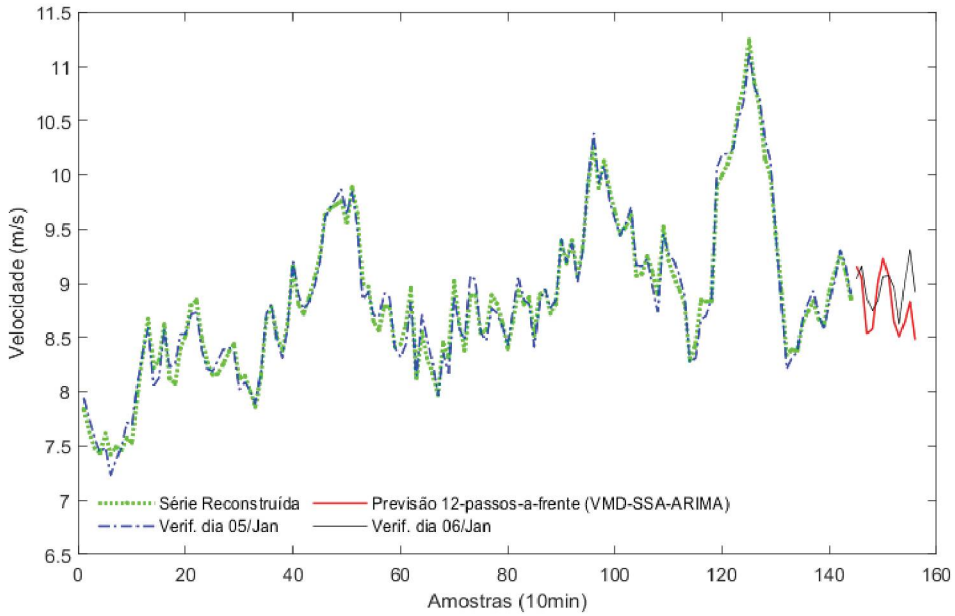
diretamente a série original, cujo resultado divergiu bastante do comportamento da série real, conforme resultados apresentados na Figura 118.

Figura 113: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 05/jan/2019 através do método de Box & Jenkins



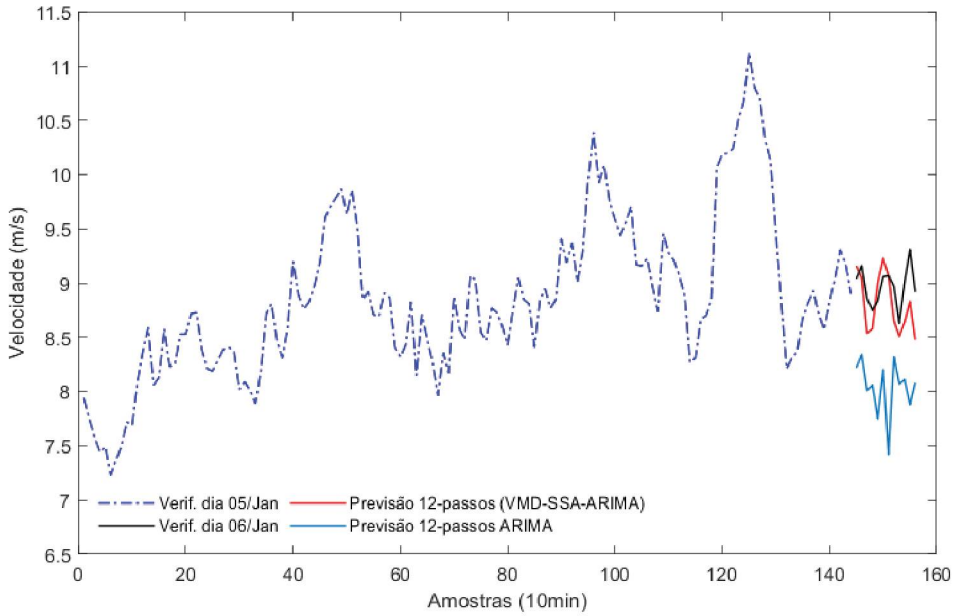
Fonte: O autor(2019).

Figura 114: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 04 e 05/jan/19.



Fonte: O autor(2019).

Figura 115: Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 05 e 06/jan/19.

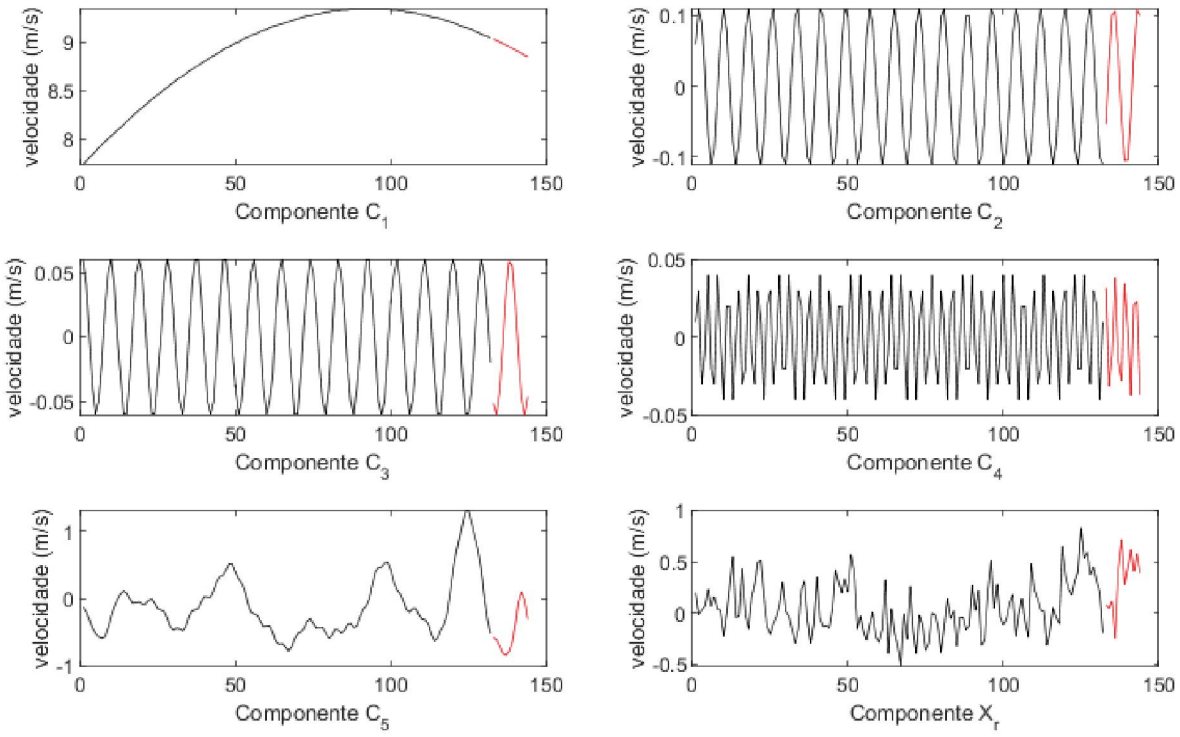


Fonte: O autor(2019).

O modelo VMD-SSA-LSTM, assim como o VMD-SSA-ARIMA, também produziu previsões alinhadas ao comportamento da série real, enquanto que os resultados obtidos pela rede neural LSTM, aplicada diretamente aos dados, não foram adequados, conforme

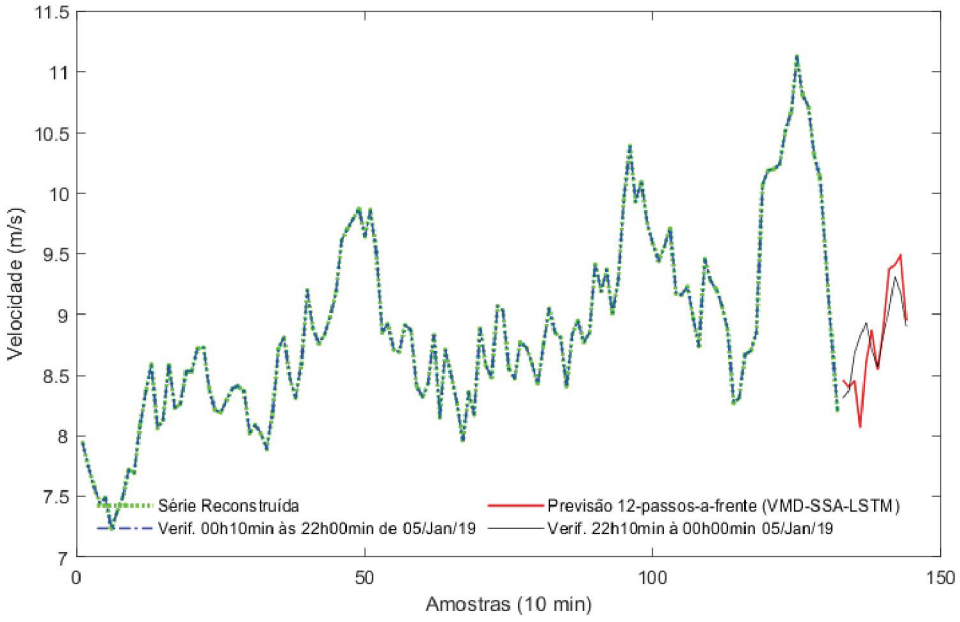
Figura 118.

Figura 116: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 05/jan/2019 através da LSTM



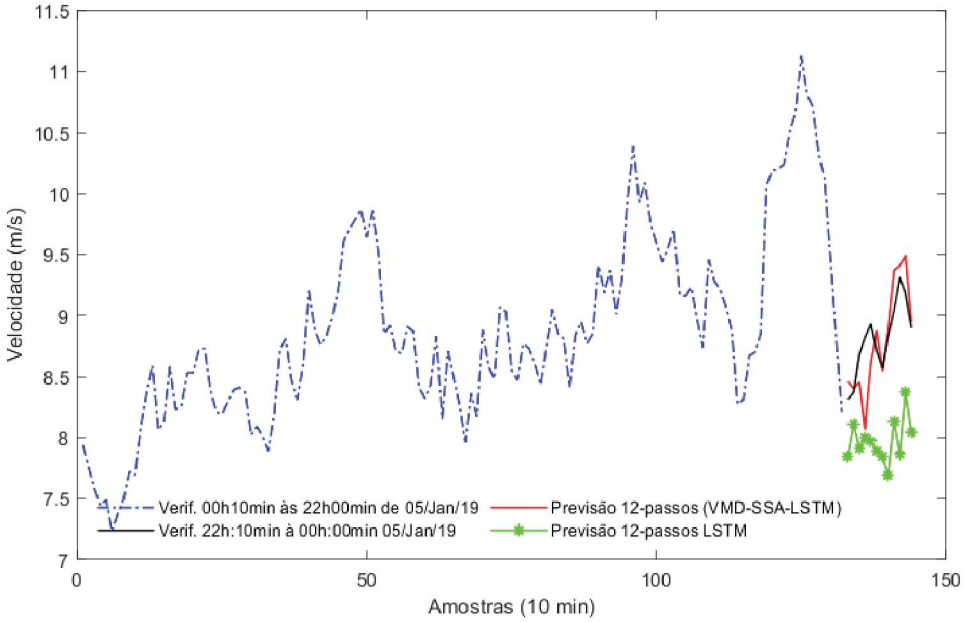
Fonte: O autor(2019).

Figura 117: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas no dia 05/jan/19.



Fonte: O autor(2019).

Figura 118: Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores 10min observado no dia 05/jan/19.



Fonte: O autor(2019).

Os índices de qualidade para o modelo ARIMA são relativamente elevados, com MSE superando a unidade em alguns casos, enquanto que para o VMD-SSA-ARIMA, os valores máximos do MSE foi da ordem $0,074$. Desta forma, fica evidente a robustez alcançada pelo modelo previsor ARIMA ao ser acoplado ao modelo de decomposição proposto nesta tese.

Tabela 22: Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA para as duas horas iniciais do dia 06/Jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-ARIMA	$\hat{y}_{t+1}$	**0,120**	**0,014**	**0,120**	**0,120**	**1,33**	**1,33**
	$\hat{y}_{t+2}$	**-0,105**	**0,013**	**0,113**	**0,113**	**1,24**	**1,24**
	$\hat{y}_{t+3}$	**-0,322**	**0,043**	**0,208**	**0,182**	**2,04**	**2,30**
	$\hat{y}_{t+4}$	**-0,170**	**0,039**	**0,199**	**0,179**	**2,01**	**2,22**
	$\hat{y}_{t+5}$	**0,156**	**0,036**	**0,191**	**0,175**	**1,96**	**2,14**
	$\hat{y}_{t+6}$	**0,173**	**0,035**	**0,188**	**0,174**	**1,95**	**2,10**
	$\hat{y}_{t+7}$	**-0,002**	**0,030**	**0,174**	**0,150**	**1,68**	**1,94**
	$\hat{y}_{t+8}$	**-0,319**	**0,039**	**0,198**	**0,171**	**1,91**	**2,21**
	$\hat{y}_{t+9}$	**-0,124**	**0,037**	**0,191**	**0,166**	**1,86**	**2,14**
	$\hat{y}_{t+10}$	**0,370**	**0,047**	**0,216**	**0,186**	**2,08**	**2,42**
	$\hat{y}_{t+11}$	**-0,483**	**0,064**	**0,252**	**0,213**	**2,37**	**2,81**
	$\hat{y}_{t+12}$	**-0,435**	**0,074**	**0,272**	**0,232**	**2,58**	**3,03**
ARIMA	$\hat{y}_{t+1}$	-0,823	0,678	0,823	0,823	9,11	9,11
	$\hat{y}_{t+2}$	-0,822	0,677	0,823	0,823	9,04	9,04
	$\hat{y}_{t+3}$	-0,853	0,694	0,833	0,833	9,24	9,23
	$\hat{y}_{t+4}$	-0,691	0,640	0,800	0,797	8,90	8,93
	$\hat{y}_{t+5}$	-1,088	0,749	0,865	0,856	9,58	9,69
	$\hat{y}_{t+6}$	-0,860	0,747	0,864	0,856	9,57	9,66
	$\hat{y}_{t+7}$	-1,646	1,027	1,014	0,969	10,79	11,30
	$\hat{y}_{t+8}$	-0,651	0,952	0,976	0,929	10,35	10,88
	$\hat{y}_{t+9}$	-0,562	0,881	0,939	0,888	9,92	10,51
	$\hat{y}_{t+10}$	-0,901	0,874	0,935	0,890	9,93	10,46
	$\hat{y}_{t+11}$	-1,428	0,980	0,990	0,939	10,42	11,03
	$\hat{y}_{t+1}$	-0,840	0,957	0,978	0,930	10,34	10,91

Fonte: O autor(2019).

Tal qual os resultados apresentados na Tabela 22 para destacar o modelo VMD-SSA-ARIMA em relação ao ARIMA, os resultados da Tabela 23 apresentam o comparativo entre os modelos VMD-SSA-LSTM e a rede neural LSTM. Os valores do MSE para a rede neural LSTM apresentam ordem de grandeza de 10 vezes os valores obtidos para o modelo híbrido. Outros índices tais como MAE e MAPE também ratificam a superioridade do modelo proposto.

Tabela 23: Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM para as duas horas finais do dia 05/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-lSTM	$\hat{y}_{t+1}$	**0,145**	**0,021**	**0,145**	**0,145**	**1,75**	**1,75**
	$\hat{y}_{t+2}$	**0,027**	**0,011**	**0,105**	**0,086**	**1,03**	**1,25**
	$\hat{y}_{t+3}$	**-0,228**	**0,025**	**0,157**	**0,133**	**1,56**	**1,85**
	$\hat{y}_{t+4}$	**-0,755**	**0,161**	**0,401**	**0,289**	**3,31**	**4,69**
	$\hat{y}_{t+5}$	**-0,319**	**0,149**	**0,386**	**0,295**	**3,36**	**4,48**
	$\hat{y}_{t+6}$	**0,156**	**0,128**	**0,358**	**0,272**	**3,10**	**4,15**
	$\hat{y}_{t+7}$	**-0,022**	**0,110**	**0,332**	**0,236**	**2,69**	**3,84**
	$\hat{y}_{t+8}$	**0,064**	**0,097**	**0,311**	**0,214**	**2,45**	**3,59**
	$\hat{y}_{t+9}$	**0,332**	**0,098**	**0,313**	**0,227**	**2,58**	**3,60**
	$\hat{y}_{t+10}$	**0,101**	**0,089**	**0,299**	**0,215**	**2,43**	**3,41**
	$\hat{y}_{t+11}$	**0,313**	**0,090**	**0,300**	**0,224**	**2,52**	**3,41**
	$\hat{y}_{t+12}$	**0,048**	**0,083**	**0,288**	**0,209**	**2,36**	**3,27**
LSTM	$\hat{y}_{t+1}$	-0,467	0,218	0,467	0,467	5,61	5,61
	$\hat{y}_{t+2}$	-0,259	0,142	0,377	0,363	4,36	4,53
	$\hat{y}_{t+3}$	-0,766	0,291	0,539	0,497	5,85	6,38
	$\hat{y}_{t+4}$	-0,822	0,387	0,622	0,578	6,71	7,28
	$\hat{y}_{t+5}$	-0,959	0,493	0,702	0,655	7,52	8,15
	$\hat{y}_{t+6}$	-0,819	0,523	0,723	0,682	7,83	8,37
	$\hat{y}_{t+7}$	-0,727	0,524	0,724	0,688	7,93	8,39
	$\hat{y}_{t+8}$	-1,139	0,620	0,788	0,745	8,55	9,10
	$\hat{y}_{t+9}$	-0,910	0,644	0,802	0,763	8,72	9,23
	$\hat{y}_{t+10}$	-1,450	0,789	0,888	0,832	9,40	10,15
	$\hat{y}_{t+11}$	-0,812	0,778	0,882	0,830	9,35	10,03
	$\hat{y}_{t+12}$	-0,856	0,774	0,880	0,832	9,37	9,99

Fonte: O autor(2019).

6.2.6 PREVISÕES PARA DADOS OBSERVADOS NO TRIMESTRE 01/NOV/2018 à 31/JAN/2019

As previsões obtidas utilizando como dados realizados no trimestre de 01/nov/2018 à 31/jan/2019, representam a velocidade média diária. Lembrando que até aqui as previsões se referiam a valores de 10min ao longo de um dia. Ao analisar o comportamento da série para o trimestre na Figura 120, pode-se observar que a série é mais comportada, principalmente devido ao operador de média amostral ($E(\mu)$) atuar similarmente a um filtro passa baixa (apenas uma abstração), eliminando componentes de alta frequência.

O intuito ao utilizar esta série com valores médio-diários, é para averiguar se

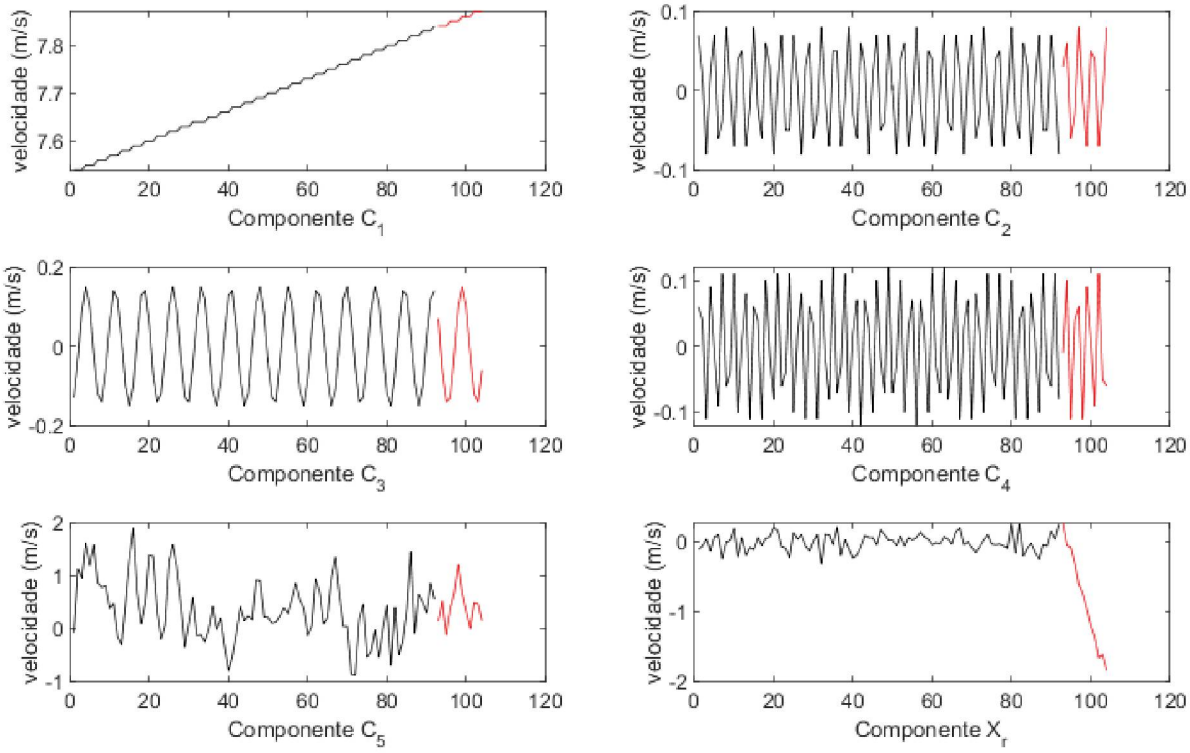
o modelo de decomposição VMD-SSA é eficiente em detectar e extrair as componentes relativa a tendência, sazonalidade e componentes harmônicas. Dado que a série tem uma inércia maior contra variações abruptas. Outra possibilidade a ser investigada com esta aplicação é se o VMD-SSA determina e consegue extrair sinais que possuem pequena amplitude oscilatória com curta permanência no tempo.

Os resultados das previsões obtidas pelo VMD-SSA-ARIMA e ARIMA estão em destaque nas Figuras 119 à 121 e são para 12 passos-à-frente, ou seja, o resultado contem valores médios de velocidades de vento para os doze dias iniciais de fevereiro/2019. Já para combinação VMD-SSA-LSTM e LSTM a amostra de dados ficou restrita ao trimestre, de forma que foram retirados os 10 últimos dias da série, resultando em uma série curta de 82 dias. O resultado das previsão de ambos os modelos para esta série estão destacados nas Figuras 122 à 124.

A análise de desempenho de cada modelo de previsão é apresentada nas Tabelas 24 e 25.

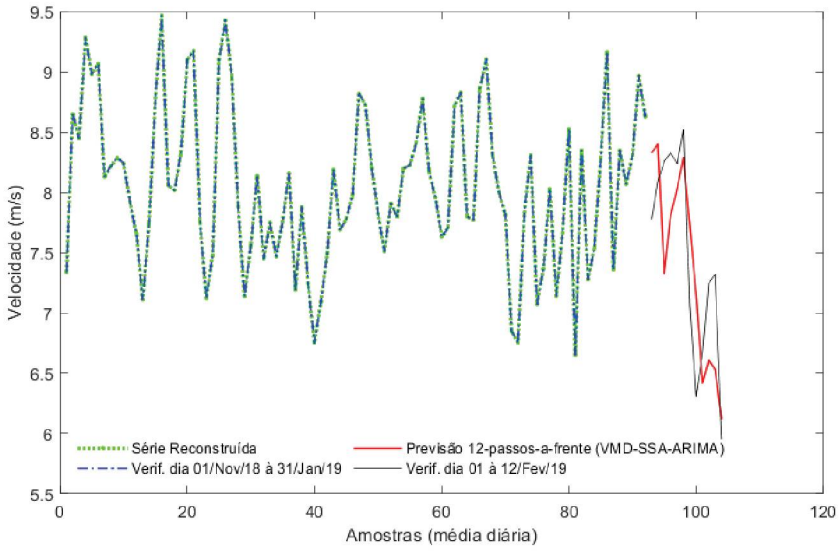
Os resultados apresentados na Figura 121 comparam os modelos ARIMA e VMD-SSA-ARIMA. Pode-se observar que o modelo ARIMA apresentou um fraco desempenho para a previsão diária de 12 passos-à-frente. Já os resultados do modelo proposto mantiveram a tendência e variação de amplitude do sinal observado.

Figura 119: Previsão 12 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento de 01/nov/2018 à 31/jan/2019 através do método de Box & Jenkins



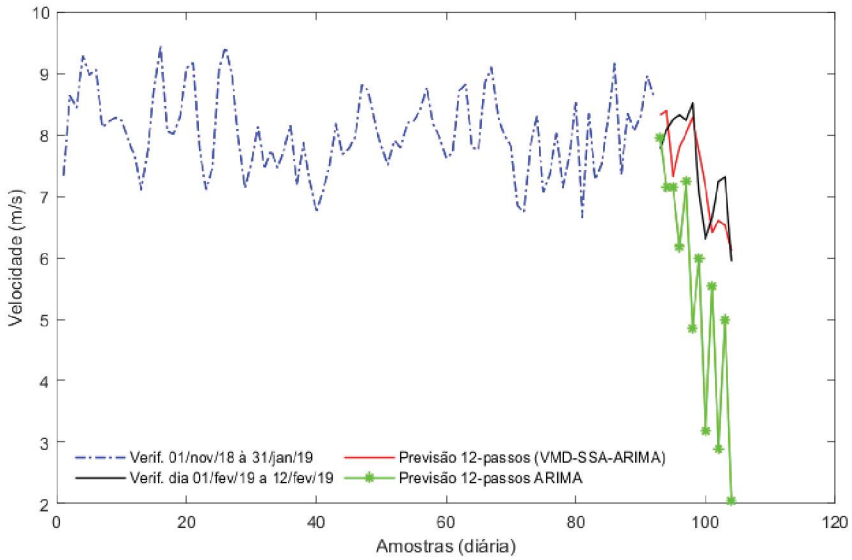
Fonte: O autor(2019).

Figura 120: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 01/nov/2018 e 31/jan/2019.



Fonte: O autor(2019).

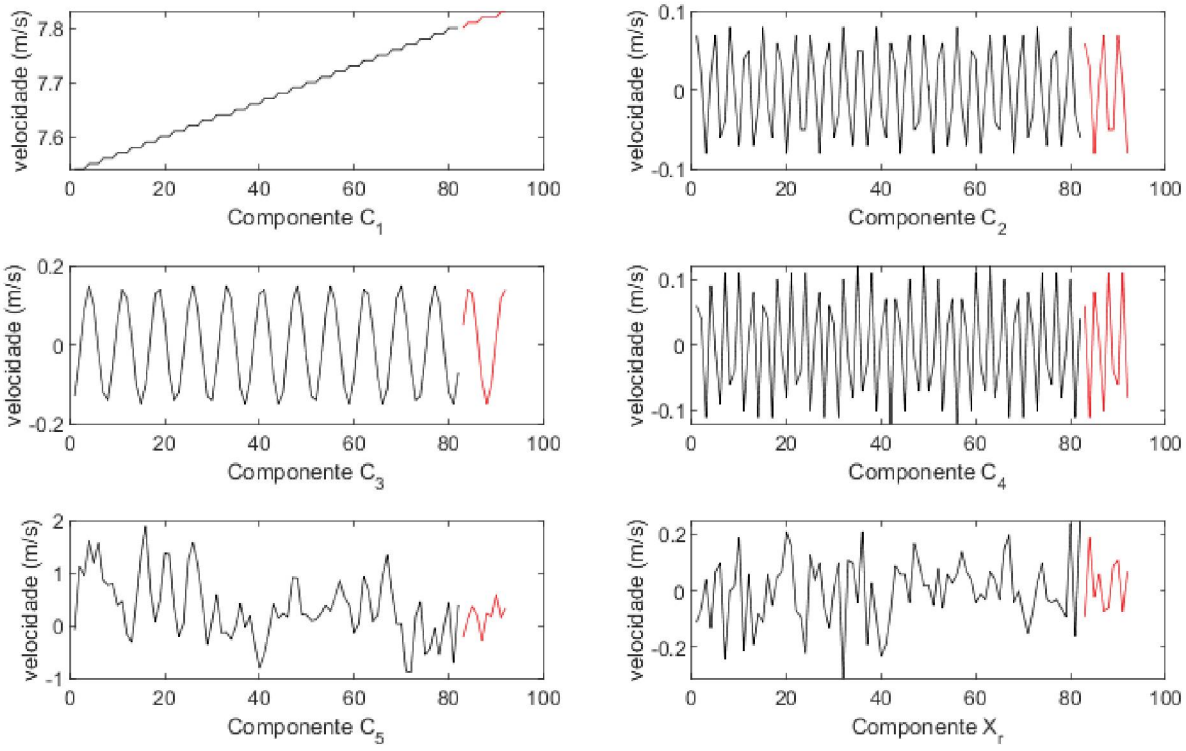
Figura 121: Comparação das previsões VMD-SSA-ARIMA *versus* ARIMA para os valores 10min observado entre 01/nov/2018 e 31/jan/2019.



Fonte: O autor(2019).

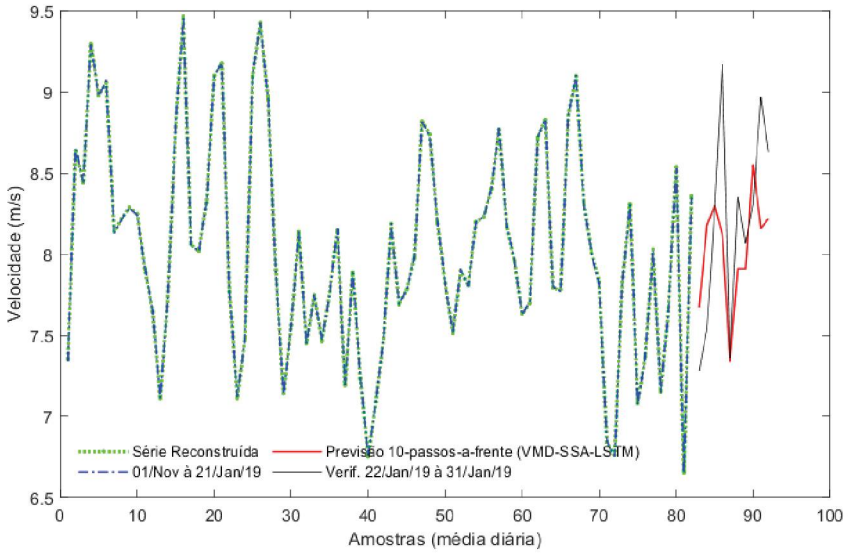
A Figura 124 apresenta a comparação de resultados entre a rede neural LSTM e o modelo VMD-SSA-LSTM. Este último apresentou resultados de previsão mais aderentes a partir do horizonte $\hat{y}_{t+2}$, mantendo o MSE estável a partir daí até o horizonte final de 10 passos-à-frente.

Figura 122: Previsão 10 passos-à-frente para as componentes C_1 à X_r (vermelho) da série velocidades de vento entre 01/nov/18 à 21/jan/2019 através da LSTM



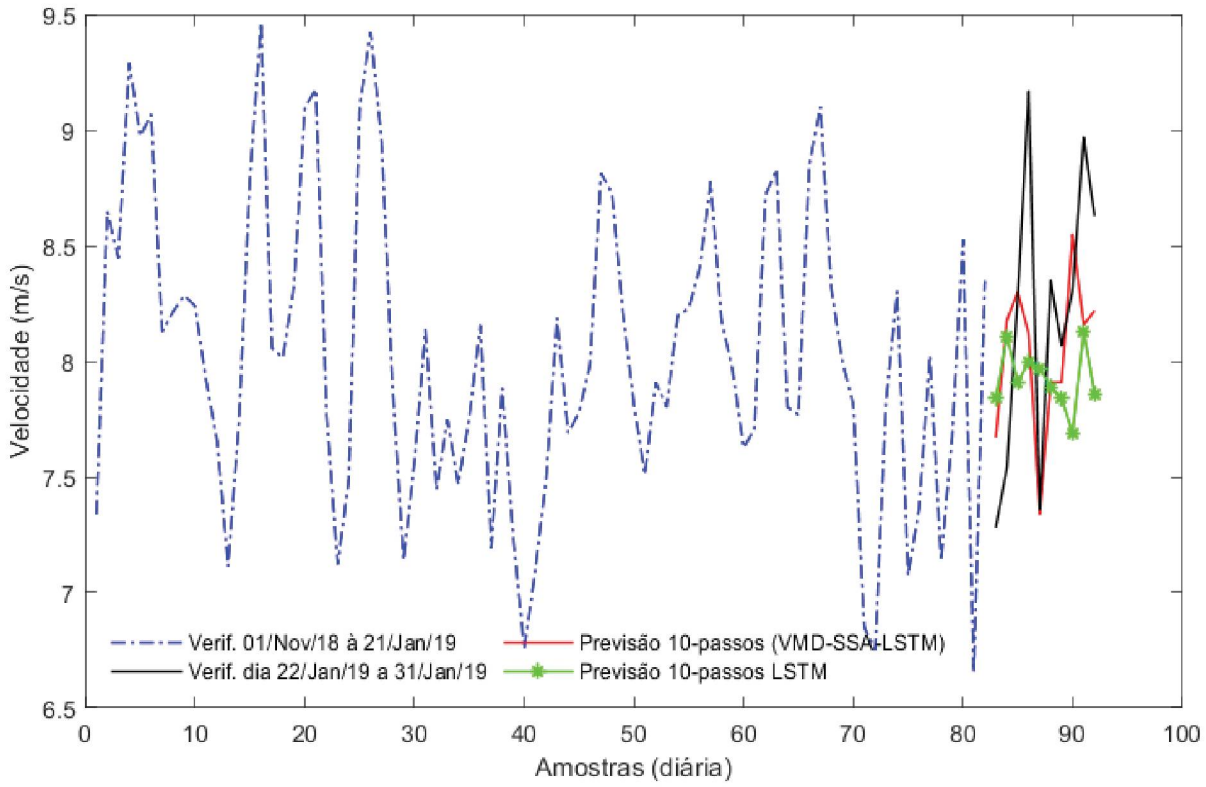
Fonte: O autor(2019).

Figura 123: Reconstrução das componentes C_1 à X_r (vermelho) originais e previstas da série velocidades *versus* observadas entre 01/nov/2018 à 21/jan/2019.



Fonte: O autor(2019).

Figura 124: Comparação das previsões VMD-SSA-LSTM *versus* LSTM para os valores diários observados no período de 01/nov/2018 à 21/jan/2019.



Fonte: O autor(2019).

Os resultados apresentados na Tabela 24 comparam o desempenho dos modelos VMD-SSA-ARIMA e ARIMA. Pode-se observar que o modelo VMD-SSA-ARIMA não apresentou bom desempenho apenas para a previsão diária de 1 passo-à-frente. Já os resultados para as previsões de $\hat{y}_{t+2}$ à $\hat{y}_{t+12}$ foram muito superiores em termos de MSE, MAPE, MAE e RRMSE.

Tabela 24: Medidas de qualidade das previsões VMD-SSA-ARIMA *versus* ARIMA no trimestre 01/nov/2018 à 31/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-ARIMA	$\hat{y}_{t+1}$	0,551	0,304	0,551	0,551	7,09	7,09
	$\hat{y}_{t+2}$	**0,310**	**0,200**	**0,447**	**0,431**	**5,46**	**5,64**
	$\hat{y}_{t+3}$	**-0,929**	**0,421**	**0,649**	**0,597**	**7,39**	**8,07**
	$\hat{y}_{t+4}$	**-0,512**	**0,381**	**0,618**	**0,576**	**7,08**	**7,61**
	$\hat{y}_{t+5}$	**-0,209**	**0,314**	**0,560**	**0,502**	**6,17**	**6,88**
	$\hat{y}_{t+6}$	**-0,237**	**0,271**	**0,521**	**0,458**	**5,61**	**6,35**
	$\hat{y}_{t+7}$	**0,669**	**0,296**	**0,544**	**0,488**	**6,16**	**6,77**
	$\hat{y}_{t+8}$	**0,835**	**0,346**	**0,588**	**0,532**	**7,04**	**7,52**
	$\hat{y}_{t+9}$	**-0,240**	**0,314**	**0,561**	**0,499**	**6,66**	**7,28**
	$\hat{y}_{t+10}$	**-0,638**	**0,324**	**0,569**	**0,513**	**6,87**	**7,43**
	$\hat{y}_{t+11}$	**-0,797**	**0,352**	**0,593**	**0,539**	**7,24**	**7,78**
	$\hat{y}_{t+12}$	**0,177**	**0,325**	**0,570**	**0,509**	**6,88**	**7,62**
ARIMA	$\hat{y}_{t+1}$	**0,182**	**0,033**	**0,182**	**0,182**	**2,3**	**2,3**
	$\hat{y}_{t+2}$	-0,942	0,460	0,678	0,562	7,0	8,5
	$\hat{y}_{t+3}$	-1,107	0,715	0,846	0,744	9,1	10,5
	$\hat{y}_{t+4}$	-2,140	1,681	1,297	1,093	13,3	16,0
	$\hat{y}_{t+5}$	-0,985	1,539	1,241	1,071	13,0	15,2
	$\hat{y}_{t+6}$	-3,659	3,515	1,875	1,503	18,0	22,9
	$\hat{y}_{t+7}$	-1,076	3,178	1,783	1,442	17,6	22,2
	$\hat{y}_{t+8}$	-3,133	4,008	2,002	1,653	21,6	25,6
	$\hat{y}_{t+9}$	-1,121	3,702	1,924	1,594	21,1	25,0
	$\hat{y}_{t+10}$	-4,355	5,228	2,287	1,870	25,0	29,9
	$\hat{y}_{t+11}$	-2,324	5,244	2,290	1,911	25,6	30,1
	$\hat{y}_{t+12}$	-3,904	6,077	2,465	2,077	28,9	33,0

Fonte: O autor(2019).

Da mesma forma que o desempenho do modelo VMD-SSA-ARIMA, o VMD-SSA-LSTM apenas não desempenhou bem para as previsão de 1 passo-a-frente. Já para as demais, considerando que neste cenário de previsão diária, foi considerado um horizonte de apenas 10 dias à frente, a técnica apresentou excelente desempenho.

Tabela 25: Medidas de qualidade das previsões VMD-SSA-LSTM *versus* LSTM no trimestre 01/nov/2018 à 31/jan/2019

Modelo	$k-step$	Erro	MSE	RMSE	MAE	MAPE (%)	RRMSE (%)
VMD-SSA-LSTM	$\hat{y}_{t+1}$	0.390	0.152	0.390	0.390	5.35	5.35
	$\hat{y}_{t+2}$	**0.650**	**0.287**	**0.536**	**0.520**	**6.99**	**7.23**
	$\hat{y}_{t+3}$	**0.017**	**0.191**	**0.438**	**0.352**	**4.73**	**5.68**
	$\hat{y}_{t+4}$	**-1.053**	**0.421**	**0.649**	**0.527**	**6.42**	**8.04**
	$\hat{y}_{t+5}$	**-0.015**	**0.337**	**0.580**	**0.425**	**5.17**	**7.32**
	$\hat{y}_{t+6}$	**-0.437**	**0.312**	**0.559**	**0.427**	**5.18**	**6.99**
	$\hat{y}_{t+7}$	**-0.161**	**0.271**	**0.521**	**0.389**	**4.73**	**6.51**
	$\hat{y}_{t+8}$	**0.247**	**0.245**	**0.495**	**0.371**	**4.51**	**6.15**
	$\hat{y}_{t+9}$	**-0.810**	**0.291**	**0.539**	**0.420**	**5.01**	**6.62**
	$\hat{y}_{t+10}$	**-0.412**	**0.279**	**0.528**	**0.419**	**4.99**	**6.44**
LSTM	$\hat{y}_{t+1}$	**0.270**	**0.073**	**0.270**	**0.270**	**3.70**	**3.70**
	$\hat{y}_{t+2}$	0.722	0.297	0.545	0.496	6.65	7.36
	$\hat{y}_{t+3}$	0.402	0.252	0.502	0.465	6.05	6.52
	$\hat{y}_{t+4}$	-0.379	0.225	0.474	0.443	5.57	5.88
	$\hat{y}_{t+5}$	0.875	0.333	0.577	0.530	6.83	7.28
	$\hat{y}_{t+6}$	-0.312	0.294	0.542	0.493	6.32	6.78
	$\hat{y}_{t+7}$	-0.585	0.301	0.548	0.506	6.45	6.85
	$\hat{y}_{t+8}$	-0.675	0.320	0.566	0.528	6.66	7.03
	$\hat{y}_{t+9}$	-1.226	0.452	0.672	0.605	7.44	8.25
	$\hat{y}_{t+10}$	-0.975	0.501	0.708	0.642	7.83	8.64

Fonte: O autor(2019).

Os resultados são analisados no próximo capítulo, porém pode-se observar para todos os cenários de previsões, obtidas pelo acoplamento do modelo de Box & Jenkins ou da rede neural LSTM ao modelo de decomposição, provocou uma estabilidade nos erros do modelo de previsão.

6.3 RESUMO DO CAPÍTULO

O capítulo apresentou os resultados da previsão de séries de velocidades de vento para um horizonte máximo de 12 passos-à-frente. Os modelos de previsão utilizado foram o paramétrico de Box & Jenkins e um modelo não paramétrico, constituído por uma rede neural LSTM. As previsões obtidas para cada componente extraída pelo modelo híbrido foi comparada ao modelo canônico (ou puro) para avaliar o impacto da etapa de decomposição na qualidade das previsões.

Pode-se concluir que o modelo proposto estabiliza o erro, de forma que as previsões apresentam bons resultados com MSE, MAE, MAPE e RRMSE bem inferior aos obtidos sem a etapa de decomposição.

7 CONCLUSÃO E PESQUISA FUTURA

Neste capítulo é feito uma análise dos resultados obtidos através da aplicação do modelo proposto, destacando os pontos fortes, as dificuldades e possíveis melhorias. A seção inicia com à análise do resultado das decomposições, a correlação entre os dados e seu efeitos. Em seguida discute-se a resposta dos modelos de previsão, com uma análise da qualidade do resultado obtido por cada variante do modelo. E por fim, é estabelecido um rol de possíveis tópicos para pesquisa futura, visando tornar mais robusto o método proposto.

7.1 ANÁLISE DAS DECOMPOSIÇÕES

Os resultados apresentados na Seção 5.1 indicam o potencial do modelo proposto na aplicação de decomposição de sinais. Os resultados para as séries temporais de velocidades de vento, amostradas a cada 10min, possuem variância elevada, que é a característica de sinais não estacionários. Ao aplicar a técnica para extrair as 5 componentes selecionadas, o modelo foi capaz de obter componentes com elevado grau de separação na maioria dos casos. Essa eficácia é comprovada pela resultado das correlações apresentadas para cada componente obtida após o processo de decomposição.

Analisando alguns dos resultados onde a correlação foi a superior a 15%, pode-se perceber que as componentes resultantes apresentam em sua estrutura componentes oscilatórias de alta frequência em alguns casos, e em outros de baixíssima frequência. Como exemplo a estrutura da componente C_5 da série do dia 02/janeiro/19 comparada a da componente C_2 da Figura 41 e também à Figura 42. A solução para esse problema passa por aumentar o valor de α, de forma que o filtro seja mais seletivo, porém a dificuldade em ajustar um valor α que possibilite que todas as componentes estejam centradas efetivamente em frequências distintas, ainda que harmônicas, é um problema em aberto ainda.

Cabe ressaltar que grande parte das correlações acima desse percentual citado,

normalmente envolve uma das duas últimas componentes C_5 ou X_r. Uma grande vantagem porém do modelo, é que o nível de informação da série original contida nessas duas componentes é desprezível na maioria das vezes, surgindo então a possibilidade de se descartar uma delas, naturalmente a componente X_r parece ser uma escolha mais razoável. O descarte de X_r não utilizado nesta tese, dado que a hipótese que se deseja investigar era a acurácia introduzida na previsão através do modelo híbrido de decomposição, sem descartar qualquer componente. Outra opção seria proporcionar o aumento do número de modos a serem extraídos do sinal original, porém há um *trade-off* que deve ser considerado, para o caso específico de previsão de séries de temporais, dado que com aumento do número de componentes têm-se um crescimento expressivo do número de parâmetros do modelo previsor, como já discutido ao longo da tese. É importante notar também que, as séries com menor variabilidade de amplitude são as que apresentaram maiores correlações com a componente C_1, que normalmente mapeia a tendência da série, e com a componente X_r que é a componente residual após a decomposição. Apesar de em alguns casos, a matriz de correlação apresentar valores superior a 40%, veremos na análise dos resultados das previsões que o impacto dessa correlação um pouco elevada foi mínimo. Pode-se concluir desta forma, que o modelo proposto tem potencial para ser aplicado na etapa de decomposição e filtragem de séries de velocidades de vento para previsões de curto prazo. Desta afirmação resulta um outro ponto que pode ser explorado em pesquisa futura, ou seja, a análise do valor máximo para correlações entre componentes que não provoque instabilidade nos erros resultantes do modelo de previsão.

Diante dos resultados obtidos, pode-se afirmar que os objetivos relacionados a etapa de decomposição, proposto no início desta tese, foram atingidos em sua plenitude.

7.2 ANÁLISE DAS PREVISÕES

A conclusão elaborada na seção anterior é fundamentada em grande parte na qualidade das previsões e na análise das tabelas dos indicadores de qualidade das previsões. Independentemente da técnica de previsão utilizada, uma característica importante incorporada pelo modelo proposto é a estabilização dos erros de previsão. Percebe-se que a medida que o horizonte de previsão aumenta, os índices MSE, MAE, MAPE e RRMSSE se mantém comportados, oscilando lentamente, enquanto que esses índices aplicado às respostas do modelo de previsão sem a etapa de decomposição e filtragem, resultam em um comportamento crescente com o horizonte de previsões, além de apresentarem oscilações bruscas entre um horizonte qualquer e o subsequente. Desta análise conclui-se

que o modelo proposto vai de encontro a hipótese levantada no início desta pesquisa, que considera a etapa de decomposição uma possibilidade na redução dos erros de previsão. Acrescente-se aqui, que o fato de provocar a estabilização do erro, advém do modelo de previsão estar mapeando apenas o comportamento individual local contido em cada componente, tal como a sazonalidade, tendência, as relações harmônicas, etc., resultado da soma das contribuições de vários mecanismos independentes formador da série, e não um comportamento amplo contido na série original.

Outro ponto relevante é a melhoria da capacidade dos modelos de previsão que a partir da série decomposta, consegue mapear e prever variações bruscas no comportamento futuro da série. Este fato é evidente quando é analisado os resultados dos índices de desempenho para os dados previstos dos dias 03 e 05 de janeiro. Para estes dois dias, as velocidades verificadas possuem uma característica muito impactante para o Setor Elétrico, mais especificamente em locais onde há forte presença de produção de energia eólica, como já pontuado em seção anterior deste trabalho, que é a rampa acentuada de subida ou descida em certos intervalos. Desta forma, possuir modelos que tenham a capacidade de estimar a tendência dessas rampas, torna a operação do Sistema Elétrico mais segura, dado que as previsões dessas rampas são utilizadas em análise de segurança e de estabilidade de sistemas elétricos, bem como nos ajustes dos reguladores de tensão e de velocidades das máquinas das hidrelétricas despachadas.

Avaliando ainda as respostas do modelo ARIMA, quanto as obtidas pela rede LSTM para o caso específico sem a etapa de decomposição, os mesmos apresentaram índices MSE, MAE, MAPE e RRMSE relativamente ruins, demonstrando que o mecanismo formador da série é muito complexo e não pode ser explicado pelo modelo previsor, dado que o padrão previsto por ambos está totalmente descorrelacionado com o verificado. Em contrapartida, tem-se a estabilização ou pouca variação do erro para os resultados obtidos pelo modelo híbrido VMD-SSA-ARIMA e VMD-SSA-LSTM para estes dois casos.

Outra análise importante é a qualidade das previsões para séries curtas como a utilizada no trimestre, relativo aos dados diários de 01/nov/2018 à 31/jan/2019. Esta série contém apenas 92 amostras no cenário para previsão 12 passos-a-frente, e 82 amostras no cenário de 10 passos-a-frente. Os resultados obtidos pelo modelo ARIMA apresentam índices de qualidade relativamente ruins, para o caso em que a etapa de decomposição não é considerada, enquanto que o modelo proposto consegue explicar o mecanismo formador da série relativamente bem, com índices MSE, MAE, MAPE e RRMSE bem estáveis e aceitáveis. No cenário de 10 passos-a-frente, a rede LSTM devido a sua característica de

retropropagação, conseguiu resultados melhores se comparado ao modelo ARIMA. Porém ao utilizar a rede LSTM acoplado ao modelo VMD-SSA, a qualidade das previsões foi superior, conforme os valores dos índices MSE, MAE, MAPE, RMSE e RRMSE para cada horizonte de previsão.

Desta forma, resta concluir que a hipótese desta tese, bem como todos os objetivos propostos foram atingidos em sua plenitude, resultando desta pesquisa, um modelo de decomposição híbrido multiestágio capaz de estabilizar os erros de previsão de séries temporais de velocidades de vento.

7.3 PESQUISA FUTURA

Apesar dos resultados obtidos, não foi possível a generalização da aplicação do método a outros tipos de séries temporais. Desta forma, os pontos em aberto, que podem ser tratados em pesquisas futura são elencados nessa seção. Os temas apontados aqui, são correlatos ao fato desta pesquisa ter sido aplicado apenas a séries curtas de velocidades de vento. Desta forma, os problemas identificados e que necessitam de aprofundamento para tornar a técnica robusta, quando aplicada a outros tipos de séries temporais podem ser enumerados como sendo a necessidade de:

1. Determinar um método explícito para a escolha do número de componentes a serem extraídas do sinal original, e avaliar se o número específico de componentes guarda alguma relação e influência na determinação do valor de α;

2. Avaliar outros métodos para a definição do número de componentes L a ser extraída no método SSA, de forma a tornar o modelo mais robusto, pois ao extrair um número elevado de componentes, pode ocorrer a saturação de algumas componentes, como citado em Golyandina et al. (2001);

3. Conduzir análise relativa ao valor máximo para correlações entre componentes, de forma que não resulte em instabilidade no erro resultante do modelo de previsão para cada horizonte de previsão;

4. Aplicar diferentes técnicas paramétricas ou de Inteligência Computacional (CI) na etapa de previsão: Ao generalizar a aplicação do modelo híbrido, utilizando técnicas de previsão como ANFIS, GPR, GMDH, entre outras, busca-se confirmar que o modelo proposto realmente potencializa a capacidade de estabilizar os erros de previsão para qualquer técnica e não só daquelas aqui aplicada e;

5. Explorar com maior profundidade a contribuição de cada componente na explicação do fenômeno que se está buscando prever, aplicando o modelo em diferentes séries temporais.

REFERÊNCIAS

ABEEóLICA. *Boletim Anual De Geração Eólica 2016.* [S.l.]: Abeeólica, 2017. 31 p.

ABEEóLICA. *Boletim Anual De Geração Eólica 2018.* [S.l.]: Abeeólica, 2018. 15 p.

ARTECHE, J.; GARCíA-ENRíQUEZ, J. Singular spectrum analysis for signal extraction in stochastic volatility models. *Econometrics and Statistics*, v. 1, p. 85 – 98, 2017.

AYALA, J. C. C. *Estimação da estrutura a prazo da curva de rendimentos para Colômbia: aplicação empírica com análise de espectro singular.* 44 f. Mestrado em Economia — Faculdade de Ciências Econômicas, Porto Alegre, RS, 2016.

BARUNíK, J.; KřEHLíK, T. Combining high frequency data with non-linear models for forecasting energy market volatility. *Expert Systems with Applications*, v. 55, p. 222 – 242, 2016.

BENDRE, M.; MANTHALKAR, R. ime series decomposition and predictive analytics using mapreduce framework. *Expert Systems with Applications*, v. 116, p. 108–120, 2019.

BONFIL, G. S.; BALLESTEROS, A. R.; GERSHENSON, C. Wind speed forecasting for wind farms: A method based on support vector regression. *Renewable Energy*, v. 85, p. 790 – 809, 2016.

BOX, G.; JENKINS, G.; REINSEL, G. *Time Series Analysis, Forecasting and Control.* Third edition. [S.l.]: Holden-Day, Inc, 1976.

BROOMHEAD, D. S.; KING, G. P. Extracting qualitative dynamics from experimental data. *Physica D*, v. 20, p. 3217–236, 1986.

CASSIANO, K. M. *Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e Clusterização de Suas Componentes Baseada em Densidade.* 172 f. Doutorado em Engenharia Elétrica — Faculdade de Engenharia Eletrica, Rio de Janeiro, RJ, 2014.

CAVALCANTE, R. C.; BRASILEIRO, R. C.; SOUZA, V. L.; NOBREGA, J. P.; OLIVEIRA, A. L. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, v. 55, p. 194 – 211, 2016.

CHATEFIELD, C. *Time-series forecasting.* [S.l.]: Chapman Hall Crc, 2000.

CHEN, S.; DONG, X.; PENG, Z.; ZHANG, W.; MENG, G. Nonlinear chirp mode decomposition: A variational method. *IEEE Transactions on Signal Processing*, v. 65, n. 22, p. 6024–6037, Nov 2017.

CHENG, H.; TAN, P.-N.; GAO, J.; SCRIPPS, J. Multistep-ahead time series prediction. In: *Advances in Knowledge Discovery and Data Mining.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 765–774.

COHEN, M. X. *Analyzing Neural Time Series Data: Theory and Practice*. [S.l.]: MIT Press, 2014.

COOLEY, J. M.; TUKEY, J. W. An algorithm for the machine calculation of complex fourier transforms. *Mathematics Computation*, n. 12, p. 297–301, 1965.

DANTAS, P. V. *Identificação por decomposição de sinais de consumo de energia elétrica*. 78 f. Mestrado em Engenharia Elétrica — Faculdade de Engenharia Eletrica, Manaus, AM, 2016.

DESPOTOVIC, M.; NEDIC, V.; DESPOTOVIC, D.; CVETANOVIC, S. Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. *Renewable and Sustainable Energy Reviews*, v. 56, p. 246–260, 2016.

DRAGOMIRETSKIY, K. V.; ZOSSO, D. Variational mode decomposition. *IEEE Transactions on Signal Processing*, v. 62, n. 3, p. 531–544, Feb 2014.

ELSNER, J. B.; TSONIS, A. A. *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. [S.l.]: Plenum Press, 1996.

ESQUIVEL, R. de M. *Análise Singular Espectral: Modelagem de séries temporais através de estudos comparativos usando diferentes estratégias de previsão*. 174 f. Mestrado em Modelagem Computacional e Tecnologia Industrial — SENAI CIMATEC, Salvador, BA, 2012.

GOLYANDINA, N.; NEKRUTKIN, V.; ZHIGLJAVSKY, A. *Analysis of Time Series Structure*: Ssa and related techniques. [S.l.]: Chapman and Hall CRC, 2001.

GONG, X.; SI, Y.-W.; FONG, S.; BIUK-AGHAI, R. P. Financial time series pattern matching with extended {UCR} suite and support vector machine. *Expert Systems with Applications*, v. 55, p. 284 – 296, 2016.

HIRANSHA, M.; GOPALAKRISHNAN, E.; MENON, V. K.; SOMAN, K. Nse stock market prediction using deep-learning models. *Procedia Computer Science*, v. 132, p. 1351 – 1362, 2018. International Conference on Computational Intelligence and Data Science.

HOCHREITER, S.; SCHMIDHUBER, J. Long short term memory. *Neural Computation*, v. 9, n. 8, p. 1735–1780, 1997.

HUANG, N.; SHEN, Z.; LONG, S.; WU, M.; SHIH, H.; ZHANG, Q.; YEN, N.; TUNG, C.; LIU, H. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Royal Society of London Proceedings Series A*, n. 454, p. 903–995, 1998.

JENKINS, G. M.; WATTS, D. G. *Spectral Analysis and its applications*. [S.l.]: Holden-Day, 1968.

JIANG, Y.; HUANG, G. Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction. *Energy Conversion and Management*, v. 144, p. 340 – 350, 2017.

JOLLIFFE, I. T. *Principal Component Analysis*: Series: Springer series in statistics. 2nd. ed. [S.l.]: Springer, 2002.

JUNIOR, L. A. T. *Combinação SSA-WAVELET de Métodos Preditivos com ajuste numérico MINIMAX, na geração de previsões e de cenários*. 114 f. Doutorado em Engenharia Elétrica — Faculdade de Engenharia de Elétrica, Pontificia Universidade Católica do Rio de Janeiro, PUC Rio, Rio de Janeiro, RJ, 2013. Disponível em: ⟨https://www.maxwell.vrac.puc-rio.br/Busca_etds.php?strSecao=resultado&nrSeq=22892@1⟩.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: *Conference Track Proceedings*. San Diego, CA, USA: [s.n.], 2015.

KLINE, D. M. Methods for multi-step time series forecasting neural networks. In: . [S.l.]: IGI Global, 2004. p. 226–250.

KUSIAK, A.; ZHENG, H.; SONG, Z. On-line monitoring of power curves. *Renewable Energy*, v. 34, n. 6, p. 1487 – 1493, 2009.

LEI, M.; SHIYAN, L.; CHUANWEN, J.; HONGLING, L.; YAN, Z. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, v. 13, n. 4, p. 915 – 920, 2009.

LI, Y.; ZHU, Z.; KONG, D.; HAN, H.; ZHAO, Y. Ea-lstm: Evolutionary attention-based lstm for time series prediction. *Knowledge-Based Systems*, 2019.

LIU, H.; TIAN, H. qi; CHEN, C.; LI, Y. fei. An experimental investigation of two wavelet-mlp hybrid frameworks for wind speed prediction using {GA} and {PSO} optimization. *International Journal of Electrical Power e Energy Systems*, v. 52, p. 161 – 173, 2013. ISSN 0142-0615.

LIU, H.; TIAN, H. qi; LI, Y. fei. Comparison of two new arima-ann and arima-kalman hybrid methods for wind speed prediction. *Applied Energy*, v. 98, p. 415 – 424, 2012.

LOPES, Z. F. *Detecção de mudança hidroclimática na bacia hidrográfica do rio São Francisco*. 165 f. Doutorado em Geografia — Faculdade de Geografia, 2014. Disponível em: ⟨http://repositorio.ufpe.br/handle/123456789/11134⟩.

LUCENA, D. S. de; MORENO, S.; MARIANI, V. C.; COELHO, L. Support vector machine optimized by artificial bee colony applied to eeg pattern recognition. In: *Proceedings in Lecture Notes in Artificial Intelligence*. [S.l.]: Springer, Heidelberg, 2017. v. 1, p. 1–2.

MA, X.; JIN, Y.; DONG, Q. A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting. *Applied Soft Computing*, v. 54, p. 296 – 312, 2017.

MALLAT, S. G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Pattern Analysis and Machine Intelligence*, v. 11, n. 7, p. 674–693, 1989.

MATSUOKA, D. H. *Previsões da inflação brasileira a partir do método não-paramétrico de Análise Espectral Singular*. 34 f. Mestrado em Engenharia Elétrica — Faculdade de Ciências Econômicas, Porto Alegre, RS, 2016.

MEHRKANOON, S. Deep shared representation learning for weather elements forecasting. *Knowledge-Based Systems*, 2019.

MEN, Z.; YEE, E.; LIEN, F.-S.; WEN, D.; CHEN, Y. Short-term wind speed and power forecasting using an ensemble of mixture density neural networks. *Renewable Energy*, v. 87, Part 1, p. 203 – 211, 2016.

MENEZES, M. L. de. *Abordagem PAR(p) e Singular Spectrum Analysis na Modelagem e Geração de Cenários*. 122 f. Doutorado em Engenharia Elétrica — Faculdade de Engenharia Eletrica, Rio de Janeiro, RJ, 2014.

MI, X.; LIU, H.; LI, Y. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Conversion and Management*, v. 180, p. 196–205, 2019.

MOORE, K. J.; KURT, M.; ERITEN, M.; MCFARLAND, D. M.; BERGMAN, L. A.; VAKAKIS, A. F. Wavelet-bounded empirical mode decomposition for measured time series analysis. *Mechanical Systems and Signal Processing*, v. 99, p. 14–29, 2018.

MORENO, S.; KLEIN, C.; MARIANI, V. C.; COELHO, L. D. S.; ALOTTO, P.; BAUERNFEIND, T.; BAUMGARTNER, P. Non-uniformly spaced linear antenna array design by means of peec approach applying cheetah optimization algorithm. In: *15th International Workshop on Optimization and Inverse Problems in Electromagnetism (OIPE 2018)*. [S.l.: s.n.], 2018. v. 1, p. 38–39.

MORENO, S.; KLEIN, C. E.; MARIANI, V. C.; COELHO, L.; ALOTTO, P.; BAUERNFEIND, T.; BAUMGARTNER, P. Non-uniformly spaced linear antenna array design by means of peec approach applying cheetah optimization algorithm. *International Journal of Applied Electromagnetics and Mechanics*, v. 60, p. 15–24, 2019.

MORENO, S. R.; COELHO, L. dos S. Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renewable Energy*, v. 126, p. 736–754, 2018.

MORENO, S. R.; JUNIOR, O. da S.; COELHO, L. dos S. Adaptive imperialist competitive algorithm applied to wind farm layout optimization. In: *L Simpósio Brasileiro de Pesquisa Operacional, L SBPO*. [S.l.]: Sociedade Brasileira Pesquisa Operacional, SOBRAPO, Rio de Janeiro, RJ, 2018. v. 50.

MORETTIN, P. A.; TOLOI, C. M. *Análise de séries temporais*. 2 edição. ed. [S.l.]: Editora Edgard Blucher Ltda, 2006.

NETO, J. X. V.; JUNIOR, E. J. G.; MORENO, S. R.; AYALA, H. V. H.; MARIANI, V. C.; COELHO, L. dos S. Wind turbine blade geometry design based on multi-objective optimization using metaheuristics. *Energy*, v. 162, p. 645 – 658, 2018.

NYQUIST, H. Certain topics in telegraph transmission theory. *Proceedings of the IEEE*, v. 90, n. 2, p. 280–305, Feb 2002.

ONS. *Boletim Mensal de Geração Eólica Março 2017*. Operador Nacional do Sistema, 2017. 65 p. Disponível em: ⟨http://www.ons.org.br/download/resultados_operacao/boletim_mensal_geracao_eolica/Boletim_Eolica_mar_2017.pdf⟩.

PAPAILIAS, F.; THOMAKOS, D. Exssa: Ssa-based reconstruction of time series via exponential smoothing of covariance eigenvalues. *International Journal of Forecasting*, v. 33, n. 1, p. 214 – 229, 2017.

QIU, X.; REN, Y.; SUGANTHAN, P. N.; AMARATUNGA, G. A. Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Applied Soft Computing*, v. 54, p. 246 – 255, 2017.

RIOS, R. A.; MELLO, R. F. de. A systematic literature review on decomposition approaches to. estimate time series components. *Journal of Computer Science*, v. 11(3-4), p. 31–46, 2012.

SALAZAR, A. E. C. *Mineração de estruturas musicais e composição automática utilizando redes complexas*. 171 f. Doutorado em Ciências Matemáticas e de Computação — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, USP SP, São Carlos, SP, 2015.

SANEI, S.; HASSANI, H. *Singular Spectrum Analysis of Biomedical Signals*. [S.l.]: Taylor Francis Inc, United States, 2015.

SANTOS, F. P. *Identicação de sistemas neurais com redes bayesianas dinâmicas e transferência de entropia*. 78 f. Doutorado em Engenharia Elétrica — Faculdade de Engenharia Eletrica, São Carlos, SP, 2017.

SANTOS, J. E. S. dos. *Controle Preditivo Não-Linear para Sistemas de Hammerstein*. 156 f. Doutorado em Engenharia Elétrica — Faculdade de Engenharia Elétrica, Florianópolis, SC, 2007.

SORJAMAA, A. C.; LENDASSE, A. Time series prediction using dirrec strategy. In: . Bruges, Belgium: M. Verleysen (Ed.), 2006. p. 143–148.

TAIEB, S. B.; BONTEMPI, G.; ATIYA, A. F.; SORJAMAA, A. A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition. *Expert Systems with Applications*, Elsevier, v. 39, n. 8, p. 7067–7083, 6 2012.

TAIEB, S. B.; BONTEMPI, G.; SORJAMAA, A.; LENDASSE, A. Long-term prediction of time series by combining direct and mimo strategies. In: . [S.l.: s.n.], 2009. p. 3054–3061.

TEIXEIRA, A. R.; TOME, A. M.; LANG, E. W.; GRUBER, P.; SILVA, A. M. da. On the use of clustering and local singular spectrum analysis to remove ocular artifacts from electroencephalograms. In: *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* [S.l.: s.n.], 2005. v. 4, p. 2514–2519.

TESSIER, T. H.; ARMSTRONG, J. S. Decomposition of time-series by level and change. *Journal of Business Research*, v. 68, n. 8, p. 1755 – 1758, 2015. Special Issue on Simple Versus Complex Forecasting.

TIAO, G.; TSAY, R. Some advances in non-linear and adaptive modelling in time-series. *Journal of Forecasting*, v. 13, n. 2, p. 109–131, 1994.

TITCHMARSH, E. C. *Introduction to the Theory of Fourier Integrals, 2nd edition.* [S.l.]: Oxford University Press, 1948.

TSILI, M.; AMOIRALIS, E. I.; LEITE, J. V.; MORENO, S. R.; COELHO, L. dos S. A novel multiobjective lognormal-beta differential evolution approach for the transformer design optimization. *Engineering Computations*, v. 35, n. 2, p. 955–978, 2018.

VARGAS, S. A. *Previsão da Distribuição da Densidade de Probabilidade da Geração de Energia Eólica usando Técnicas não Paramétricas.* 191 f. Doutorado em Engenharia Elétrica — Faculdade de Engenharia Eletrica, Rio de Janeiro, RJ, 2014.

WANG, J.; HU, J. A robust combination approach for short-term wind speed forecasting and analysis – combination of the arima (autoregressive integrated moving average), elm (extreme learning machine), svm (support vector machine) and lssvm (least square svm) forecasts using a gpr (gaussian process regression) model. *Energy*, v. 93, p. 41 – 56, 2015.

WANG, J.; WANG, J. Forecasting stochastic neural network based on financial empirical mode decomposition. *Neural Networks*, v. 90, p. 8 – 20, 2017.

WANG, S.; CHEN, X.; CAI, G.; CHEN, B.; LI, X.; HE, Z. Matching demodulation transform and analysis. *IEEE Transactions on Signal Processing*, v. 62, n. 1, p. 69–84, 2014.

WANG, Y.; WANG, J.; WEI, X. A hybrid wind speed forecasting model based on phase space reconstruction theory and markov model: A case study of wind farms in northwest china. *Energy*, v. 91, p. 556 – 572, 2015.

XIAO, J.; LI, Y.; XIE, L.; LIU, D.; HUANG, J. A hybrid model based on selective ensemble for energy consumption forecasting in china. *Energy*, v. 159, p. 534–546, 2018.

XIAO, L.; QIAN, F.; SHAO, W. Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm. *Energy Conversion and Management*, v. 143, p. 410–430, 2017.

YASEEN, Z. M.; SULAIMAN, S. O.; DEO, R. C.; CHAU, K.-W. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, v. 569, p. 387–408, 2019.

YU, S.; MA, J. Complex variational mode decomposition for slop-preserving denoising. *IEEE Transactions on Geoscience and Remote Sensing*, v. 56, n. 1, p. 586–597, Jan 2018.

ZHANG, X.; WANG, J.; ZHANG, K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm. *Electric Power Systems Research*, v. 146, p. 270 – 285, 2017.

ZIEL, F.; CROONENBROECK, C.; AMBACH, D. Forecasting wind power – modeling periodic and non-linear effects under conditional heteroscedasticity. *Applied Energy*, v. 177, p. 285 – 297, 2016.

APÊNDICE A – ESTRATÉGIAS DE PREVISÃO

A etapa de previsão pode ser realizada através da aplicação de técnicas paramétricas, tais como os modelos de Box & Jenkins, ou de modelos não paramétricos tais como relação linear de recorrência (LRR), redes neurais artificiais, sistemas nebuloso adaptativo de inferência neuro (ANFIS), GMDH, SVM, GPR entre outras técnicas. Como o foco desta pesquisa está relacionada ao método de decomposição da série temporal, as técnicas aplicadas à previsão não serão tratadas de forma detalhada. As mesmas estão descritas no Apêndice deste trabalho, buscando dessa forma facilitar o entendimento conceitual de cada técnica utilizada.

Ao contrário das técnicas de previsão, as estratégias de previsão para diferentes horizontes ou k-passos a frente ($k = 1, \ldots, K$), também conhecida como previsão de longo termo (*long term forecasting*) são apresentadas com maiores detalhes nesta seção, principalmente devido a escassez de materiais que apresentem essa modelagem, tal como apontado em (TAIEB et al., 2012).

A.1 ESTRATÉGIA RECURSIVA

Talvez esta seja a mais intuitiva estratégia de previsão encontrada na literatura. A mesma consiste em ajustar um único modelo f de previsão de 1 passo-a-frente (TIAO; TSAY, 1994) e aplicá-lo para obter as previsões dos K passos-a-frente, conforme Equação A.1, onde ϵ representa o erro do modelo.

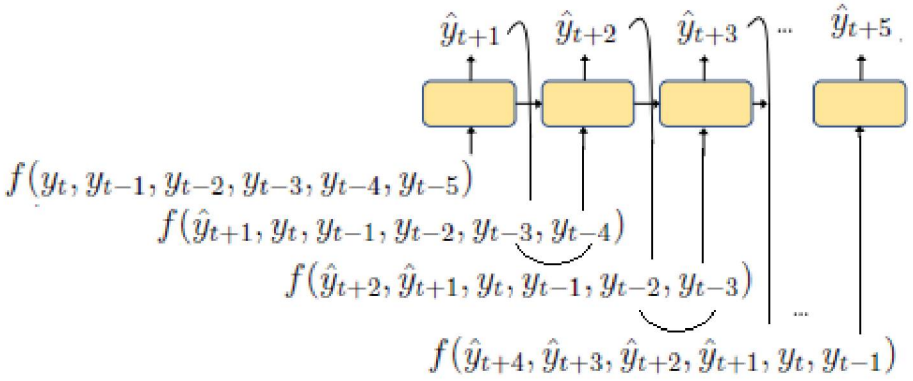
$$y_{t+1} = f(y_t, \ldots, y_{t-d+1}) + \epsilon, \quad t \in \{d, \ldots, N-1\} \tag{A.1}$$

Uma desvantagem, apesar da simplicidade do modelo, é a propagação do erro ao longo do horizonte de previsões, dado que a previsão da etapa atual é utilizada como entrada para a previsão da etapa seguinte. A qualidade deteriora severamente quando o horizonte de previsão K excede a dimensão da janela deslizante (*embedding dimension*)

d. A estratégia recursiva para 5 passos-a-frente é apresentada na Equação A.2 e na Figura A.1. A Equação A.3 apresenta a generalização para K passos-a-frente de uma série de tamanho N,

$$\begin{aligned}
\hat{y}_{t+1} &= f(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}) \\
\hat{y}_{t+2} &= f(\hat{y}_{t+1}, y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}) \\
\hat{y}_{t+3} &= f(\hat{y}_{t+2}, \hat{y}_{t+1}, y_t, y_{t-1}, y_{t-2}, y_{t-3}) \\
\hat{y}_{t+4} &= f(\hat{y}_{t+3}, \hat{y}_{t+2}, \hat{y}_{t+1}, y_t, y_{t-1}, y_{t-2}) \\
\hat{y}_{t+5} &= f(\hat{y}_{t+4}, \hat{y}_{t+3}, \hat{y}_{t+2}, \hat{y}_{t+1}, y_t, y_{t-1})
\end{aligned} \tag{A.2}$$

Figura A.1: Estratégia recursiva para previsão 5-passos-a-frente



Fonte: O autor(2019).

$$\hat{y}_{N+k} = \begin{cases} f(y_N, \ldots, y_{N-d+1}), & \text{se } k = 1 \\ f(\hat{y}_{N+k-1}, \ldots, \hat{y}_{N+1}, y_N, \ldots, y_{N-d+k}), & \text{se } k \in \{2, \ldots, d\} \\ f(\hat{y}_{N+k-1}, \ldots, y_{N+k-d}), & \text{se } k \in \{d+1, \ldots, K\} \end{cases} \tag{A.3}$$

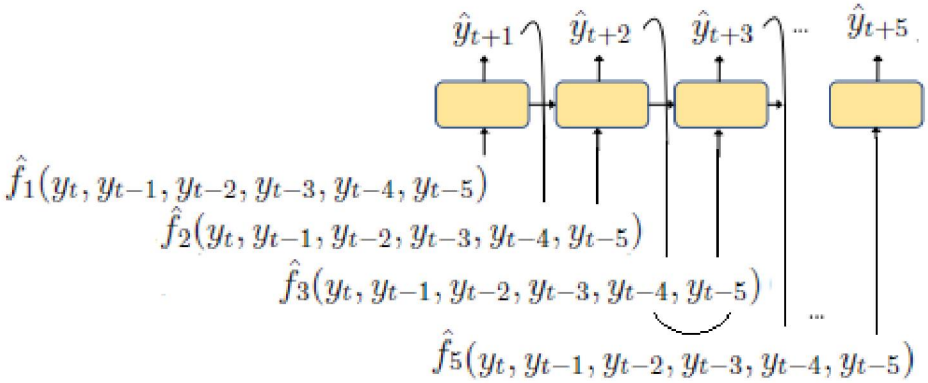
A.2 ESTRATÉGIA DIRETA

A estratégia direta, também conhecida como estratégia independente (TIAO; TSAY, 1994; CHENG et al., 2006), consiste em realizar a previsão de cada passo independentemente da previsão anterior. A mesma utiliza diferentes modelos $\hat{f}$ para cada horizonte de previsão, que são ajustados aos dados reais a medida que estes se tornam disponíveis. A complexidade desta estratégia está na necessidade de identificar os parâmetros do modelo de previsão para cada passo de previsão desejado. Tal estratégia é representada

na Figura A.2 para um horizonte de 5 passos-a-frente, e na Equação A.4,

$$\begin{aligned}
\hat{y}_{t+1} &= \hat{f}_1(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}) \\
\hat{y}_{t+2} &= \hat{f}_2(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}) \\
\hat{y}_{t+3} &= \hat{f}_3(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}) \\
\hat{y}_{t+4} &= \hat{f}_4(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}) \\
\hat{y}_{t+5} &= \hat{f}_5(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5})
\end{aligned} \tag{A.4}$$

Figura A.2: Estratégia direta para previsão 5-passos-a-frente



Fonte: O autor(2019).

A Equação A.5 e A.6 apresenta a generalização da estratégia para K-passos-a-frente, onde é necessário obter K modelos $\hat{f}_k$ (um para cada horizonte) para a série temporal $[y_1, \ldots, y_N]$, onde

$$y_{t+k} = \hat{f}_k(y_t, \ldots, y_{t-d+1}) + \epsilon, \quad t \in \{d, \ldots, N-K\}, \quad e\ k \in \{1, \ldots, K\} \tag{A.5}$$

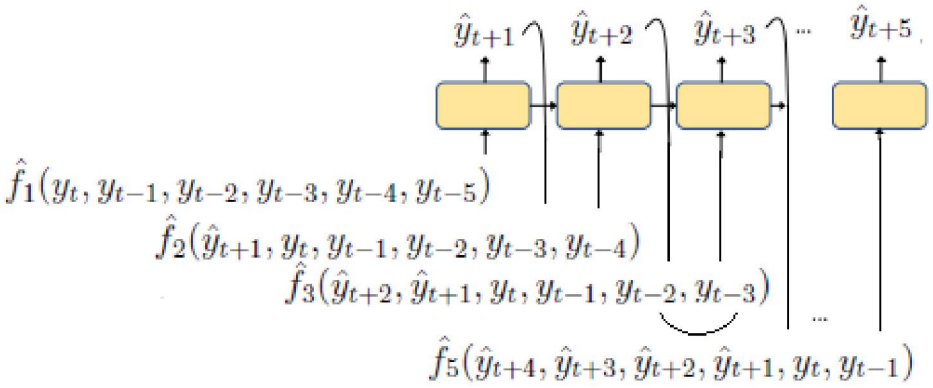
$$\hat{y}_{N+k} = \hat{f}_k(y_N, \ldots, y_{N-d+1}) \tag{A.6}$$

A.3 ESTRATÉGIA DIRREC

A estratégia *DirRec* foi proposta por Sorjamaa e Lendasse (2006) e combina os princípios e arquitetura das estratégias direta e recursiva. Segundo os autores os resultados obtidos pela estratégia *DirRec* supera os obtidos pela recursiva e pela direta. A mesma é apresentada na Equação A.7 e na Figura A.3 para um horizonte de 5 passos-a-frente,

$$
\begin{aligned}
\hat{y}_{t+1} &= \hat{f}_1(y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}) \\
\hat{y}_{t+2} &= \hat{f}_2(\hat{y}_{t+1}, y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}) \\
\hat{y}_{t+3} &= \hat{f}_3(\hat{y}_{t+2}, \hat{y}_{t+1}, y_t, y_{t-1}, y_{t-2}, y_{t-3}) \\
\hat{y}_{t+4} &= \hat{f}_4(\hat{y}_{t+3}, \hat{y}_{t+2}, \hat{y}_{t+1}, y_t, y_{t-1}, y_{t-2}) \\
\hat{y}_{t+5} &= \hat{f}_5(\hat{y}_{t+4}, \hat{y}_{t+3}, \hat{y}_{t+2}, \hat{y}_{t+1}, y_t, y_{t-1})
\end{aligned}
\tag{A.7}
$$

Figura A.3: Estratégia dirrec para previsão 5-passos-à-frente



Fonte: O autor(2019).

A.4 ESTRATÉGIA MIMO

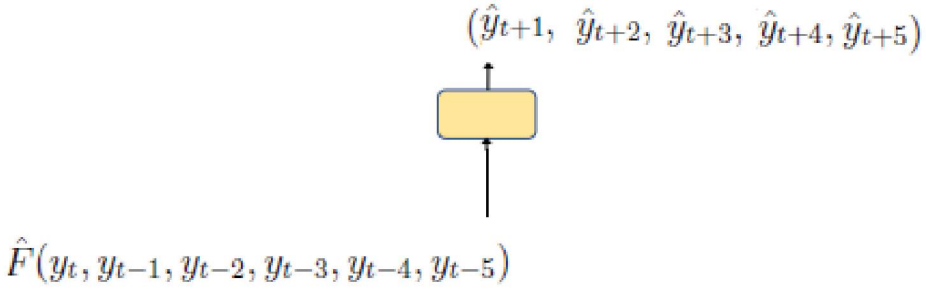
A estratégia com múltiplas entradas e múltiplas saídas (*multiple inputs and multiple output* - MIMO), também chamada de estratégia conjunta (*Joint Strategy*) (KLINE, 2004), possui capacidade de múltiplas saídas. Na estratégia MIMO se faz necessário apenas identificar um único modelo F de múltiplas entradas e múltiplas saídas, mantendo a dependência estocástica entre os valores futuros. A estratégia MIMO possui a estrutura apresentada na Equação A.8,

$$
[y_{t+K}, \ldots, y_{t+1}] = F(y_t, \ldots, y_{t-d+1}) + \epsilon \tag{A.8}
$$

onde $t \in \{d, \ldots, N-K\}$; $F : \mathbb{R}^d \longrightarrow \mathbb{R}^K$ é uma função vetorial e; ϵ é o vetor de ruídos. As previsões resultantes da estratégia MIMO são retornadas em um passo a frente por um modelo com múltiplas saídas $\hat{F}$, conforme Equação A.9 e Figura A.4. O principal fundamento desta estratégia é preservar, entre os valores previstos, a dependência estocástica que caracteriza a série temporal.

$$[\hat{y}_{t+K}, \ldots, \hat{y}_{t+1}] = \hat{F}(y_N, \ldots, y_{N-d+1}) \tag{A.9}$$

Figura A.4: Estratégia mimo para previsão 5-passos-à-frente



Fonte: O autor(2019).

A.5 ESTRATÉGIA DIRMO

Esta estratégia busca preservar os aspectos positivos encontrados nas estratégias direta e MIMO. A DIRMO prevê o horizonte K em blocos, onde cada bloco é previsto de forma MIMO. Assim a tarefa de previsão de K passos-a-frente é decomposta em n previsões de múltiplas saídas, tal que $n = K/s$ e cada saída de tamanho s tal que $s \in \{1, \ldots, K\}$ (TAIEB et al., 2009).

Na estratégia DIRMO quando o valor do parâmetro s é igual a 1 resulta em $n = K$, que corresponde a estratégia direta. Por outro lado, quando s é igual ao valor de K, n será igual a 1, correspondendo a estratégia MIMO. Configurações intermediárias a esses dois valores extremos dependem do valor do parâmetro s. A estratégia DIRMO identifica n modelos F_p a partir da série temporal $[y_1, \ldots, y_N]$ de forma que,

$$[y_{t+p*s}, \ldots, y_{t+(p-1)*s+1}] = F_p(y_t, \ldots, y_{t-d+1}) + \epsilon \tag{A.10}$$

onde $t \in \{d, \ldots, N-K\}$; $p \in \{1, \ldots, N-K\}$; $F_p : \mathbb{R}^d \longrightarrow \mathbb{R}^s$ é uma função vetorial se $s > 1$ e; ϵ é o vetor de ruídos.

As previsões para K passos-a-frente são retornadas pelos n modelos identificados como representado na Equação A.11,

$$[\hat{y}_{N+p*s}, \ldots, \hat{y}_{t+(p-1)*s+1}] = \hat{F}_p(y_N, \ldots, y_{N-d+1}) \tag{A.11}$$

Em Taieb et al. (2012) é possível encontrar um comparação entre as 5 estratégias apresentadas. Os autores avaliaram o custo computacional, as vantagens e desvantagens de cada estratégia.

APÊNDICE B – MODELAGEM BOX E JENKINS

A relação temporal considerada na modelagem e Box & Jenkins (BOX et al., 1976) é representada formalmente por um conjunto de processos estocásticos genericamente denominados modelos ARIMA, normalmente classificados como modelos univariados. O modelo resulta da combinação de três componentes, que atuam como filtros, tem-se a componente Auto Regressiva (AR), o filtro de Integração (I) e a componente de Médias Móveis (MA). Uma série temporal pode conter os três filtros ou apenas um subconjunto deles, resultando em várias alternativas de modelos passíveis de análise pela metodologia Box & Jenkins. O modelo ARIMA é fundamentado em conceitos estatísticos, cuja fundamentação é apresentada a seguir.

Seja a série temporal $\{Y_t\}_{t=1}^{T}$, onde T é o comprimento da série, e $\{y_t\}_{t=1}^{T}$ uma realização do processo estocástico, ou seja, uma série temporal estacionária de segunda ordem que não apresenta movimentos sazonais. O modelo Box & Jenkins que representa o comportamento dinâmico desta série é definido pela Equação B.1

$$y_t = \varphi y_{t-1} + \ldots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} \tag{B.1}$$

Este modelo é denominado de modelo ARMA(p,q), onde p representa a ordem auto regressiva, q a ordem das médias móveis, $\varphi_1, \ldots, \varphi_p$ representa os parâmetros auto-regressivos e $\theta_1, \ldots, \theta_q$ os parâmetros de médias móveis. Caso derivados: Se $q = 0$ e $p \neq 0$, tem-se o modelo Auto Regressivo AR(p), representado na Equação B.2, onde c é uma constante a ser estimada e ε_t é considerado aproximadamente um ruído branco $\approx RB(0, \sigma^2)$; Se $p = 0$ e $q \neq 0$ o modelo resultante é o de Médias Móveis MA(q) conforme Equação B.3, onde μ é uma constante (BOX et al., 1976).

$$y_t = c + \varphi_1 y_{t-1} + \ldots + \varphi_p y_{t-p} + \varepsilon_t \tag{B.2}$$

$$y_t = \mu - \theta_1 y_{t-1} - \ldots - \theta_q y_{t-q} + \varepsilon_t \tag{B.3}$$

A Equação B.1 pode ser reescrita em termos de dois polinômios, basta aplicar o operador de diferença, representado por $\nabla = (1 - B)$, de forma que B é o operador de atraso (lag), e é definido como apresentado na Equação B.4, resultando desta forma no modelo da Equação B.5

$$B^d y_t = y_{t-d} \quad d = 1, , 2, \ldots \tag{B.4}$$

$$(1 - \varphi_1 B - \ldots - \varphi_p B^p)(y_t - \mu) = (1 - \theta_1 B - \ldots - \theta_q B^q)\varepsilon_t \tag{B.5}$$

A verificação da estacionariedade de segunda ordem pode ser realizada através da análise gráfica da função de Auto Correlação (fac) ρ, que é a Auto Covariância padroniza (γ_d), representada na Equação B.6, que serve para medir o comprimento e a memória de um processo, ou seja, a extensão para qual o valor tomado no tempo t depende daquele tomado no tempo $t - d$,

$$\rho_d = \frac{\gamma_d}{\gamma_0} = \frac{Cov[Y_t, Y_{t-d}]}{\sqrt{Var(Y_t)Var(Y_{t-d})}} \qquad d = 0, 1, 2, \ldots \tag{B.6}$$

É facil provar que $\rho_0 = 1$ e $\rho_d = \rho_{-d}$, pois a função de Auto Correlação é uma função par, desta forma a Equação B.6 é um estimador amostral da Auto Correlação de atraso (ou lag) d é necessário determinar γ_d apenas para $d > 0$. Lembrando que o operador de atraso (lag) é dado pela Equação B.4.

Através do gráfico da função de Auto Correlação (fac) pode-se identificar a estrutura do modelo, de forma que:

- Nas estruturas AR(p): a fac decai de acordo com exponenciais e/ou exponenciais e/ou senóides amortecidas, infinita em extensão;
- Nas estruturas MA(q): a fac é finita para $lag < q$, no sentido de que ela apresenta um corte, ou se anula bruscamente, após o $lag > q$;
- Nas estruturas ARMA(p,q): a fac tem infinita extensão, a qual decai de acordo com exponenciais e/ou senóides amortecidas após o $lag(q - p)$.

Existe ainda a função de Autocorrelação Parcial (*facp*) que mede a correlação entre duas observações da série temporal, Y_t e Y_{t+d} por exemplo, eliminando a dependência dos termos intermediários $Y_{t+1}, Y_{t+2}, \ldots, Y_{t+d-1}$, tem-se a autocorrelação parcial, representada por $Cov\Big(Y_t, Y_{t-d}|Y_{t-1}, \ldots, Y_{t-(d+1)}\Big)$. Em linhas gerais, a idéia da função Autocorrelação Parcial é medir quanto Y_t e Y_{t+d} estão correlacionados, mas com os efeitos dos valores intermediários controlados. Por exemplo, caso se deseje avaliar o relacionamento entre o par (Y_t, Y_{t+2}), mantém-se fixo o efeito de Y_{t+1} sobre Y_{t+2}. Em seguida, desejando-se avaliar o relacionamento para o par (Y_t, Y_{t+3}) deve-se manter fixo os efeitos de Y_{t+1} e Y_{t+2} sobre Y_{t+3}, e assim sucessivamente.

A definição da *facp* parte da solução do Sistema de Equações de Yule-Walker, representada na Equação B.7,

$$\begin{cases} \rho_1 &= \varphi_1 + \varphi_2\rho_1 + \ldots + \varphi_p\rho_{p-1} \\ \rho_2 &= \varphi_1\rho_1 + \varphi_2 + \ldots + \varphi_p\rho_{p-2} \\ \vdots & \qquad\qquad \vdots \\ \rho_p &= \varphi_1\rho_{p-1} + \varphi_2\rho_{p-2} + \ldots + \varphi_p \end{cases} \tag{B.7}$$

Reescrevendo a Equação B.7 na forma representada pela Equação B.8, e resolvendo-a para $k = 1, 2, \ldots$, tem-se a definição da função de Autocorrelação Parcial.

$$\begin{vmatrix} 1 & \rho_1 & \ldots & \rho_{k-1} \\ \rho_1 & 1 & \ldots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \ldots & 1 \end{vmatrix} \times \begin{vmatrix} \varphi_{k1} \\ \varphi_{k2} \\ \ldots \\ \varphi_{kk} \end{vmatrix} = \begin{vmatrix} \rho_1 \\ \rho_2 \\ \ldots \\ \rho_k \end{vmatrix} \tag{B.8}$$

Para caso da série temporal não possuir estacionariedade, a metodologia de Box & Jenkins pode ser aplicada, desde que a mesma se torne estacionária após a aplicação do operador de diferenças ∇. O número de diferenças necessárias para tornar uma série estacionária é chamado ordem de integração da série. Um processo é dito ser integrado de ordem d se é não estacionária, mas a série resultante após d diferenças é estacionária. Desta forma, a série resultante das d diferenças pode ser modelada por um modelo ARMA(p,q), conforme Equação B.9, dizemos neste caso que a série temporal y_t é um processo Auto Regressivo Integrado de Médias Móveis - ARIMA(p,d,q) com $d > 0$ e inteiro.

$$(1 - \varphi_1 B - \ldots - \varphi_p B^p)[(1 - B^d)y_t - \mu)] = (1 - \theta_1 B - \ldots - \theta_q B^q)\varepsilon_t \tag{B.9}$$

Quando d assume valores fracionários, o processo pode ter comportamento de longa dependência, neste caso, $d \in (0, 0.5)$, para o caso de curta dependência $d \in (-0.5, 0)$. Modelos com essas características são denominados como modelos Autor Regressivos Integrados Fracionalmente e de Médias Móveis (ARFIMA).

Para o caso de séries temporais que apresentem sazonalidade, os modelos de Box & Jenkins também podem ser aplicados. Neste caso, o modelo é representado conforme Equação B.10

$$\varphi_B(1 - \phi_1 B - \ldots - \phi_p B^p S)\nabla^d(1 - B^S)^D y_t = \theta_B(1 - \Theta_1 B - \ldots - \Theta_Q B^Q S)\varepsilon_t \quad \text{(B.10)}$$

onde $\varphi_B = (1 - \varphi_1 B - \ldots - \varphi_p B^p)$; $\theta_B = (1 - \theta_1 B - \ldots - \theta_q B^q)$; d são os polinômios das diferenças simples; D é a ordem das diferenças sazonais; S o período sazonal. Os coeficientes dos polinômios não sazonais são representados por $\phi_k \in \mathbb{R}$ e $\theta_j \in \mathbb{R}$, enquanto que $\phi_m \in \mathbb{R}$ e $\Theta_n \in \mathbb{R}$ representam os polinômios sazonais. Os modelos sazonais são descritos como Modelos Sazonais Autor Regressivos Integrados e de Médias Móveis - SARIMA(p,d,q)$\times(P, D, Q)_S$. Se pelo menos um dois parâmetros d ou D assume valores fracionários, o modelo é conhecido como Sazonal Auto Regressivo Fracionalmente Integrado e de Médias Móveis - SARFIMA(p,d,q)$\times(P, D, Q)_S$.

B.1 CRITÉRIO DE INFORMAÇÃO BAYESIANO - BIC

Quando a identificação dos modelos de Box & Jenkins não é trivial através da análise dos gráficos da fac ou da facp, uma possível solução é utilizar o Critério de Informação Bayesiano (BIC) para selecionar os graus p e q de um modelo ARMA. Para isso deve-se estimar vários modelos com diferentes valores de p e q, porém para cada modelo estimado, o valor da função objetivo deve ser mínimo *loglikelihood*. O BIC penaliza a complexidade do modelo, ou seja, os modelos que possuem o menor número de parâmetros apresentam o menor valor BIC, que é calculado através da EquaçãoB.11

$$BIC = -2\log f(x_n|\theta) + p\log n \quad \text{(B.11)}$$

em que $f(x_n|\theta)$ é o modelo escolhido, p é o número de parâmetros a serem estimados e n é o número de observações da amostra.

APÊNDICE C – RELAÇÃO DE RECORRÊNCIA LINEAR - LRR

O algoritmo de Relação de Recorrência Linear (LRR) basicamente consiste em considerar $\mathbf{X}_N = \mathbf{X}_N^{(1)} + \mathbf{X}_N^{(2)}$ e supor que $\mathbf{X}_N^{(1)}$ admite continuação recorrente. Então, denotando por d a dimensão da miníma relação de recorrência que governa $\mathbf{X}_N^{(1)}$, se $d < min(L, N - L + 1)$ então $d = rank_L(\mathbf{X}_N^{(1)})$. Desta forma, se $\mathbf{X}_N^{(1)}$ e $\mathbf{X}_N^{(2)}$ são fortemente separáveis para alguma janela de comprimento L, então o espaço de trajetórias de $\mathbf{X}_N^{(1)}$ pode ser encontrado e a continuação recorrente da série $\mathbf{X}_N^{(1)}$ pode ser feita. Essa afirmação, conforme (GOLYANDINA et al., 2001), advém do fato de que se o último valor da última componente x_L do vetor $\mathbf{X} = (x_1, \ldots, x_L)^T \in L$ é uma combinação linear das primeiras componentes $x_1, \ldots, x_{L-1}$, isso significa que x_L pode ser escrito conforme Equação C.1

$$x_L = a_1 x_{L-1} + a_2 x_{L-2} + \ldots + a_{L-1} x_1 \tag{C.1}$$

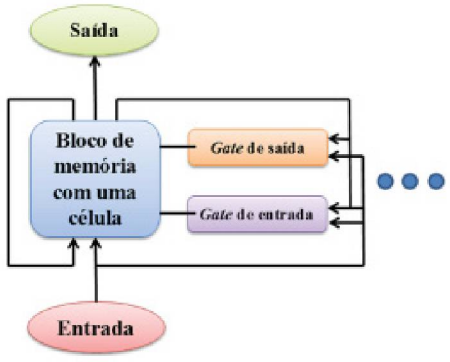
Se não há garantias de exata separabilidade, então o algoritmo de previsão usualmente retorna diferentes valores. Cabe aqui ressaltar que o procedimento para a escolha da janela L não é trivial, como já citado anteriormente. Como o método de previsão utiliza L para encontrar a relação de recorrência LRR, se a escolha das componentes de $\mathbf{X}_N^{(1)}$ é adequada então $r = d$, isto significa que o *rank* de todas as matrizes $\mathbf{X}^{(r)}$ é igual a dimensão da minima relação de recorrência que governa $\mathbf{X}_N^{(1)}$.

Por outro lado, se a escolha não é adequada, produzirá erro na LRR que governa as séries utilizadas na previsão.

APÊNDICE D – REDES NEURAIS LSTM

As rede neurais LSTM (*Long Short Term Memory*) possuem uma característica importante que a permite aprender ou guardar dependências de informações no longo termo. Elas foram introduzidas em 1997 por Hochreiter e Schmidhuber (1997), e a partir da última década vários pesquisadores contribuíram para o desenvolvimento e aprimoramento da mesma. A rede LSTM tem uma característica peculiar que minimiza o problema de desvanecimento do gradiente, forçando um fluxo de erro constante através de unidades que conservam o sinal de erro, por isso tem larga aplicações em dados sequências como séries temporais, por exemplo. Essas unidades estão localizadas dentro dos blocos de memória e evitam o decaimento do fluxo do erro quando o mesmo é retroprogado (LI et al., 2019).

Figura D.1: Estrutura da rede LSTM como 1 bloco de memória na camada escondida

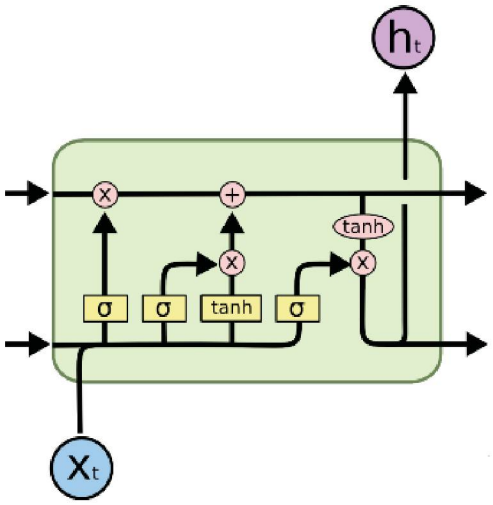


Fonte: O autor(2019) - adaptado de (SALAZAR, 2015).

A idéia geral de uma rede LSTM e a forma com que a informação se propaga através da mesma é apresentada a seguir. Como o intuito desta seção é apresentar uma idéia básica sobre os conceitos que envolvem a rede LSTM, não será detalhado todas as variantes existentes da mesma. Para os leitores que desejam mais detalhes por favor se dirija as referências (HOCHREITER; SCHMIDHUBER, 1997; SALAZAR, 2015; LI et al., 2019).

As principais etapas de fluxo da informação internamente à rede são a etapa ou passo de propagação e a retroprogação. A propagação na LSTM equivale ao mesmo princípio encontrado nas tradicionais redes neurais recorrentes (RNN) (SALAZAR, 2015). Porém a diferença entre ambas esta no fato da LSTM possuir módulos de repetição, que são compostos por quatro camadas que interagem de forma especial, como apresentado na Figura D.2,

Figura D.2: Estrutura do módulo de repetição da LSTM



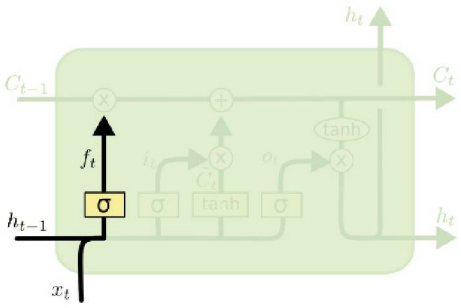
Fonte: O autor(2019) - adaptado de (SALAZAR, 2015).

O estado célula representado é pela linha horizontal superior que cruza o diagrama do módulo de repetição. Ele atravessa toda a cadeia inteira sofrendo alterações provocadas pelas estruturas denominadas de *gates* (LI et al., 2019). A rede LSTMA tem habilidade de remover ou adicionar informação ao estado da célula, sendo este processo regulado pelos *gates* que são compostos por uma camada *sigmoidal* e uma operação de multiplicação. Lembrando que função sigmoidal apresenta saídas limitadas em 0 e 1, e no LSTM são utlizadas na formação do *forget gate layer* que representa a operação de deixar (valor 1) ou não o sinal passar (valor 0). O LSTM possui três *gates* como esse para controlar e proteger o estado da célula, o mesmo é representado pela operação vetorial representada na Equação D.1 e graficamente na Figura D.3 (SALAZAR, 2015).

$$f_t = \sigma\big(W_f(h_{t-1}, x_t) + b_f\big) \tag{D.1}$$

onde W_f é a matriz de parâmetros que recebe h_{t-1} e x_t; b_f é um valor de correção.

Figura D.3: Estrutura do módulo de repetição da LSTM



Fonte: O autor(2019) - adaptado de (SALAZAR, 2015).

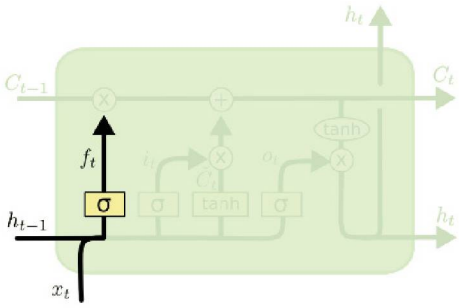
O passo seguinte consiste decidir qual nova informa será armazenada no estado da célula. Este processo é divido em duas etapas, onde na primeira etapa consiste em decidir quais valores serão atualizados, isso é feito através da camada *sigmoid* chamada de *input gate*. A etapa seguinte consiste em criar um vetor com valores candidatos de $\tilde{C}_t$ que podem ser adicionadas ao estado, representado pelas Equações D.2 e D.3 e pela Figura D.4, sendo a combinação desses dois utilizada para criar uma atualização para o estado.

$$i_t = \sigma\big(W_i(h_{t-1}, x_t) + b_i\big) \tag{D.2}$$

$$\tilde{C}_t = \tanh\big(W_c(h_{t-1}, x_t) + b_c\big) \tag{D.3}$$

onde W_i e W_c são matrizes de parâmetros que recebem h_{t-1} e x_t; b_i e b_c são valores de correção.

Figura D.4: Estrutura do *input gate* da LSTM



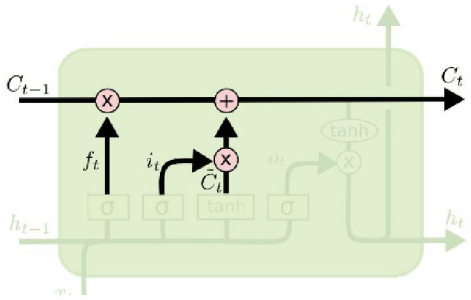
Fonte: O autor(2019) - adaptado de (SALAZAR, 2015).

A atualização do estado anterior da célula para o estado atual, consiste em atu-

alizar C_{t-1} para o valor de C_t. Isto é feito multiplicando o estando anterior f_t, que na realidade representa a operação de *esquecer* o se não é relevante, e enão adiciona-se o termo $i_t * \tilde{C}_t$. Esta operação esta representada na Equação ??

$$C_t = f_t * C_{t-1} + i_t \tilde{C}_t \tag{D.4}$$

Figura D.5: Estrutura da etapa de atualização da célula LSTM



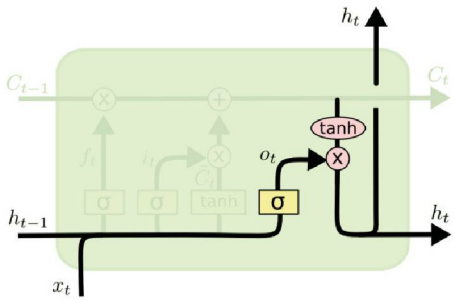
Fonte: O autor(2019) - adaptado de (SALAZAR, 2015).

A ultima etapa do fluxo de informação, consiste na decisão sobre qual informação será a saída da célula. Este processo se dá na camada *sigmoid* que decidirá quais partes comporão a saída. Em seguida o estado atual da célula é apresentado a função *tanh*, de forma que os valores sejam mapeados para os intervalos $[-1, 1]$, multiplica-se este resultado pela saída do *sigmoid gate*, conforme Equação D.6, resultando desta forma, na saída desejada.

$$o_t = \sigma\big(W_o(h_{t-1}, x_t) + b_o\big) \tag{D.5}$$

$$h_t = o_t * \tanh\big(C_t\big) \tanh \tag{D.6}$$

Figura D.6: Estrutura da etapa de saída da célula LSTM



Fonte: O autor(2019) - adaptado de (SALAZAR, 2015).

APÊNDICE E – OTIMIZADOR Adam

O otimizador baseado na estimativa adaptativa do momento de menor ordem, cuja abreviatura é a sigla Adam (do inglês, *adaptive moment estimation*), é usado para treinar os pesos desconhecidos na rede neural LSTM. O algoritmo Adam foi proposto por Kingma e Ba (2015) para otimização baseada em gradiente de primeira ordem de funções objetivas estocásticas, baseado em estimativas adaptativas de momentos de menor ordem. O método pode ser divido em 5 passos e é bem simples de implementar, além de ser computacionalmente eficiente, pois tem poucos requisitos de memória, é invariante para o reescalonamento diagonal dos gradientes e é adequado para problemas que são grandes em termos de dados e/ou parâmetros.

O método também é apropriado para objetivos não estacionários e problemas com gradientes muito ruidosos e/ou esparsos. Os hiper-parâmetros possuem interpretações intuitivas e normalmente requerem pouco ajuste. Para melhor entendimento do método, considere a função objetivo $f(\theta)$, α o tamanho do passo, β_1 e β_2 a taxa decaimento exponencial compreendida no intervalo $[0, 1)$. Os símbolos m_t e v_t são os vetores dos momentos de primeira e segunda ordem na $t - th$ iteração, respectivamente. Desta forma, o vetor θ de valores desconhecidos pode ser atualizado iterativamente utilizando as Equações E.1 à E.5,

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla f(\theta_t) \tag{E.1}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\nabla^2 f(\theta_t) \tag{E.2}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{E.3}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{E.4}$$

$$\theta_{t+1} = \theta_t + \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{E.5}$$

onde η é a taxa de aprendizagem e ϵ é uma constante para evitar divisão por zero. Os valores nornalmente utilizados para β_1 e β_2 são $0,9$ e $0,999$; respectivamente. Já η assume valor normalmente igual à $0,001$. Os passos do Algoritmo E.1 a seguir representam o método iterativo de cálculo.

Algoritmo E.1 Otimizador Adam

1: Entradas: Parâmetros iniciais
2: Calcule o gradiente e seu quadrado elementar usando os parâmetros iniciais (Eq. E.1 e E.2)
3: Atualize a média móvel exponencial do momento de primeira ordem e do momento de segunda ordem
4: Calcule uma média (não-viesada) do momento de primeira ordem e do momento de segunda ordem (Eq. E.3 e E.4)
5: Atualização de peso de cálculo: média do momento de primeira ordem dividida pela raiz quadrada de média do momento de segunda ordem multiplicada pelo fator de escala da taxa de aprendizado (Eq. E.5)
6: Aplique a atualização aos pesos

Fonte: O autor(2019)

APÊNDICE F – PUBLICAÇÕES E PREMIAÇÃO

Ao longo do desenvolvimento desta tese vários trabalhos foram publicados, englobando a pesquisa desenvolvida não só relacionada ao tema da tese, mas também a aprendizado de máquina, otimização, identificação de sistemas, entre outros tópicos. O autor também teve um de seus artigos reconhecido e premiado como o melhor artigo do 50 Simpósio Brasileiro de Pesquisa Operacional - SOBRAPO. A seguir é apresentado a lista de contribuições do autor ao longo do desenvolvimento da tese.

F.1 TRABALHOS PUBLICADOS

MORENO, S.; KLEIN, C. E.; MARIANI, V. C.; COELHO, L.; ALOTTO, P.; BAUERNFEIND, T.; BAUMGARTNER, P. Non-uniformly spaced linear antenna array design by means of peec approach applying cheetah optimization algorithm. *International Journal of Applied Electromagnetics and Mechanics*, v. 60, p. 15–24, 2019.

MORENO, S. R.; COELHO, L. dos S. Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renewable Energy*, v. 126, p. 736–754, 2018.

TSILI, M.; AMOIRALIS, E. I.; LEITE, J. V.; MORENO, S. R.; COELHO, L. dos S. A novel multiobjective lognormal-beta differential evolution approach for the transformer design optimization. *Engineering Computations*, v. 35, n. 2, p. 955–978, 2018.

NETO, J. X. V.; JUNIOR, E. J. G.; MORENO, S. R.; AYALA, H. V. H.; MARIANI, V. C.; COELHO, L. dos S. Wind turbine blade geometry design based on multiobjective optimization using metaheuristics. *Energy*, v. 162, p. 645 – 658, 2018.

MORENO, S. R.; JUNIOR, O. da S.; COELHO, L. dos S. Adaptive imperialist competitive algorithm applied to wind farm layout optimization. In: *L Simpósio Brasileiro de Pesquisa Operacional, L SBPO.* [S.l.]: Sociedade Brasileira Pesquisa Operacional, SOBRAPO, Rio de Janeiro, RJ, 2018. v. 50.

MORENO, S.; KLEIN, C.; MARIANI, V. C.; COELHO, L. D. S.; ALOTTO, P.; BAUERNFEIND, T.; BAUMGARTNER, P. Non-uniformly spaced linear antenna array design by means of peec approach applying cheetah optimization algorithm. In: *15th International Workshop on Optimization and Inverse Problems in Electromagnetism (OIPE 2018).* [S.l.: s.n.], 2018. v. 1, p. 38–39.

LUCENA, D. S. de; MORENO, S.; MARIANI, V. C.; COELHO, L. Support vector machine optimized by artificial bee colony applied to eeg pattern recognition. In: *Proceedings in Lecture Notes in Artificial Intelligence.* [S.l.]: Springer, Heidelberg, 2017. v. 1, p. 1–2.

F.2 PREMIAÇÃO

O autor recebeu o Prêmio Roberto Diéguez Galvão (PRDG) de melhor artigo no 50 Simpósio Brasileiro de Pesquisa Operacional - L SBPO, pelo trabalho publicado relativo a otimização do *layout* de parques eólicos utilizando heurísticas populacionais, conforme referência a seguir:

MORENO, S. R.; JUNIOR, O. da S.; COELHO, L. dos S. Adaptive imperialist competitive algorithm applied to wind farm layout optimization. In: *L Simpósio Brasileiro de Pesquisa Operacional, L SBPO.* [S.l.]: Sociedade Brasileira Pesquisa Operacional, SOBRAPO, Rio de Janeiro, RJ, 2018. v. 50..