

Redução das incertezas na projeção de longo prazo da temperatura máxima na Neoenergia Pernambuco

Tema: Planejamento da Expansão

Autores: Luiz Fernando dos Santos, João Augusto Hackerott, Dirceu Luis Herdies, Helber Barros Gomes, Robinson Semolini, Bruno Dantas Cerqueira, Djanilton Henrique Moura Junior

Co-Autores: -

Empresa: Elektro Redes S/A - NEOENERGIA Elektro

Resumo

Este estudo buscou aprimorar a projeção da temperatura máxima do ar na concessão da Neoenergia Pernambuco, reduzindo incertezas em projeções climáticas para até 10 anos e otimizando o planejamento de demanda energética. As métricas de erro (MAPE, MAE e MSE) foram calculadas para os resultados puramente de modelos do CMIP6 e para os seis métodos propostos. Os modelos selecionados puramente do CMIP6 apresentaram MAPE em torno de 9,1%, MAE de até 2,8°C e MSE acima de 9°C². Por outro lado, o método V, que combina aprendizado de máquina com dados observacionais do SAMet e MERGE, histórico e projeção do CMIP6 e previsões do CFSv2, destacou-se por apresentar os menores erros, indicando valores de MAPE, MAE e MSE de aproximadamente 1,6%; 0,5°C e 0,4°C², respectivamente. Esses resultados evidenciam o potencial do aprendizado de máquina para aumentar a precisão das projeções climáticas e energéticas, consolidando sua relevância como ferramenta para o planejamento eficiente e estratégico do setor elétrico

1. Introdução

Ao longo do último século, a expansão econômica do Brasil devido à industrialização, à evolução tecnológica, ao aumento na densidade demográfica de grandes centros, além de evidências sobre mudanças climáticas oriundas da ação antropogênica, desencadeou um aumento considerável no consumo de energia primária. No último relatório do Painel Intergovernamental de Mudanças Climáticas (IPCC, 2021) foram apresentadas evidências quanto ao aumento na frequência de eventos climáticos intensos em diversas partes do globo, incluindo a América do Sul, com ocorrência de extremos de precipitação e temperatura. A temperatura do ar, em particular, é uma das principais variáveis preditoras na estimativa da demanda de carga energética devido à sua influência direta no consumo e na distribuição de energia. Por exemplo, em janeiro de 2015, a demanda de energia elétrica ficou bem acima do programado e isso culminou na interrupção na distribuição de energia em dez estados e no Distrito Federal (ONS, 2015). Um dos fatores responsáveis pela grande demanda de carga naquela época foi atribuído ao uso exagerado do ar-condicionado (ALMEIDA, 2019). Ademais, estudos mostram que pessoas que se expõem em demasia ao ar-condicionado tendem a desenvolver menor tolerância às temperaturas mais altas (DE VECCHI et al., 2012).

No que concerne às concessionárias de energia elétrica do Brasil, a Neoenergia distribui energia elétrica para mais de 37 milhões de pessoas nos estados da Bahia, Pernambuco, Rio Grande do Norte, São Paulo, Mato Grosso do Sul e Distrito Federal. Alinhadas com o compromisso de gerar valor para os clientes, as distribuidoras da Neoenergia têm investido no aperfeiçoamento das áreas de atendimento e introduzido novas ferramentas que ofereçam maior comodidade e eficiência na prestação de serviços. Para isso, tem investido também em estudos para o aperfeiçoamento do planejamento energético para que seja garantido o fornecimento de energia elétrica aos consumidores a qualquer condição de demanda.

No que diz respeito à previsão de variáveis climáticas, os modelos numéricos de tempo e clima são as principais ferramentas para simular variáveis meteorológicas, como a temperatura do ar, variável climática indispensável nos modelos de projeção de carga. No entanto, os modelos numéricos podem apresentar alguns desvios em seus resultados dadas às aproximações matemáticas, parametrizações de processos físicos e incertezas nas condições iniciais e de contorno, o que pode resultar em erros sistemáticos.

Nesse cenário, a combinação de modelos numéricos com técnicas de aprendizado de máquina (*machine learning*) tem emergido como uma abordagem promissora para lidar com essas limitações. Enquanto os modelos numéricos fornecem informações baseadas em processos físicos bem estabelecidos, os algoritmos de *machine learning* têm a capacidade de aprender padrões complexos e não lineares a partir de dados históricos, observacionais ou de modelos numéricos. Essa integração permite não apenas a redução das incertezas nas projeções de temperatura do ar, mas também uma melhoria significativa na previsão de variáveis críticas para o planejamento energético. Assim, a abordagem combinada oferece uma solução robusta para aprimorar a acurácia das previsões climáticas e, conseqüentemente, contribuir para um planejamento energético mais eficiente.

Sendo assim, a fim de aprimorar a projeção da temperatura máxima do ar na área de concessão da Neoenergia Pernambuco, que atende os 184 municípios de Pernambuco, o Arquipélago de Fernando de Noronha e a cidade de Pedras de Fogo, na Paraíba, o presente estudo tem como objetivo reduzir as incertezas no horizonte de previsão de até 10 anos e, desta maneira, contribuir para um planejamento mais eficiente da demanda de carga energética nessa concessão. Vale ressaltar que esses são resultados relacionados ao projeto de pesquisa e desenvolvimento (P&D) da Neoenergia, regulado pela Agência Nacional de Energia Elétrica (ANEEL), em parceria com a Tempo OK Tecnologia em Meteorologia Ltda e Universidade Federal de Alagoas (UFAL).

2. Desenvolvimento

2.1) Dados e Métodos

Os dados diários de temperatura máxima do *South American Mapping of Temperature* (SAMeT) (ROZANTE et al., 2022) e de precipitação do MERGE (ROZANTE et al., 2010) foram considerados como dados observacionais no período de 2000-2023. O SAMeT possui dados a cada 0,5° de latitude e longitude e o MERGE apresenta resolução de 0,1° de latitude por longitude, ambos sobre a América do Sul. Eles foram utilizados para o treinamento do algoritmo de *machine learning*, tal como para comparar os resultados das projeções. Para contribuir na projeção de temperatura no horizonte de até 10 anos, foi utilizada a temperatura máxima do ar, histórica de 2000 a 2014 e projeção após 2015, de modelos selecionados da fase 6 e mais recente do *Coupled Model Intercomparison Project* (CMIP6). O CMIP6 é um esforço global que padroniza experimentos com modelos climáticos para simular cenários passados, presentes e futuros (EYRING et al., 2016; FIRPO et al., 2022). Sua importância para projeções de temperatura de até 10 anos está na capacidade de fornecer dados consistentes e detalhados, essenciais para análises climáticas de médio prazo, permitindo

maior precisão no planejamento energético e em estudos de impactos climáticos. Aqui foram analisados dois cenários diferentes de projeção: SSP2-4.5 e SSP5-8.5, considerados cenário intermediário e pessimista, respectivamente.

Para previsão de temperatura máxima, precipitação e umidade específica foi utilizado o *Climate Forecast System version 2* (CFSv2) - modelo norte-americano operacional acoplado oceano-continente-atmosfera, integrado a um sistema de assimilação de dados (SAHA et al. 2014), apresenta simulações periódicas podendo alcançar um horizonte de previsão de até 9 meses. O período utilizado foi de 2012-2022, de acordo com a disponibilidade dos dados.

Por fim, o algoritmo de *machine learning* abordado foi o *Gradient Boosting Regressor* (GBR), o qual considera uma previsão por conjunto a partir de uma única árvore de decisão e as demais são geradas reduzindo os erros da árvore anterior (MALAKOUTI, 2023). Inclusive, a Tabela 1 mostra o resumo dos modelos e dados utilizados:

Tabela 1 - Resumo dos dados observados, previsão e projeção de temperatura máxima utilizadas nos métodos de projeção para a Neoenergia Pernambuco.

	CONJUNTO DE DADOS	PERÍODO	PARTICULARIDADES
OBSERVAÇÃO	South American Mapping of Temperature (SAMET)	2000-2022	dados diários
HINDCAST DO MODELO DE 9 MESES	Climate Forecast System (CFS)	1982-2010	dados mensais
PREVISÃO DO MODELO DE 9 MESES	Climate Forecast System (CFS)	2012-2022	dados diários; 5 previsões consecutivas e não-sobrepostas
HISTÓRICO DOS MODELOS DE LONGO PRAZO	Coupled Model Intercomparison Project Phase 6 (CMIP6)	2000-2014	Análise de componentes principais (PCA) de dados diários e com modelos selecionados
MACHINE LEARNING	Gradient Boosting Regressor (GBR)		

Foi calculada a média aritmética de cada variável climática na área que corresponde à Neoenergia Pernambuco e construída uma série temporal de cada conjunto de dados no período de 2000-2022. Além disso, para combinar os resultados dos modelos numéricos do CMIP6 e CFSv2 com o GBR, foram desenvolvidos e testados 6 métodos diferentes, mostrados a seguir:

Método I - Neste método são configurados três processos com *machine learning*, os quais serão identificados por IA1, IA2 e IA3.

İ Dados diários de temperatura (média e máxima) a partir do conjunto de dados observacionais escolhido, na região de concessão da Neoenergia Pernambuco no período de 2012-2019 (8 anos), considerados dados de referência para o treinamento da média diária de 5 previsões consecutivas e não-sobrepostas das respectivas temperaturas no horizonte de nove meses do CFSv2. Este processo resulta na IA1;

İ Dados diários de temperatura, histórico e projeção (cenário SSP2-4.5 e SSP5-8.5) dos modelos selecionados do CMIP6 são treinados com base na temperatura média ou máxima do dado observacional no período de 2000-2019 (20 anos), definindo assim a IA2;

İ Os dados do CMIP6 ajustados (SSP2-4.5 e SSP5-8.5 separadamente) pela IA2 são combinados pelo IA1 (CFSv2 ajustado) de 2012-2019 (8 anos), concluindo a IA3 para o período de 2000-2019;

İ Por fim, para testar o ajuste e a projeção, a IA3 ajusta a projeção do CMIP6 dos cenários mencionados de 2020-2022 (3 anos), período de teste da projeção.

Método II - Neste método apresenta somente um processo com *machine learning*.

İ Dados diários de temperatura máxima observada na região de concessão da Neoenergia Pernambuco no período de 2012-2019 (8 anos), considerados dados de referência para o treinamento da média diária de 5 previsões consecutivas e não-sobrepostas das respectivas temperaturas no horizonte de nove meses do CFSv2 concomitantemente aos dados diários de temperatura, histórico e projeção (cenário SSP2-4.5 e SSP5-8.5), dos modelos selecionados do CMIP6 a partir do GBR;

İ Por fim, para testar o ajuste e a projeção, o modelo de *machine learning* treinado no item anterior ajusta a projeção do CMIP6 dos cenários mencionados de 2020-2022 (2 anos), período de teste da projeção.

Método III - Neste método apresenta dois processos com *machine learning* similar ao Método 2, mas da seguinte maneira:

İ Temperatura máxima diária do conjunto de dados observacionais na região de concessão da Neoenergia Pernambuco no período de 2012-2019 (8 anos), considerados dados de referência para o treinamento de:

İ média diária de 5 previsões consecutivas e não-sobrepostas, a partir da quinta previsão individual, das respectivas temperaturas no horizonte de nove meses do CFSv2 concomitantemente aos dados diários de temperatura, histórico e projeção (cenário SSP2-4.5 e SSP5-8.5), dos modelos selecionados do CMIP6 a partir do *machine learning* (IA1);

İ Posteriormente, a média diária das 5 primeiras previsões no horizonte de nove meses do CFSv2 da temperatura foi utilizada para aprimorar o IA1, resultando no IA2.

İ Por fim, para testar o ajuste e a projeção, o modelo de *machine learning* treinado no item anterior ajustará a projeção do CMIP6 dos cenários mencionados de 2020-2022 (2 anos), período de teste da projeção.

Método IV - Neste método apresenta três processos com *machine learning*, sendo uma variação entre os métodos 2 e 3:

İ Dados diários de temperatura máxima observada na região de concessão da Neoenergia Pernambuco no período de 2012-2019 (8 anos), considerados dados de referência para o treinamento da média diária de 5 previsões consecutivas e não-sobrepostas das respectivas temperaturas no horizonte de nove meses do CFSv2 concomitantemente aos dados diários de temperatura, histórico e projeção (cenário SSP2-4.5 e SSP5-8.5), dos modelos selecionados do CMIP6 a partir do GBR (IA1);

İ Posteriormente, as 5 primeiras médias das 5 primeiras previsões no horizonte de nove meses do CFSv2 da precipitação, ajustada pelos dados do MERGE a partir do *machine learning* (IA2), foram utilizadas para aprimorar o IA1, resultando no IA3.

İ Por fim, para testar o ajuste e a projeção, o modelo de *machine learning* treinado no item anterior ajustará a projeção do CMIP6 dos cenários mencionados de 2020-2022 (2 anos), período de teste da projeção.

Método V - Neste método apresenta três processos com *machine learning* similar ao Método 4, mas para os seguintes dados:

İ Dados diários de temperatura máxima do conjunto observacional na região de concessão da Neoenergia Pernambuco no período de 2012-2019 (8 anos), considerados os dados de referência para o treinamento da média diária de 5 previsões consecutivas e não-sobrepostas das respectivas temperaturas no horizonte de nove meses do CFSv2 concomitantemente aos dados diários de temperatura, histórico e projeção (cenário SSP2-4.5 e SSP5-8.5), dos modelos selecionados do CMIP6 a partir do GBR (IA1);

İ 4 primeiras médias diárias das 5 primeiras previsões no horizonte de nove meses do CFSv2 da precipitação ajustada pelos dados do MERGE a partir do GBR (IA2);

İ Posteriormente, as 4 primeiras médias diárias das 5 primeiras previsões no horizonte de nove meses do CFS da umidade específica foram utilizadas para aprimorar o IA1 e IA2 concomitantemente, resultando no IA3.

Por fim, para testar o ajuste e a projeção, o modelo de *machine learning* treinado no item anterior ajustará a projeção do CMIP6 dos cenários mencionados de 2020-2022 (2 anos), período de teste da projeção.

Método VI - Neste método apresenta três processos com *machine learning* similar aos método 4 e 5, mas para os seguintes dados:

Dados diários de temperatura máxima do conjunto de dado observacional escolhido na seção 8.2 na região de concessão da Neoenergia Pernambuco no período de 2012-2019 (8 anos), considerados os dados de referência para o treinamento da média diária de 5 previsões consecutivas e não-sobrepostas das respectivas temperaturas no horizonte de nove meses do CFSv2 concomitantemente aos dados diários de temperatura, histórico e projeção (cenário SSP2-4.5 e SSP5-8.5), dos modelos selecionados do CMIP6 a partir do GBR (IA1);

5 primeiras previsões individuais no horizonte de nove meses do CFS da precipitação ajustada pelos dados do MERGE a partir do GBR (IA2);

Posteriormente, as 4 primeiras médias diárias das 5 primeiras previsões no horizonte de nove meses do CFSv2 da umidade específica foram utilizadas para aprimorar o IA1 e IA2 concomitantemente, resultando no IA3.

Por fim, para testar o ajuste e a projeção, o modelo de *machine learning* treinado no item anterior ajustará a projeção do CMIP6 dos cenários mencionados de 2020-2022 (2 anos).

A partir dos resultados de cada método foi mensurado erro percentual absoluto médio (MAPE, em %), erro médio absoluto (MAE, em °C) e erro quadrático médio (MSE, °C²) (LE e ZIDEK, 2006; WILKS, 2019) no período de teste (2020-2022) para a Neoenergia Pernambuco. Inclusive, a Equação 1, 2 e 3 denotam o MAPE, MAE e MSE, respectivamente:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - o_i}{o_i} \right| \quad (1)$$

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

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

sendo y o resultado da projeção de temperatura máxima de cada método, o a temperatura máxima observada, i cada dado individual, e n o número total de dados no período de 2020-2022.

A escolha dos melhores métodos estará atrelada àqueles que apresentarem os menores valores de MAPE, MAE e MSE no período de teste (2020-2022).

2.2) Avaliação dos métodos de projeção

Com base nos resultados dos métodos supracitados e para os cenários intermediário e pessimista, foi mensurado erro percentual absoluto médio (MAPE, em %), erro médio absoluto (MAE, em °C) e erro quadrático médio (MSE, °C²) no período de 2020-2022 para Neoenergia Pernambuco, como pode ser visto na Figura 1:

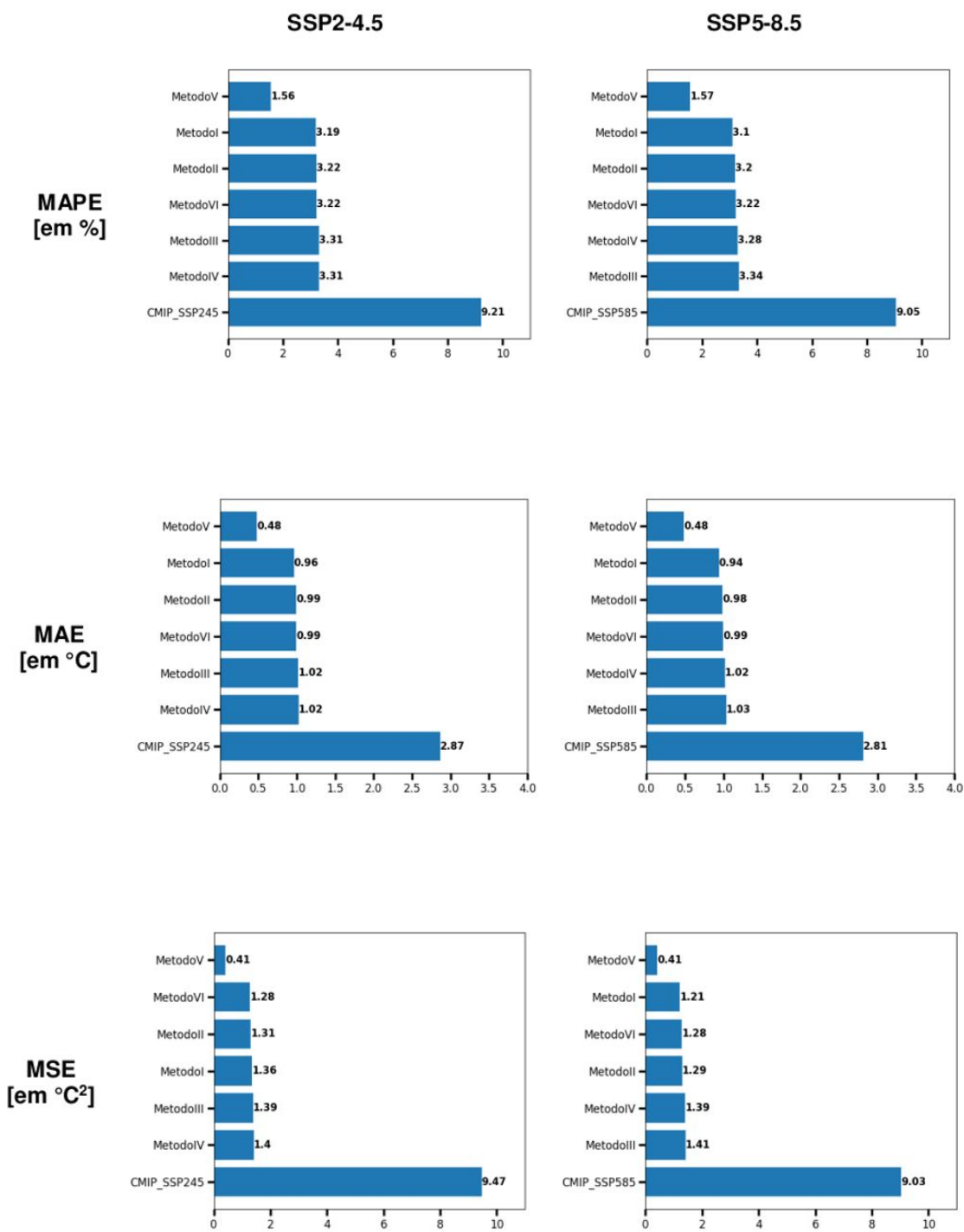


Figura 1 - Métricas estatísticas de erro (MAPE, MAE e MSE) da temperatura máxima para Pernambuco, a partir dos métodos propostos e nos cenários intermediário (SSP2-4.5) e pessimista (SSP5-8.5) do CMIP6.

Os resultados ajustados são significativamente superiores quando é aplicado o método V, onde as métricas estatística de erro MAPE, MAE e MSE (1,6%, 0,5 °C e 0,4 °C², respectivamente) são inferiores a 50% quando da aplicação dos outros cinco métodos que ficaram em torno de 3,2%, 1 °C e 1,3 °C², respectivamente. Quando esses resultados são comparados com as projeções puras do CMIP, essas reduções são ainda mais significativas, com valores inferiores a 30% dos erros atribuídos ao CMIP6, em ambos os cenários. A partir das métricas de erro, foram selecionados os dois melhores resultados para analisar a série temporal e comparar com a observação no período de teste (Figura 2).

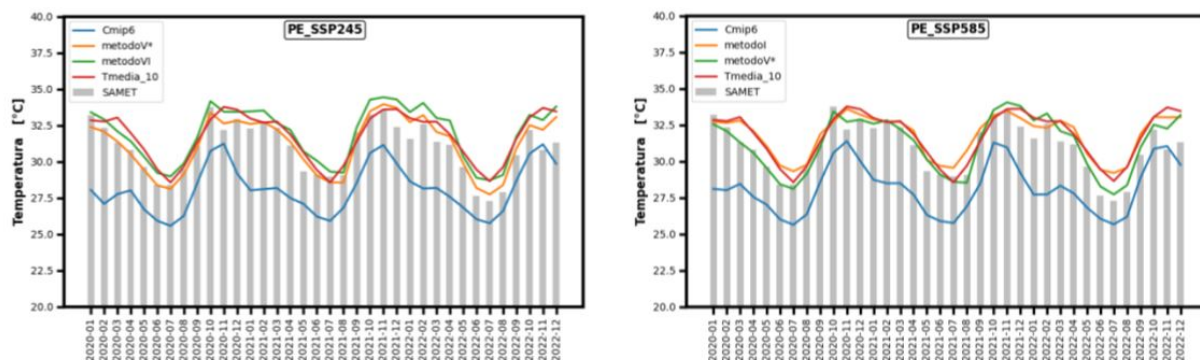


Figura 2 - Temperatura máxima para Neoenergia Pernambuco a partir dos métodos propostos e nos cenários intermediário (SSP2-4.5, à esquerda) e pessimista (SSP5-8.5, à direita) do CMIP6 e série temporal na escala mensal da temperatura máxima de 2020-2022. O asterisco (*) indica o método que apresentou os menores erros. A barra corresponde à temperatura máxima do SAMeT (dado observado), a linha azul da projeção pura do CMIP6 e a linha vermelha a média dos 2013-2022 (últimos 10 anos)

Os dois métodos que apresentaram os melhores resultados para o cenário intermediário foram os métodos V e VI e, para o cenário pessimista, os métodos V e I, respectivamente. Claramente, o ajuste das projeções utilizando aprendizado de máquina reduziu o viés do CMIP6. De modo geral, o método V foi aquele que, além de proporcionar os menores erros, representou melhor as variações da temperatura máxima para Neoenergia Pernambuco em toda a série temporal em ambos os cenários do CMIP6, superior também à média de 10 anos (método comumente utilizado pela Neoenergia para projeção a temperatura sem a utilização de modelos numéricos). Seu desempenho deve estar atrelado ao uso de temperatura, precipitação e umidade específica para treinar o modelo de *machine learning* para ajustar as projeções do CMIP6 combinadas às previsões do CFSv2.

3. Conclusão

Este estudo teve como principal objetivo aprimorar a projeção da temperatura máxima do ar para a concessão da Neoenergia Pernambuco para diminuir as incertezas dessa variável climática no horizonte de previsão de até 10 anos e reduzir os desvios no planejamento de mercado potencialmente causados na projeção de carga pela Metodologia Padrão - projeção do clima baseada na média de temperatura máxima observada nos últimos 10 anos. A combinação do estado da arte de modelos climáticos e o uso de *machine learning* alcançou uma redução de incertezas na projeção de temperatura, quando comparado aos resultados puros do CMIP6, com uma diminuição do MAPE de 9,21% para 1,56% a 3,19%, representando consequentemente uma redução do MAE de 2,8°C para 0,5°C, e do MSE de 9,5°C² para 0,4°C² na projeção da temperatura máxima considerando o método V.

As metodologias propostas e seus respectivos cenários foram capazes de alcançar resultados satisfatórios, reduzindo a incerteza nas projeções e até superando as projeções puramente oriundas de modelos numéricos de clima. Esses resultados também foram comparados à Metodologia Padrão e os resultados da combinação do estado da arte de modelos climáticos e técnicas de *machine learning* também foram superiores, confirmando que as projeções desenvolvidas neste projeto captaram de forma mais consistente os picos e quedas na série de temperatura estudada.

4. Referências bibliográficas

- ALMEIDA, P. M. **Influência da ventilação natural na sensação térmica do usuário em ambiente educacional**. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santos. Vitória, p.92, 2019.
- DE VECCHI, R.; CÂNDIDO, C.; LAMBERTS, R. Thermal history and its influence on occupants' thermal acceptability and cooling preferences in warm-humid climates: A new desire for comfort. In: **Proceedings of the 7th Windsor Conference: The Changing Context of Comfort in an Unpredictable World Cumberland Lodge, Windsor, UK**, p. 13, 2012.
- EYRING, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. **Geoscientific Model Development**, v. 9, n. 5, p. 1937-1958, 2016.
- FIRPO, M. A. F. et al. Assessment of CMIP6 models' performance in simulating present day climate in Brazil. **Frontiers in Climate**, p. 170, 2022.
- IPCC - INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE: Climate change 2021-the physical science basis. **Interaction**, v. 49, n. 4, p. 44-45, 2021.
- LE, N. D.; ZIDEK, J. V.; **Statistical analysis of environmental space-time processes**. New York: Springer, p.11, 2006.
- MALAKOUTI, S. M. Utilizing time series data from 1961 to 2019 recorded around the world and machine learning to create a Global Temperature Change Prediction Model. **Case Studies in Chemical and Environmental Engineering**, v. 7, p. 100312, 2023.
- ONS - OPERADOR NACIONAL DO SISTEMA ELÉTRICO: **Boletim diário - eventos e ocorrência**, dia 19/01/2015. Disponível em: <<http://sdro.ons.org.br/SDRO/DIARIO/index.htm>>. Acesso em: 02 fev. 2023.
- ROZANTE, J. R. et al. Combining TRMM and surface observations of precipitation: technique and validation over South America. **Weather and forecasting**, v. 25, n. 3, p. 885-894, 2010.
- ROZANTE, *et al.* A newly developed South American Mapping of Temperature with estimated lapse rate corrections. **International Journal of Climatology**, v. 42, n. 4, p. 2135-2152, 2022.
- SAHA, S. *et al.* The NCEP climate forecast system version 2. **Journal of climate**, v. 27, n. 6, p. 2185-2208, 2014.
- WILKS, D. S. Statistical methods in the atmospheric sciences. **Elsevier**, p. 840, 2019.