

Self-attention Auto Encoder para previsão de séries temporais de potência.

Mateus G. Machado

Centro de Informática (CIn)
Universidade Federal de Pernambuco (UFPE)
Recife, Brazil
mgm4@cin.ufpe.br

Renan S. S. Costa

Centro de Informática (CIn)
Centro de Energias Renováveis (CER)
Universidade Federal de Pernambuco (UFPE)
Recife, Brazil
rssc@cin.ufpe.br

Abstract—A previsão eólica a curto prazo é de extrema importância para o bom gerenciamento do Sistema Interligado Nacional, bem como para compra e venda de energia nos mercados diários e intradiários de energia. Dessa forma, o desenvolvimento de modelos acurados torna-se essencial na dinâmica do setor elétrico. Nesse contexto, modelos de aprendizagem de máquina são amplamente utilizados devido ao bom desempenho na extração das características da dinâmica dos fenômenos eólicos. Neste trabalho, nós combinamos modelos de atenção com a técnica de pré-processamento desenvolvida por Cabral [1] com foco em prever séries temporais de potência de centrais eólicas usando poucos dados. Também avaliamos os modelos com conjuntos de dados bem conhecidos na literatura para saber a aplicabilidade dos mesmos em séries temporais de outros tipos¹.

Index Terms—wind energy, wind power, machine learning, neural networks, attention

I. INTRODUÇÃO

No início de 2020, a capacidade instalada de energia eólica no mundo atingiu 650.8 GW, sendo 59,7 GW adicionados somente ao longo de 2019 [2]. No Brasil, até fevereiro de 2021, a capacidade instalada atingiu a marca de 18 GW, representando 10,3% da matriz elétrica brasileira [3]. Quanto a geração de energia, em 2019 o Brasil produziu 55,9 TWh, correspondente a 9,7% de toda a geração injetada no Sistema Interligado Nacional (SIN) referente ao período [3].

O órgão responsável pela coordenação e controle da operação das instalações de geração e transmissão da energia elétrica no SIN é o Operador Nacional do Sistema Elétrico (ONS). Dentre suas incumbências, está a de gerenciar o despacho da produção de energia das fornecedoras de modo a manter a estabilidade da rede.

Por se tratar de uma fonte renovável, a produção eólica tem prioridade no despacho. Porém, com a crescente inserção de uma fonte intermitente no SIN, o gerenciamento da rede elétrica torna-se uma tarefa árdua e faz-se necessário estimar a futura produção de energia a fim de evitar sobrecarga do sistema, bem como o déficit de fornecimento. Para a realização desta tarefa, é de extrema importância manter-se informado, com o maior nível de acurácia possível, sobre o comportamento da produção energética de modo a ter mais

tempo para conectar outras fontes de energia, que, devido a suas inércias, podem demorar horas para chegarem ao nível ideal de produção.

Nesse contexto, o desenvolvimento de bons modelos de previsão torna-se indispensável para uma melhor gestão dos recursos energéticos que consequentemente garantem uma maior estabilidade do SIN. Como os métodos estatísticos possuem melhor desempenho para previsão eólica a curto prazo [4], este trabalho se dedicará a explorar a previsão a curto prazo através de técnicas de decomposição de séries temporais e aprendizagem de máquina em dados reais.

Baseado na dissertação de mestrado de Cabral [1], propomos neste trabalho uma reformulação do modelo que se utiliza da decomposição realizada para melhorar o desempenho do método em relação a séries com menos de 18 mil observações enquanto mantém a performance para séries com maiores. Utilizamos de uma arquitetura baseada em mecanismos de *self-attention* [5] e *Transfer Learning* para modelar uma representação temporal da série. Para os resultados, serão comparados o modelo proposto e MLPs em três tipos de conjuntos de dados: séries temporais de potência utilizando um ano de observações; utilizando dois meses de observações; séries temporais utilizadas comumente na literatura com 90% menos observações que as séries de potência. Para cada conjunto de dados aplicaremos dois cenários: com o pré-processamento através da decomposição da série original e sem o pré-processamento.

II. BACKGROUND

1) Decomposição de séries temporais baseada em análise espectral: Cabral descreveu um modelo de decomposição da série temporal a partir da análise da Transformada Discreta de Fourier (DFT) da série para identificar componentes periódicas de menor frequência [1]. As N menores frequências de maior amplitude são selecionadas para fazer a decomposição da série segundo a Equação 1. Após geradas as componentes, uma nova análise e decomposição é aplicada a cada componente individual, obtendo-se as subcomponentes.

Cada subcomponente, então, passa a servir de entrada para uma rede neural cujo objetivo é modelar seu comportamento individual para que, em seguida, sua previsão seja entrada

¹O código está disponível em https://github.com/goncamateus/time_series

de uma última camada de Multi-layer Perceptron (MLP) que servirá para recompor o sinal. O diagrama esquemático do modelo é apresentado na Figura 1.

$$c_{i,j} = \frac{1}{T_i} \sum_{k=j}^{T_i+j-1} \left(O_k - \sum_{l=1}^{i-1} c_{l,j} \right) \quad (1)$$

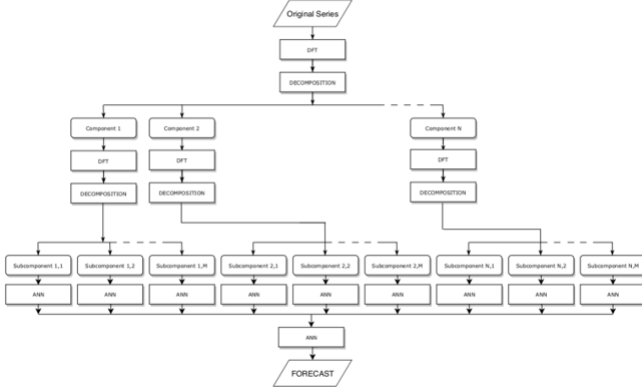


Fig. 1: Diagrama esquemático do algoritmo completo utilizado por [1].

Apesar do modelo de dupla decomposição exibir uma excelente performance quando comparado com a persistência e MLPs tradicionais [1], o algoritmo é muito custoso computacionalmente, pois, supondo M componentes e N subcomponentes cada, o modelo proposto necessita do treinamento de $(M*N)+1$ redes neurais.

2) **Redes transformers aplicadas a problemas de séries temporais:** Modelos de arquitetura Transformer foram introduzidos para previsão/classificação de texto [5]. A arquitetura utiliza mecanismos de *self-attention* com *embeddings* posicionais para atribuir uma sequência de palavras a outra sequência (modelos Seq2Seq).

Wu *et al.* [6] baseou seu modelo na arquitetura Transformer para prever séries temporais para um horizonte. Veja na Figura 2 a arquitetura geral do modelo descrito em [6]. Os resultados mostram que a utilização de modelos recursivos para os modelos Encoder combinado ao *embeddings* posicional permitiu uma representação temporal boa. Apesar dos ótimos resultados, o modelo exige grandes recursos computacionais e prevê apenas um horizonte. Nossas séries temporais serão previstas para múltiplos horizontes.

III. MÉTODO

Nesta seção descrevemos o método que intitulamos de Self-Attention Auto Encoder for Time Series (SAETS). O método se divide em um Transformer modificado e um modelo que recebe apenas o Encoder deste Transformer com seus pesos congelados treinando apenas uma camada totalmente conectada para prever os N horizontes à frente. As entradas do nosso modelo são as componentes principais geradas pela decomposição de Cabral. Na Figura 3 mostramos a arquitetura geral do nosso modelo. Optamos por aplicar apenas a primeira

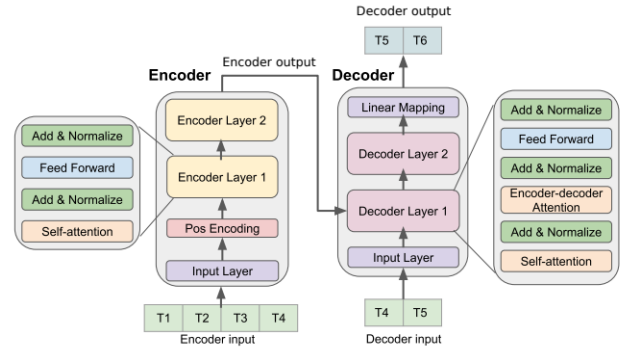


Fig. 2: Arquitetura Transformer utilizada em [6].

decomposição pelo fato do método aproximar as componentes a zero quando a série em análise tem menos que 3000 observações. Assim, esperamos que o embedding do encoder faça a expansão de dimensionalidade representando melhor as componentes principais e a série temporal.

1) **Self-Attention Auto Encoder:** Baseamos nosso modelo Transformer na arquitetura de [6]. Partimos da ideia de utilizar os horizontes a serem previstos em conjunto com o *positional encoding* como as chaves do nosso Transformer Decoder. A hipótese é que desta forma o Encoder consiga representar as componentes da série de forma que não tenha tanta ligação em relação a posição das componentes e sim aos N horizontes que queremos prever. Dessa forma haveria uma dependência temporal em relação aos horizontes porém não em relação às componentes. Optamos também por utilizar apenas MLPs na nossa arquitetura para tornar um modelo menos complexo e mais veloz, uma vez que já estamos introduzindo a dependência temporal no Decoder. O módulo está representado na Figura 3a.

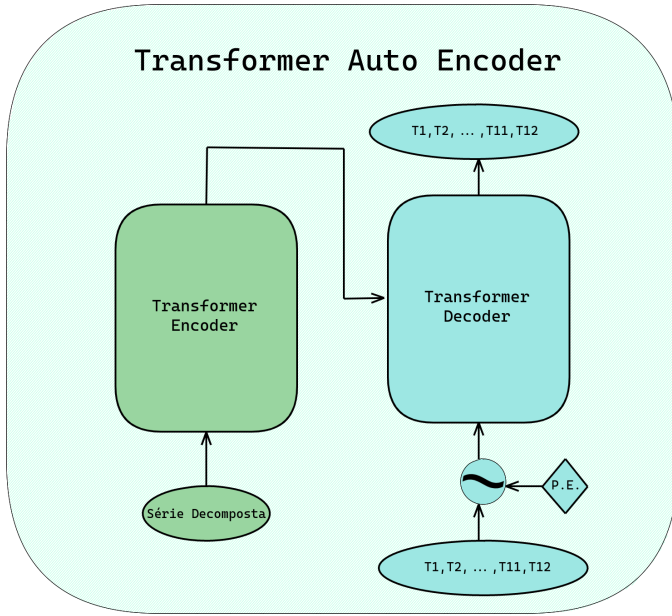
2) **Transfer Learning:** Após o Treinamento do nosso Auto Encoder (AE), utilizamos de Transfer Learning congelando os pesos do Encoder e adicionamos apenas uma camada de MLP para prever os N horizontes a frente. O processo é necessário pois nosso Decoder utiliza da série que queremos prever como entrada, algo que não é possível de ocorrer em casos reais ou no conjunto de testes. O módulo está representado na Figura 3b.

IV. CONJUNTO DE DADOS

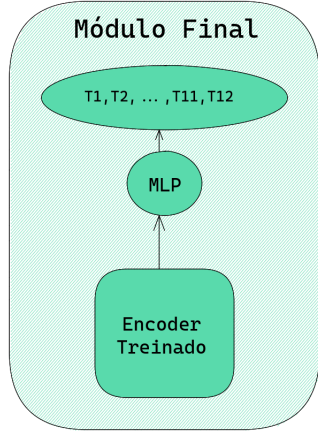
Dividimos os conjuntos de dados em dois tipos: séries de potência de centrais eólicas e outros tipos de séries temporais. Descrevemos nesta seção ambos tipos porém com mais ênfase na séries de potência, uma vez que é o foco do trabalho.

A. Séries de potência de centrais eólicas

1) **Localização das centrais:** A Austrália possui recursos eólicos significativos devido aos ventos do oeste na região sul [7]. O conjunto de dados utilizados consiste em oito séries temporais de potência de saída de centrais eólicas onshore localizadas no sul da Austrália, nas regiões de Victoria (YAMBUKWF, MERCER01 e KIATAWF1), Tasmania



(a) Transformer Auto Encoder utilizado no SAETS.



(b) Módulo final usado após o treinamento do Transformer.

Fig. 3: Arquitetura geral utilizada no nosso modelo. O módulo PE diz respeito ao *Positional Encoding*.

(MUSSELR1), South Australia (WATERLWF, STARHLWF e NBHWF1), New South Wales (GUNNING1). Os dados cobrem centrais eólicas com diferentes potências nominais: respectivamente, temos 30 MW, 131 MW, 31 MW, 168 MW, 131 MW, 34 MW, 132 MW e 46 MW. A figura 4 mostra a localização aproximada de cada parque.

2) **Séries de potência:** Os dados compreendem o período de um ano de observações com amostras em intervalos de 5 minutos, de 01/01/2018 00:00 (UTC +10) a 31/12/2018 23:55 (UTC +10) e foram disponibilizados pela Australian Energy Market Operator (AEMO). A figura 5 mostra os gráficos correspondentes a cada uma das oito séries temporais. Nota-se que são séries com o comportamento padrão para potência eólica, sem grandes períodos de parada ou limitações por parte do operador do sistema elétrico, bem como não há grandes períodos de dados ausentes, não ocasionando perdas

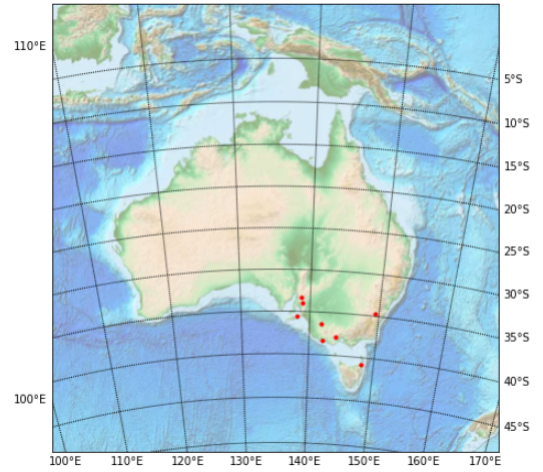


Fig. 4: Localização das centrais eólicas.

significativas no desempenho dos modelos.

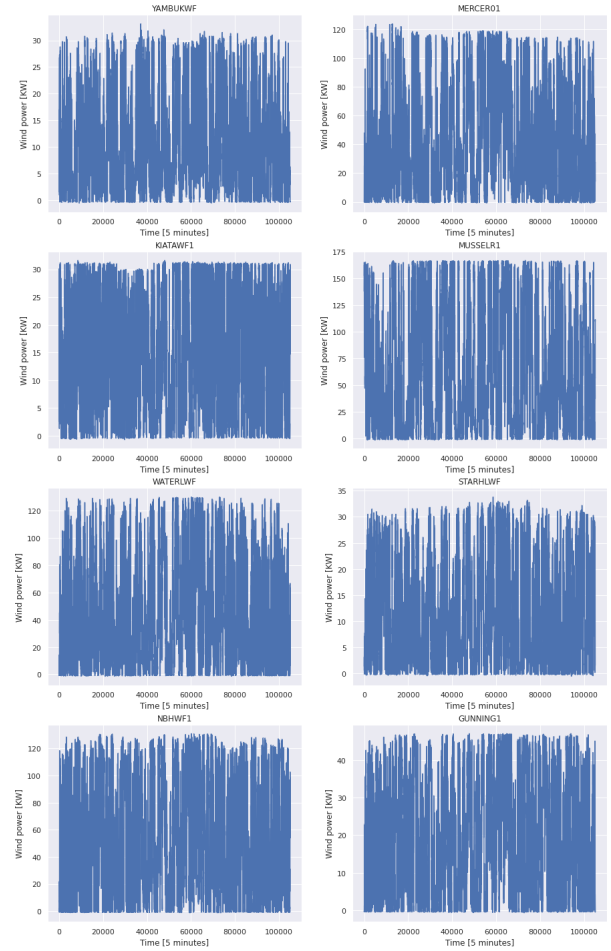


Fig. 5: Séries anuais para cada central eólica.

A análise da tendência, como mostra a figura 6, indica que o pico de produção de energia ocorre em meados do mês de agosto, que coincide com o período de maior intensidade

de velocidade do vento para a região, de acordo com dados oficiais disponibilizados pelo governo da Austrália [8]. Com excessão do comportamento médio (bias), que é influenciado pela potência nominal, as centrais possuem variações de tendência muito semelhantes. Quanto à sazonalidade, devido ao curto período avaliado, não foi possível observar comportamentos sazonais anuais e, portanto, apenas pôde-se verificar a sazonalidade diária.

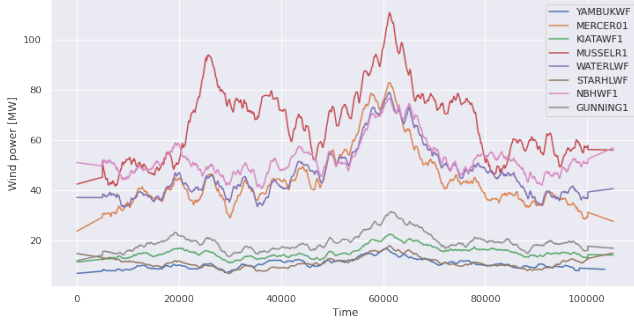


Fig. 6: Tendências das séries de potência.

Com relação ao histograma, todas as séries de potência se assemelham a figura 7, indicando um comportamento comum para função de distribuição de probabilidade de produção de energia eólica [9].

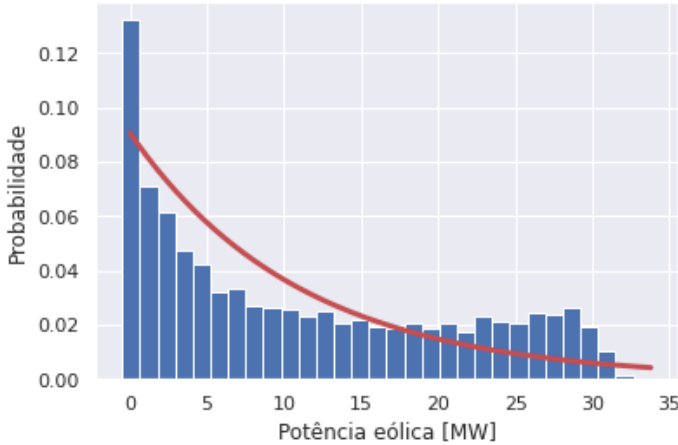


Fig. 7: Distribuição de probabilidade para a potência eólica para a central STARHLWF.

B. Outros tipos de séries temporais

Escolhemos alguns conjuntos de dados pequenos e tradicionais usados na literatura para analisarmos a performance do nosso modelo em relação a outros tipos de séries temporais². As séries estão descritas na Tabela I e ordenadas de acordo com a quantidade de observações.

²Disponíveis em https://github.com/Eraysongaldino/dataset_time_series

V. EXPERIMENTOS

Os experimentos foram realizados sobre as bases de dados descritas na Seção IV. Dividimos em dois experimentos para as séries de potência: usando o conjunto inteiro de dados e usando apenas os meses de Novembro e Dezembro de 2018 de cada central. Para as séries temporais tradicionais avaliamos apenas para o conjunto inteiro de cada uma e os resultados foram comparados com os dados normalizados de 0 a 1. Nos experimentos utilizando o conjunto inteiro, os dados foram divididos em 66,67% para treino e 33,33% para teste. Para o outro experimento, utilizamos o mês de Novembro como conjunto de treino e Dezembro para teste.

Para cada experimento utilizamos quatro modelos:

- MLP-NDEC: MLP simples com apenas uma camada escondida de 256 unidades
- SAETS-NDEC: Nosso método sem utilizar como entrada as componentes principais do modelo de [1].
- MLP: Modelo proposto por [1].
- SAETS: Nosso método utilizando como entrada as componentes principais do modelo de [1].

A. Métricas

Com objetivo de avaliar diferentes perspectivas do desempenho dos modelos, foram utilizadas diferentes métricas, i.e., mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE). As equações são mostradas a seguir:

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

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

Em que n é o tamanho da amostra, y_i é o valor observado no instante i e \hat{y}_i é a estimativa correspondente ao instante i .

B. Diagrama de Taylor

O diagrama de Taylor é uma ferramenta objetiva para visualização do desempenho de um modelo ou conjunto de modelos, em relação a uma observação, no qual é possível compará-los de um modo simples. O diagrama é capaz de representar quatro estatísticos em um único gráfico: correlação como coordenada radial, o desvio padrão e a raiz quadrática do erro quadrático médio (RMSD) como coordenadas azimutais e o skill score [10].

TABLE I: Descrição dos dados de outros tipos de séries temporais.

Data set	Tamanho	Descrição	Comportamento
Nordic	3121	Consumo de energia por hora em Titak nos países nórdicos	Não-estacionária e sazonalidade
Traffic	1657	Tráfego agregado por hora (em bits) no backbone da rede acadêmica do Reino Unido	Não-estacionária e sazonalidade
Lake	600	Níveis mensais do Lago Erie	Não-estacionária e sazonalidade
Electricity	486	Eletricidade da Austrália	Não-estacionária, tendência, sazonalidade

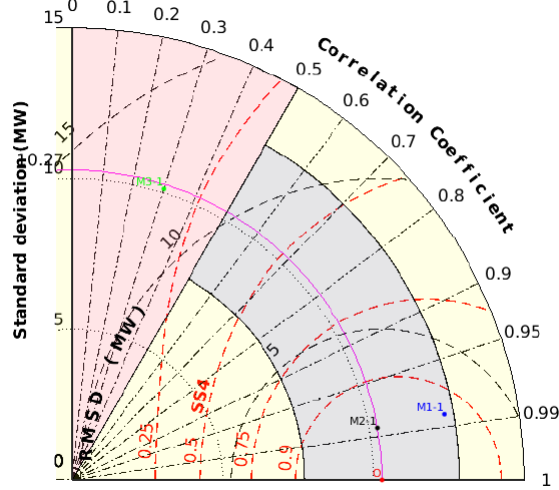


Fig. 8: Exemplo de diagrama de Taylor.

VI. RESULTADOS E DISCUSSÃO

A. Centrais eólicas

1) **Conjunto de dados completo:** Os resultados dos experimentos estão representados nas Tabelas II, III e IV e nas Figuras 9 e 10. Os diagramas de Taylor traduzem bem que a ocorre uma dispersão das previsões nos 3 maiores horizontes para o modelo que utiliza a decomposição da série, algo que não ocorre no modelo base. As tabelas também mostram o mesmo comportamento porém é visto que nosso modelo não obtém resultados muito distintos dos resultados do modelo base. Os modelos que não usaram a decomposição se mantiveram muito similares porém o SAETS conseguiu um desvio padrão mais próximo que as observações originais.

2) **Novembro-Dezembro:** Os resultados dos experimentos estão representados nas Tabelas V, VI e VII e nas Figuras 11 e 12. Na Figura 11 é perceptível a melhora do SAETS sobre o modelo base, o qual perde a performance em relação ao modelo da Figura 9 enquanto o nosso modelo mantém o mesmo padrão. As tabelas tornam mais perceptível ainda a performance do SAETS sobre todos as centrais analisadas. Os modelos que não utilizaram o método de Cabral mantiveram o mesmo padrão do primeiro experimento.

B. Séries tradicionais

Os resultados dos experimentos estão representados nas Tabelas VIII, IX e X e nas Figuras 13 e 14. Como mostra a Figura 13 e as tabelas, o SAETS se mostrou superior para menores horizontes e se dispersa muito nos maiores

horizontes. O modelo base não conseguiu bons resultados por conta de *overfitting* e suas componentes serem muito próximas de zero. As componentes quase zeradas são decorrência do método de decomposição precisar de uma grande quantidade de observações para ter um impacto melhor na série e esta é a principal motivação do nosso trabalho. Para os modelos que não utilizaram o método de Cabral, apesar da grande dispersão dos horizontes de previsão, manteve-se o mesmo padrão de previsão, um comparado ao outro.

VII. CONCLUSÃO

Este trabalho introduz o modelo SAETS baseado em mecanismos de self-attention e transfer learning para previsão de múltiplos horizontes de séries temporais de potência de centrais eólicas de poucas instâncias. Nós comprovamos experimentalmente que nosso método reproduz resultados comparáveis ou melhores ao modelo adotado como base em séries de potência longas e mantém o desempenho para séries 80% menores. Também mostramos que o modelo consegue se adaptar para outros tipos de séries com 90% a menos (ou menos ainda) observações que as séries de potência analisadas.

Apesar do ótimo desempenho nas séries temporais analisadas neste, comparamos apenas com métodos básicos e tomamos como base o trabalho de Cabral [1]. Para trabalhos futuros pretendemos comparar com métodos mais usuais na literatura assim como métodos que utilizam do mesmo mecanismo de self-attention que o SAETS.

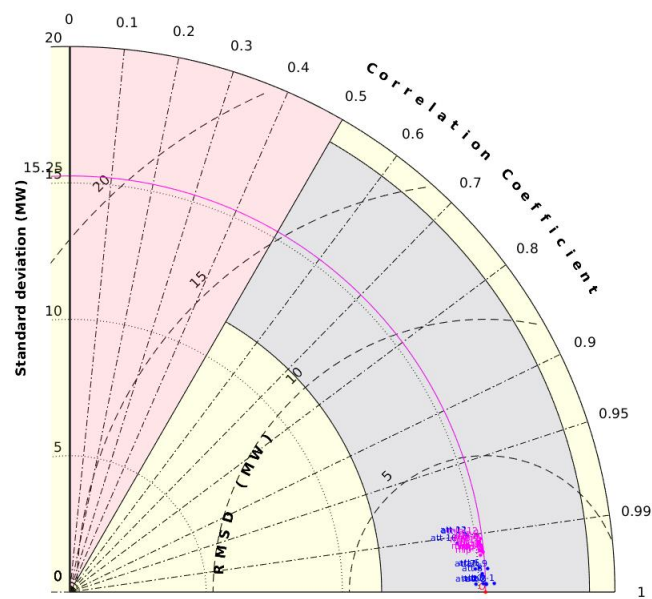


Fig. 9: Resultados para modelos com decomposição da série. Diagrama de Taylor para a central GUNNING1. O modelo MLP na cor rosa e o SAETS na cor azul.

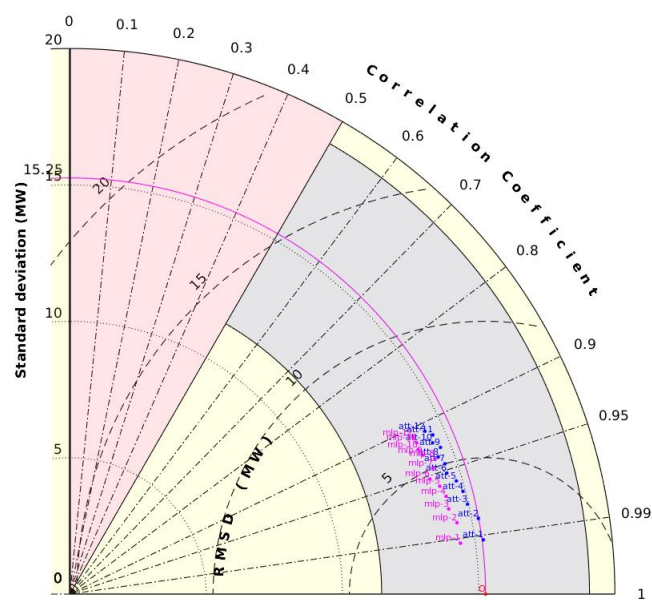


Fig. 10: Resultados para modelos sem decomposição da série. Diagrama de Taylor para a central GUNNING1. O modelo MLP na cor rosa e o SAETS na cor azul.

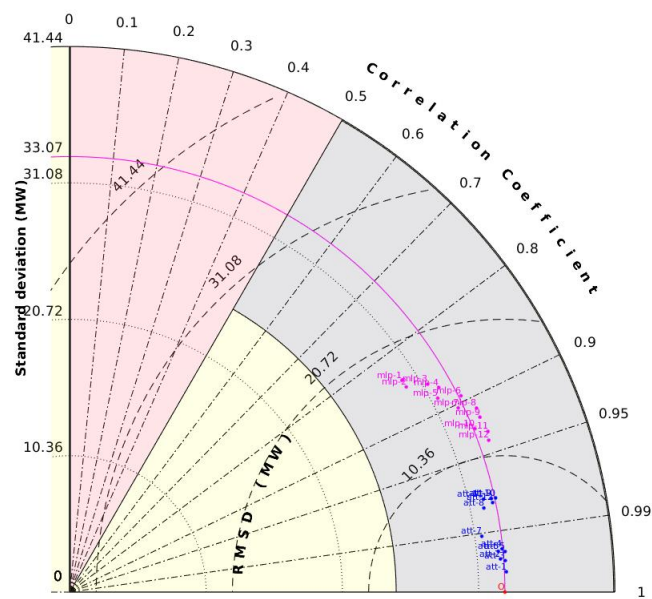


Fig. 11: Resultados para modelos com decomposição da série. Diagrama de Taylor para a central WATERLWF usando apenas os meses de Novembro e Dezembro como conjunto de dados. O modelo MLP na cor rosa e o SAETS na cor azul.

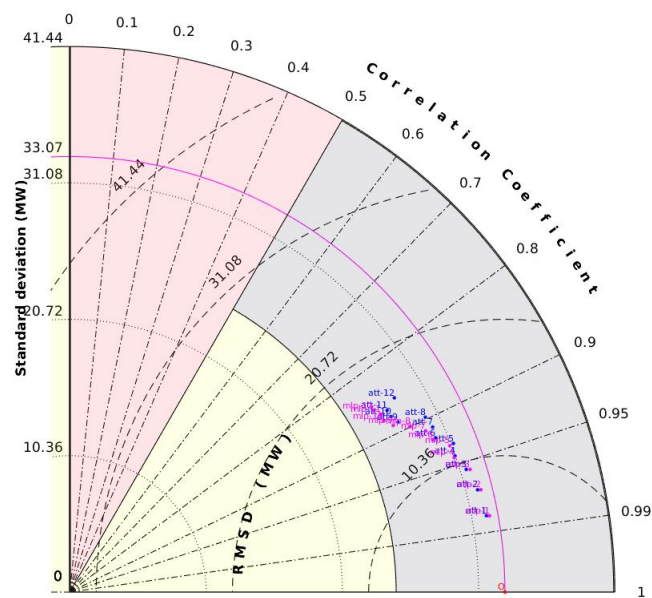


Fig. 12: Resultados para modelos sem decomposição da série. Diagrama de Taylor para a central WATERLWF usando apenas os meses de Novembro e Dezembro como conjunto de dados. O modelo MLP na cor rosa e o SAETS na cor azul.

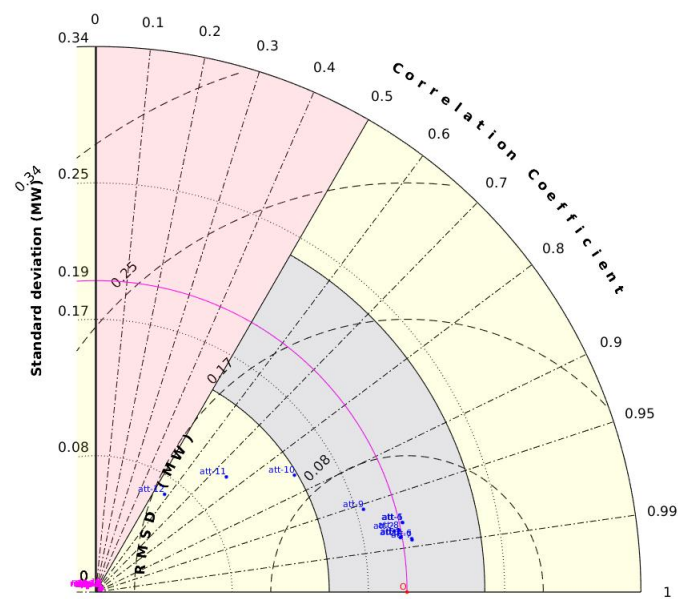


Fig. 13: Resultados para modelos com decomposição da série. Diagrama de Taylor para os dados de Nordic. O modelo MLP na cor rosa e o SAETS na cor azul.

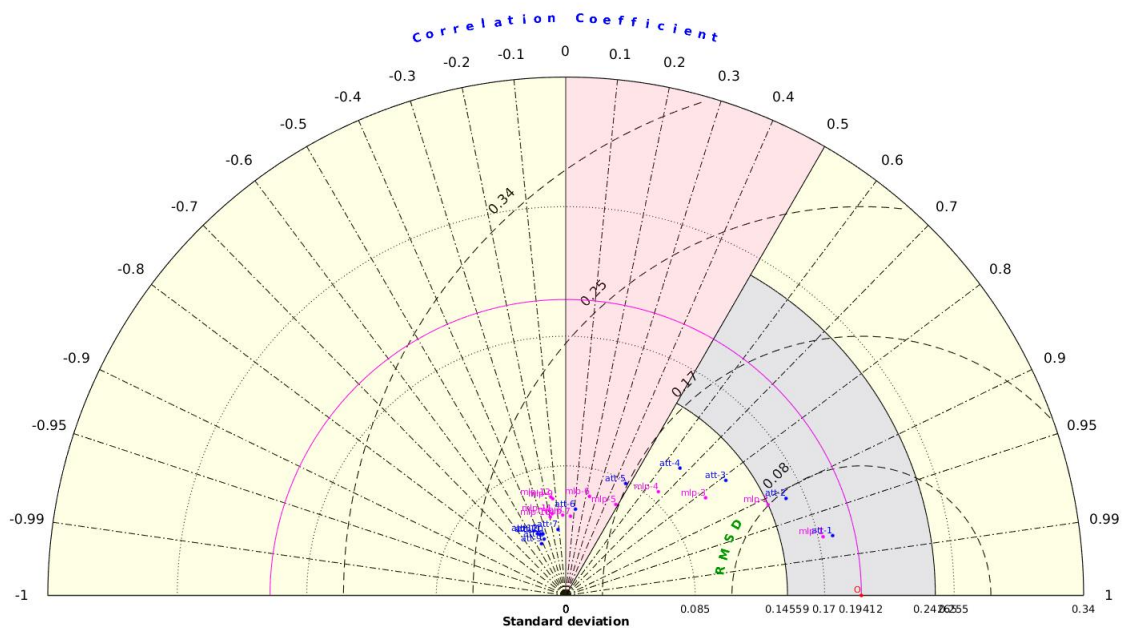


Fig. 14: Resultados para modelos sem decomposição da série. Diagrama de Taylor para os dados de Nordic. O modelo MLP na cor rosa e o SAETS na cor azul.

TABLE II: MAE para os dados de teste do dataset completo das centrais eólicas. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
GUNNING1	1.63	1.46	1.24	0.4	3.24	3.37	1.2	0.72	4.36	4.34	1.32	1.51
KIATAWF1	1.58	1.23	1.44	0.26	2.92	2.89	1.34	2.09	3.74	3.73	1.35	1.76
MERCER01	3.17	2.91	3.28	0.88	7.6	7.87	2.84	1.17	10.39	10.55	4.16	10.59
MUSSELR1	3.33	2.99	6.8	0.74	8.86	12.82	2.52	2.44	13.02	12.93	5.08	3.67
NBHWF1	2.93	3.05	5.01	2.93	8.46	9.36	3.01	1.33	12.23	12.15	3.16	2.16
STARHLWF	0.93	1.03	0.58	0.37	2.06	2.1	0.64	0.84	2.84	3.04	0.69	1.86
WATERLWF	4.11	3.04	2.74	0.66	7.93	8.36	2.03	1.45	11.32	11.33	3.66	7.3
YAMBUKWF	0.95	0.97	0.22	0.33	2.31	2.61	0.51	1.07	3.05	3.0	0.69	1.48

TABLE III: MSE para os dados de teste do dataset completo das centrais eólicas. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
GUNNING1	5.55	4.5	3.89	0.23	23.11	23.32	3.52	1.06	40.25	40.71	4.29	5.2
KIATAWF1	5.74	4.69	4.29	0.13	18.85	18.29	3.88	7.76	29.88	29.42	4.37	6.08
MERCER01	26.72	23.18	28.42	1.47	154.39	161.72	22.72	3.06	247.82	260.01	47.89	201.69
MUSSELR1	26.93	22.33	131.18	1.05	204.79	283.21	19.72	7.76	409.01	394.48	71.48	38.57
NBHWF1	21.91	22.47	53.71	9.83	163.78	186.33	22.95	4.76	316.84	327.95	26.55	10.17
STARHLWF	2.12	2.2	0.78	0.16	10.0	10.15	0.87	0.94	17.42	18.25	1.15	6.77
WATERLWF	40.1	30.3	20.61	0.82	153.78	159.88	13.34	3.79	280.86	292.36	29.48	136.18
YAMBUKWF	2.53	2.51	0.14	0.14	13.12	14.35	0.71	2.31	20.37	20.64	1.28	3.6

TABLE IV: RMSE para os dados de teste do dataset completo das centrais eólicas. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1				Horizonte 6				Horizonte 12			
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
GUNNING1	5.55	4.5	3.89	0.23	23.11	23.32	3.52	1.06	40.25	40.71	4.29	5.2
KIATAWF1	5.74	4.69	4.29	0.13	18.85	18.29	3.88	7.76	29.88	29.42	4.37	6.08
MERCER01	26.72	23.18	28.42	1.47	154.39	161.72	22.72	3.06	247.82	260.01	47.89	201.69
MUSSELR1	26.93	22.33	131.18	1.05	204.79	283.21	19.72	7.76	409.01	394.48	71.48	38.57
NBHWF1	21.91	22.47	53.71	9.83	163.78	186.33	22.95	4.76	316.84	327.95	26.55	10.17
STARHLWF	2.12	2.2	0.78	0.16	10.0	10.15	0.87	0.94	17.42	18.25	1.15	6.77
WATERLWF	40.1	30.3	20.61	0.82	153.78	159.88	13.34	3.79	280.86	292.36	29.48	136.18
YAMBUKWF	2.53	2.51	0.14	0.14	13.12	14.35	0.71	2.31	20.37	20.64	1.28	3.6

TABLE V: MAE dos modelos treinados no mês de novembro e prevendo o mês de dezembro de 2018 das centrais eólicas. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
GUNNING1	1.29	1.42	2.97	1.02	3.16	3.76	1.67	0.79	4.55	4.72	2.08	2.61
KIATAWF1	1.37	1.35	1.67	0.55	2.85	2.8	1.13	0.8	3.91	3.85	1.24	1.83
MERCER01	4.31	3.41	6.12	1.27	7.96	10.29	3.54	1.57	10.65	10.27	4.51	3.5
MUSSELR1	2.97	3.58	7.73	1.77	7.73	8.88	5.14	1.69	11.57	11.96	5.61	4.05
NBHWF1	2.97	3.34	35.09	1.33	9.34	10.1	35.09	2.04	14.31	14.14	35.06	4.74
STARHLWF	1.01	1.15	9.07	0.41	2.29	2.32	9.05	0.51	3.22	3.26	9.04	1.23
WATERLWF	5.65	3.75	6.36	1.61	8.99	8.57	3.27	1.86	11.91	12.39	4.53	3.97
YAMBUKWF	0.84	0.9	7.71	0.34	1.94	1.96	7.62	0.55	2.65	2.7	7.52	1.18

TABLE VI: MSE dos modelos treinados no mês de novembro e prevendo o mês de dezembro de 2018 das centrais eólicas. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
GUNNING1	4.17	4.62	20.02	2.06	24.5	29.21	6.0	1.42	44.64	47.87	10.16	17.58
KIATAWF1	4.06	3.94	6.32	0.44	17.42	17.56	3.14	1.14	29.73	30.22	3.62	7.13
MERCER01	34.27	26.8	97.32	2.87	153.17	181.98	33.03	6.57	246.65	255.97	57.81	24.5
MUSSELR1	17.1	24.86	131.7	6.07	140.76	167.08	55.48	6.35	286.48	299.59	66.46	35.64
NBHWF1	25.5	27.86	1620.17	3.64	194.6	215.95	1608.04	8.5	381.01	380.7	1602.34	51.52
STARHLWF	2.27	2.67	109.58	0.33	11.67	11.75	109.34	0.52	21.16	21.49	108.95	3.39
WATERLWF	54.58	38.8	94.77	5.49	170.11	170.59	30.29	9.98	276.1	316.15	54.17	44.06
YAMBUKWF	2.01	2.04	78.96	0.25	10.1	10.38	78.07	0.59	17.37	17.8	77.18	3.41

TABLE VII: RMSE dos modelos treinados no mês de novembro e prevendo o mês de dezembro de 2018 das centrais eólicas. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
GUNNING1	4.17	4.62	20.02	2.06	24.5	29.21	6.0	1.42	44.64	47.87	10.16	17.58
KIATAWF1	4.06	3.94	6.32	0.44	17.42	17.56	3.14	1.14	29.73	30.22	3.62	7.13
MERCER01	34.27	26.8	97.32	2.87	153.17	181.98	33.03	6.57	246.65	255.97	57.81	24.5
MUSSELR1	17.1	24.86	131.7	6.07	140.76	167.08	55.48	6.35	286.48	299.59	66.46	35.64
NBHWF1	25.5	27.86	1620.17	3.64	194.6	215.95	1608.04	8.5	381.01	380.7	1602.34	51.52
STARHLWF	2.27	2.67	109.58	0.33	11.67	11.75	109.34	0.52	21.16	21.49	108.95	3.39
WATERLWF	54.58	38.8	94.77	5.49	170.11	170.59	30.29	9.98	276.1	316.15	54.17	44.06
YAMBUKWF	2.01	2.04	78.96	0.25	10.1	10.38	78.07	0.59	17.37	17.8	77.18	3.41

TABLE VIII: MAE para os dados de teste de outros tipos de séries temporais. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
nordic	0.0335	0.0331	0.156	0.0255	0.1516	0.16	0.1565	0.0272	0.1738	0.1779	0.1576	0.1329
traffic	0.0914	0.0949	0.2216	0.0363	0.2098	0.2085	0.222	0.0768	0.2372	0.2595	0.2236	0.1682
lake	0.1189	0.1169	0.1808	0.0849	0.1744	0.1606	0.1731	0.3162	0.1271	0.1408	0.1772	0.1113
electricity	0.2171	0.2154	0.2026	0.1015	0.2146	0.3394	0.194	0.2063	0.1473	0.3434	0.1901	0.258

TABLE IX: MSE para os dados de teste de outros tipos de séries temporais. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
nordic	0.002	0.002	0.0368	0.0012	0.036	0.039	0.0371	0.0011	0.0459	0.0475	0.0374	0.027
traffic	0.0136	0.0189	0.0852	0.0019	0.0806	0.0836	0.0859	0.0083	0.0833	0.1069	0.0866	0.0514
lake	0.0222	0.02	0.0526	0.0107	0.0474	0.0391	0.0442	0.1203	0.0235	0.0292	0.0499	0.0186
electricity	0.0795	0.0658	0.0725	0.0137	0.0679	0.1724	0.0663	0.0583	0.0332	0.1786	0.0588	0.0929

TABLE X: RMSE para os dados de teste de outros tipos de séries temporais. As colunas com "NDEC" são os resultados sem utilizar a decomposição de séries temporais [1].

	Horizonte 1			Horizonte 6			Horizonte 12					
	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS	MLP-NDEC	SAETS-NDEC	MLP	SAETS
nordic	0.002	0.002	0.0368	0.0012	0.036	0.039	0.0371	0.0011	0.0459	0.0475	0.0374	0.027
traffic	0.0136	0.0189	0.0852	0.0019	0.0806	0.0836	0.0859	0.0083	0.0833	0.1069	0.0866	0.0514
lake	0.0222	0.02	0.0526	0.0107	0.0474	0.0391	0.0442	0.1203	0.0235	0.0292	0.0499	0.0186
electricity	0.0795	0.0658	0.0725	0.0137	0.0679	0.1724	0.0663	0.0583	0.0332	0.786	0.0588	0.0929

REFERENCES

- [1] L. Cabral, “Previsão de curtíssimo prazo da potência de saída de centrais eólicas baseada em análise espectral e decomposição da série temporal,” *Universidade Federal de Pernambuco*, 08 2018.
- [2] W. W. E. Association, “World wind capacity at 650,8 gw, corona crisis will slow down markets in 2020, renewables to be core of economic stimulus programmes,” 2020, accessed: 2021-03-20. [Online]. Available: <https://wwindea.org/world-wind-capacity-at-650-gw/>
- [3] A. B. de Energia Eólica, “Infovento 19,” 2021, accessed: 2021-03-20. [Online]. Available: http://abeeolica.org.br/wp-content/uploads/2021/02/2021_02_18_InfoVento19.pdf
- [4] A. Costa, “Mathematical/statistical and physical/meteorological models for short-term prediction of wind farms output,” *Universidad Politécnica de Madrid*, 2005.
- [5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” 2017.
- [6] N. Wu, B. Green, X. Ben, and S. O’Banion, “Deep transformer models for time series forecasting: The influenza prevalence case,” 2020.
- [7] H. Outhred, “Wind energy and the national electricity market with particular reference to south australia,” *a report prepared for the Australian Greenhouse Office*, 2003.
- [8] A. Government, “Climate statistics for australian locations,” Bureau of Meteorology, accessed: 2021-03-20. [Online]. Available: http://www.bom.gov.au/jsp/ncc/cdio/cvg/av?p_stn_num=024511&p_prim_element_index=46&p_display_type=statGraph&period_of_avg=ALL&normals_years=allYearOfData&staticPage=
- [9] A. Tsikalakis and N. Hatziaargyriou, “Financial evaluation of renewable energy source production in microgrids markets using probabilistic analysis,” 07 2005, pp. 1 – 7.
- [10] K. E. Taylor, “Summarizing multiple aspects of model performance in a single diagram,” *JOURNAL OF GEOPHYSICAL RESEARCH*, pp. 7183–7192, 04 2001.