

CONSTRUÇÃO DE UM MODELO PREDITIVO DE LONGO PRAZO: PREVISÃO DA TEMPERATURA ATMOSFÉRICA MÁXIMA MÉDIA EM CAMPOS DOS GOYTACAZES UTILIZANDO REDES NEURAS ARTIFICIAIS

CONSTRUCTION OF A LONG-TERM PREDICTIVE MODEL: FORECASTING MEAN MAXIMUM ATMOSPHERIC TEMPERATURE IN CAMPOS DOS GOYTACAZES USING ARTIFICIAL NEURAL NETWORKS

André Luiz Xavier Guimarães Nasri – Pontifícia Universidade Católica de Minas Gerais
dedeluznasri@gmail.com
ORCID: <https://orcid.org/0000-0002-4913-7759>

Hans Schmidt Santos – Observatório Nacional
hans.schmidt@live.com
ORCID: <https://orcid.org/0009-0001-6509-9617>

Guy Globa Masset – Universidade do Estado do Rio de Janeiro
guyglobamasset@gmail.com
ORCID: <https://orcid.org/0009-0001-2252-7116>

Vitor Mena Pirola – Faculdade Católica Salesiana Maria Auxiliadora
vitormp.eng@gmail.com
ORCID: <https://orcid.org/0009-0002-6814-1830>

Rayssa Barcellos Paiva – Faculdade Católica Salesiana Maria Auxiliadora
rayssabpaiva@hotmail.com
ORCID: <https://orcid.org/0009-0002-0620-5587>

Resumo

A crescente complexidade dos fenômenos climáticos tem desafiado os métodos tradicionais de previsão meteorológica, exigindo a otimização das técnicas habituais e o desenvolvimento de abordagens que acompanhem o avanço tecnológico do Estado da Arte. Nesse contexto, advoga-se que as redes neurais artificiais (RNAs) têm um enorme potencial como ferramentas para previsão climática de longo prazo, pois podem ser aplicadas para capturar padrões complexos e não-lineares presentes nos dados meteorológicos. Portanto, este artigo propõe a construção de um modelo de predição climática capaz de realizar a previsão da temperatura

máxima média das próximas décadas. Para que o objetivo seja cumprido, discorre-se sobre o processo de pré-processamento que compreende o desenvolvimento de um modelo de regressão linear múltipla para imputação de dados ausentes, bem como análises de linearidade, normalidade, homocedasticidade e pontos discrepantes e de alavancagem, para aferição da qualidade do tratamento. Além disso, o trabalho descreve a arquitetura e a topologia da rede neural construída, assim como o cálculo das medidas de incerteza e das métricas de performance para apurar o nível de desempenho do modelo e garantir a qualidade da previsão realizada.

Palavras-Chave: Análise Preditiva de Temperatura Máxima; Aprendizado de Máquina; Estatística Inferencial; Redes Neurais Artificiais.

Abstract

The increasing complexity of climatic phenomena has challenged traditional methods of weather forecasting, needing the optimization of conventional techniques and the development of approaches that keep pace with the technological advancements in the State of The Art. In this context, it is argued that artificial neural networks (ANNs) hold enormous potential as tools for long-term climate prediction, as they can be applied to capture complex and non-linear patterns present in meteorological data. Therefore, this article proposes the construction of a climate prediction model capable of forecasting the average maximum temperature for the coming decades. To achieve this objective, the paper discusses the preprocessing process, which involves the development of a linear regression model for data imputation of missing values, as well as analyses of linearity, normality, homoscedasticity, and outliers and leverage points to assess the quality of the treatment. Additionally, the work describes the architecture and topology of the constructed neural network, along with the calculation of uncertainty measures and performance metrics to assess the level of model performance and ensure the quality of the forecast conducted.

Keywords: Predictive Analysis of Maximum Temperature; Machine Learning; Inferential Statistics; Artificial Neural Networks.

1. INTRODUÇÃO

Um dos objetivos da previsão do tempo é deprimir as consequências atreladas aos riscos do desconhecimento das condições atmosféricas. O entendimento e monitoramento das variações climáticas são essenciais para a compreensão dos impactos ambientais e socioeconômicos que moldam o espaço geográfico nas escalas micro, meso, regional, continental e planetária (Ynoue *et al.*, 2017). Dentre os diversos parâmetros climáticos, a temperatura máxima desempenha um papel crucial, exercendo influência não apenas sobre os ecossistemas naturais, mas também sobre a segurança pública, agricultura, consumo de energia, planejamento urbano e diversos setores econômicos, como o turismo. À medida que a preocupação global com as mudanças climáticas cresce e as atenções são voltadas para análises de pontos de não retorno, a necessidade de estudar e prever as tendências de temperatura máxima torna-se cada vez mais premente.

A fim de explanação, cita-se que as médias das temperaturas máximas diária e mensal não se limitam a refletir a dinâmica climática instantânea, pois também servem como indicadores sensíveis de mudanças de longo prazo. Para Saucier (1969), alterações nesse parâmetro podem resultar em eventos climáticos extremos, como ondas de calor, que têm implicações diretas na saúde humana, na segurança alimentar e na disponibilidade de recursos hídricos. Além disso, ainda conforme o autor citado, o aumento persistente da temperatura máxima pode desencadear impactos significativos nos ecossistemas, levando a alterações na distribuição de espécies animais, na vegetação e nos regimes de precipitação.

Conforme dados cedidos pelo IBGE em 2022, o município fluminense de Campos dos Goytacazes é, a nível estadual, o maior produtor de cana-de-açúcar, o segundo maior produtor de soja e está entre os cinco maiores produtores de batata doce, frutas e mandioca, além de ser o maior território do estado do Rio de Janeiro. Dessa forma, afirma-se que a cidade desempenha um papel crucial no cenário de agricultura do estado do Rio de Janeiro e se mostra relevantemente sensível a alterações no índice de temperatura máxima. Por conta disso, elencou-se Campos dos Goytacazes como localidade alvo do estudo aqui discorrido, sendo palco da modelagem matemática do modelo preditivo desenvolvido, uma vez que o

desconhecimento de prováveis mudanças na temperatura máxima da região poderia impactar no desempenho do principal comércio local.

Frente a esse cenário de possibilidades plurais, afirma-se que as redes neurais artificiais (RNAs) se mostram como uma das principais ferramentas para o desenvolvimento de modelos preditivos de longo prazo (Haykin, 2000). Entende-se a RNA como uma classe de algoritmos de aprendizado de máquina capaz de se adaptar dinamicamente às mudanças nas condições meteorológicas, possibilitando previsões de amplo período, haja vista que as condições podem evoluir de maneira significativa ao longo do tempo, tornando-a, portanto, uma estratégia metodológica segura para estudos climáticos.

Frisa-se que a compreensão dessas tendências climáticas locais é fundamental para o desenvolvimento de estratégias de adaptação e mitigação de danos, permitindo uma resposta mais eficaz às mudanças ambientais. Explica-se que este artigo busca relatar o processo de desenvolvimento dos modelos de pré-processamento dos dados coletados e de previsibilidade climática da variável de temperatura máxima, sendo este último responsável pelos resultados gerados nesse trabalho. Para alcançar esse objetivo, este estudo adota uma abordagem de inteligência artificial, utilizando redes neurais artificiais para modelar e extrapolar a temperatura máxima em Campos dos Goytacazes até dezembro de 2049. Ao fazê-lo, espera-se contribuir não apenas para a ciência climática, mas também para a tomada de decisões, de forma embasada e sustentável, em face dos desafios climáticos que se mostram cada vez mais próximos.

2. METODOLOGIA

Como pré-requisito para a elaboração deste estudo, coletou-se os dados meteorológicos do município de Campos dos Goytacazes através do Banco de Dados Meteorológico do Instituto Nacional de Meteorologia (BDMEP/INMET), selecionando especificamente o repositório da estação meteorológica convencional que possui código 83698 e coordenadas latitude -21,742500 e longitude -41,332778, e cujas leituras dispõem de observações de médias mensais começando em janeiro de 1960 e concluindo em agosto de 2021. No entanto,

embora as informações do INMET sejam confiáveis, os registros do instrumento apresentam alguns dados ausentes, mostrando-se necessária a adoção de um método de imputação como estratégia metodológica adicional.

Então, como forma de melhor expor as próximas etapas realizadas na elaboração desse trabalho, optou-se por ordenar separadamente os dois modelos desenvolvidos e as suas respectivas técnicas de aferição de qualidade, dividindo-os em: 1 – modelo e técnicas de tratamento do banco de dados e 2 – modelo e métricas de performance e medidas de incerteza de previsão climática.

Dessa forma, o primeiro modelo desenvolvido é uma regressão linear múltipla entre os dados de temperatura média compensada, temperatura máxima média e temperatura mínima média. A regressão desenvolvida admite a temperatura média compensada como variável dependente e, no estudo proposto, isolou-se a variável da temperatura máxima para a obtenção dos dados ausentes e preenchimento dos valores faltantes no repositório.

Buscando assegurar a qualidade da interpolação realizada, submeteu-se os resíduos do modelo de regressão as técnicas de análise gráfica de linearidade, normalidade, homocedasticidade e pontos discrepantes e de alavancagem, processos que serão precisamente percorridos na seção de Desenvolvimento. Junto a essa imputação, separou-se toda a série de dados em dois intervalos, de treino e de teste, para que possam servir as duas demandas do modelo preditivo, o treinamento e a avaliação do aprendizado. Cita-se que após a realização do pré-processamento de dados e a aferição da qualidade desses tratamentos, a construção do modelo preditivo pôde ser iniciada.

Em anteposição a menção do modelo preditivo desenvolvido neste trabalho, vale ressaltar que as fórmulas tradicionais de previsão climática possuem limitações de aplicação de seus métodos, especialmente para análises de longo prazo, resultando em previsões que não se adequam as variabilidades naturais e antrópicas e apresentam dificuldades na representação de processos críticos (Williams, 2009). Portanto, a escolha dos algoritmos de redes neurais se mostrou natural mediante a facilidade que esses apresentam para contornar uma parte significativa destes problemas (Fernandes; Navaux; Portugal, 1996). Aponta-se, em contrapartida, que a aplicação de redes neurais para este tipo de estudo não é

capaz de prever o alcance de *tipping points*, pontos de inflexão, onde mudanças significativas e irreversíveis podem ocorrer, interrompendo o caráter cíclico de muitos fenômenos atmosféricos e, conseqüentemente, obsoletando a pesquisa a partir desse marco. Reforça-se que, de mesmo modo, a adoção dessa ferramenta metodológica contorna grande parte dos problemas enfrentados pelos métodos tradicionais, justificando o seu uso nesse trabalho.

Para Pessoa, Silva e Júnior (2006) e Guimarães *et al.* (2009), as redes neurais de arquitetura *feedforward* (FF) possuem um grande potencial para previsão climática quando treinadas com séries temporais, uma vez que podem aprender padrões a partir de conjuntos de dados históricos. Logo, ao serem alimentadas com dados meteorológicos passados, as RNAs FF podem aprender as relações sazonais e os padrões complexos presentes na atmosfera, permitindo previsões mais precisas. Assumiu-se, então, o modelo de arquitetura FF para composição das camadas, otimização dos hiperparâmetros e potencialização da previsão realizada, que contempla o intervalo de janeiro de 2020 até dezembro de 2049, totalizando 30 anos de predição.

Como forma de avaliar a previsão climática, seguiu-se a linha das medidas de incertezas estatísticas, adotando o cálculo do desvio padrão para elaboração da barra de erros, a fim de verificar se os pontos reais estão contidos dentro da variabilidade dos erros esperados. E, como segunda técnica utilizada, determinou-se os limites superiores e inferiores do intervalo de confiança de 80% para constatar se os valores reais também estão inseridos nessa amplitude (James *et al.*, 2017).

$$\text{Limites do IC} = \bar{x} \pm \frac{1,28 \cdot \sigma}{\sqrt{n}} \quad (01)$$

Onde:

\bar{x} = média da amostra;

σ = desvio padrão da amostra;

n = tamanho da amostra.

De forma análoga, consoante Zheng (2016), outra maneira direta de comprovar a confiabilidade do modelo de rede neural é realizando os testes de performance que avaliam o desempenho do ajuste dos dados em relação aos valores reais. Dessa forma, dispõem-se abaixo o Erro Médio Absoluto (MAE), o Erro Quadrático Médio (MSE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Percentual Absoluto Médio (MAPE).

Dissertando em ordem respectiva das métricas de performance e desempenho citadas no parágrafo anterior, o Erro Médio Absoluto (MAE) determina a média do erro absoluto entre as previsões e os valores reais (Zheng, 2016). Afirma-se, então, que quanto menor o MAE, melhor a previsão do modelo matemático.

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (02)$$

Onde:

MAE = Erro Médio Absoluto;

n = tamanho da amostra;

i = cada valor individual da amostra;

Y_i = valores reais;

\hat{Y}_i = valores previstos.

O Erro Quadrático Médio (MSE) expressa a média dos quadrados das diferenças entre os valores previstos e os valores reais, penalizando discrepâncias maiores entre os dados comparados.

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (03)$$

Onde:

MSE = Erro Quadrático Médio;

n = tamanho da amostra;

i = cada valor individual da amostra;

Y_i = valores reais;

\hat{Y}_i = valores previstos.

A Raiz do Erro Quadrático Médio (RMSE), por outro lado, expressa a raiz quadrada do MSE, tornando-a, portanto, uma métrica que atribui penalização ainda mais pronunciada em relação a valores discrepantes, pois eles são amplificados. Por elevarem todos os erros ao quadrado, o MSE e, principalmente, o RMSE se mostram úteis para indicar se a quantidade de valores dos outliers dos resíduos estiver alta (Montavon; Orr; Müller, 2012).

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (04)$$

Onde:

$RMSE$ = Raiz do Erro Quadrático Médio;

n = tamanho da amostra;

i = cada valor individual da amostra;

Y_i = valores reais;

\hat{Y}_i = valores previstos.

Dentro desse escopo, o Erro Percentual Absoluto Médio (MAPE) expõe a média das porcentagens absolutas das diferenças entre os valores previstos e os valores reais, divididos pelos valores reais, multiplicados por 100%, de forma a serem apresentados percentualmente (Zheng, 2016). Como padrão, por ser a porcentagem absoluta do erro médio, quanto menor o valor do MAPE, melhor a previsão.

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i} \cdot \frac{100}{100} \quad (05)$$

Onde:

$MAPE$ = Raiz do Erro Quadrático Médio;

n = tamanho da amostra;

i = cada valor individual da amostra;

Y_i = valores reais;

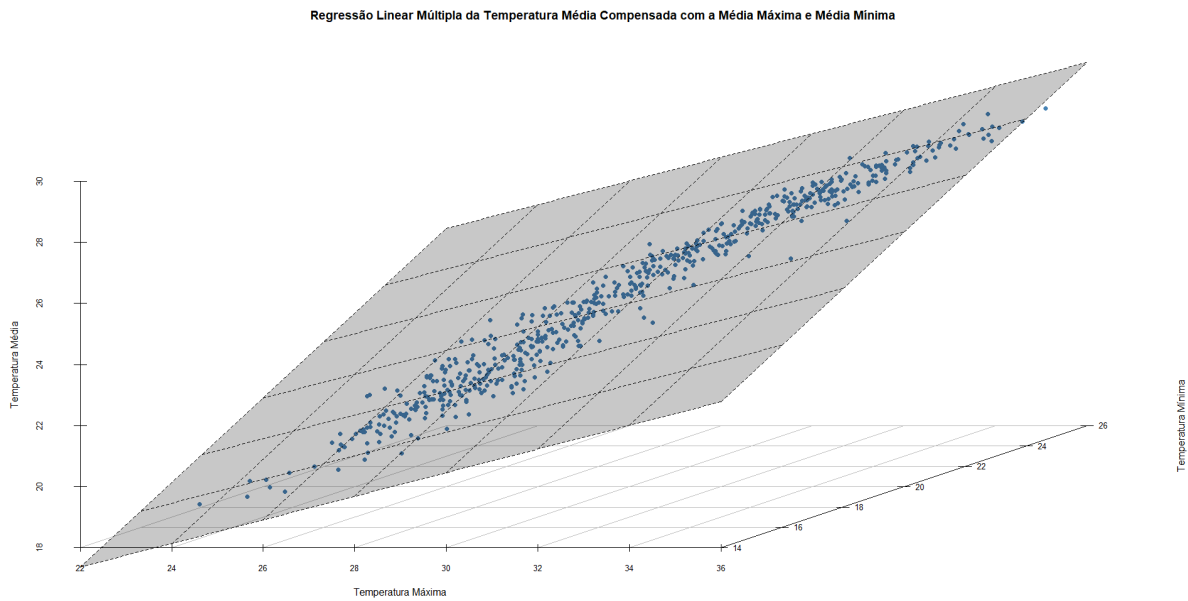
\hat{Y}_i = valores previstos.

3. DESENVOLVIMENTO

Observou-se, logo na etapa inicial de pré-processamento dos dados, o potencial para a realização de uma regressão linear múltipla para imputação dos dados ausentes no repositório. A temperatura média compensada, a temperatura máxima média e a temperatura mínima média apresentam um valor de multicolinearidade de 0,88, estando dentro do limite de aceitação de 0,9 e demonstrando uma clara correlação que possibilita esta aplicação (Seber; Lee, 2003). Explica-se que foi utilizado o Método de Spearman como técnica de regressão, pois a distribuição das variáveis não é gaussiana, desfavorecendo a aplicação do Método de Pearson (Diniz; Thiele, 2021).

Munido da regressão linear múltipla executada, analisou-se os resultados para filtragem de pontos discrepantes e refatoração dos parâmetros necessários. Após testagens realizadas a partir de algumas otimizações implementadas nos algoritmos, concluiu-se que o resultado alcançado é satisfatório e atende bem ao objetivo de tratamento dos dados que serão utilizados para treinamento da rede neural. A versão final do modelo de regressão encontra-se a seguir.

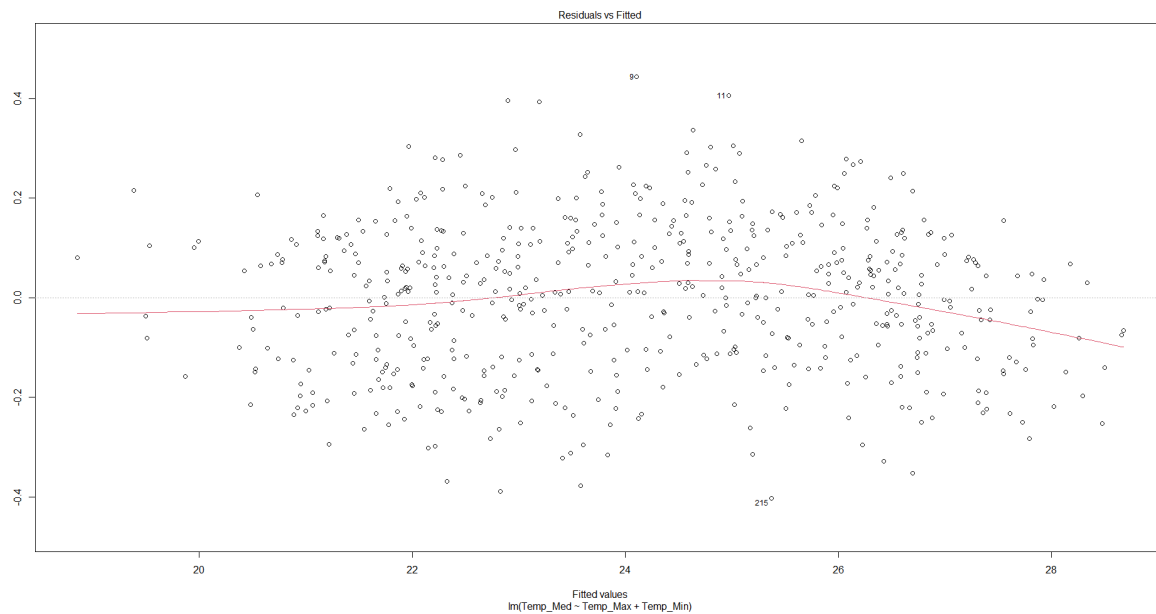
Gráfico 01 – Modelo de Regressão Linear Múltipla para Correção do Banco de Dados.



Fonte: Autores, 2023.

Com base no gráfico disponibilizado acima, nota-se que o modelo foi capaz de abranger quase todos os pontos distribuídos, comprovando a sua adequação ao problema estudado. Necessita-se, no entanto, entender se os resíduos apresentam bons resultados nos testes de qualidade do modelo, a fim de diagnosticar caso haja algum apontamento de comportamento indesejado na modelagem dos dados, atribuindo maior confiabilidade na imputação proposta pela regressão caso os resultados sejam excitatórios. Por conta disso, efetuou-se primeiro o teste de linearidade, que analisa a dispersão dos pontos dos valores dos resíduos do modelo em relação aos valores previstos para a variável dependente.

Gráfico 02 – Gráfico de Dispersão dos Resíduos do Modelo de Regressão (eixo Y) pelos Valores Previstos (eixo X).

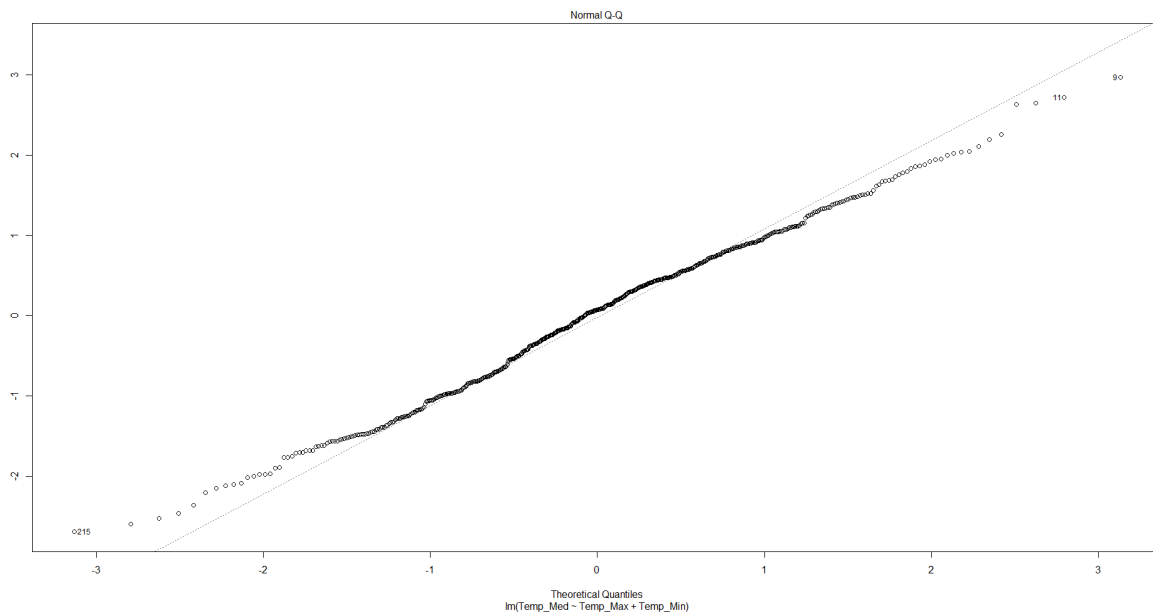


Fonte: Autores, 2023.

O gráfico expõe de forma clara que a dispersão apresenta um comportamento aproximadamente linear com 0,1 de tolerância, possibilitando interpretar que os coeficientes estipulados são apropriados, e validar os parâmetros estimados pelo modelo ao constatar que não são enviesados.

Em sequência, gerou-se o gráfico Q-Q dos resíduos do modelo comparando os quantis observados dos resíduos com os quantis esperados de uma distribuição normal, cujo resultado pode ser observado abaixo.

Gráfico 03 – Gráfico Q-Q dos resíduos do modelo de Regressão Linear Múltipla.

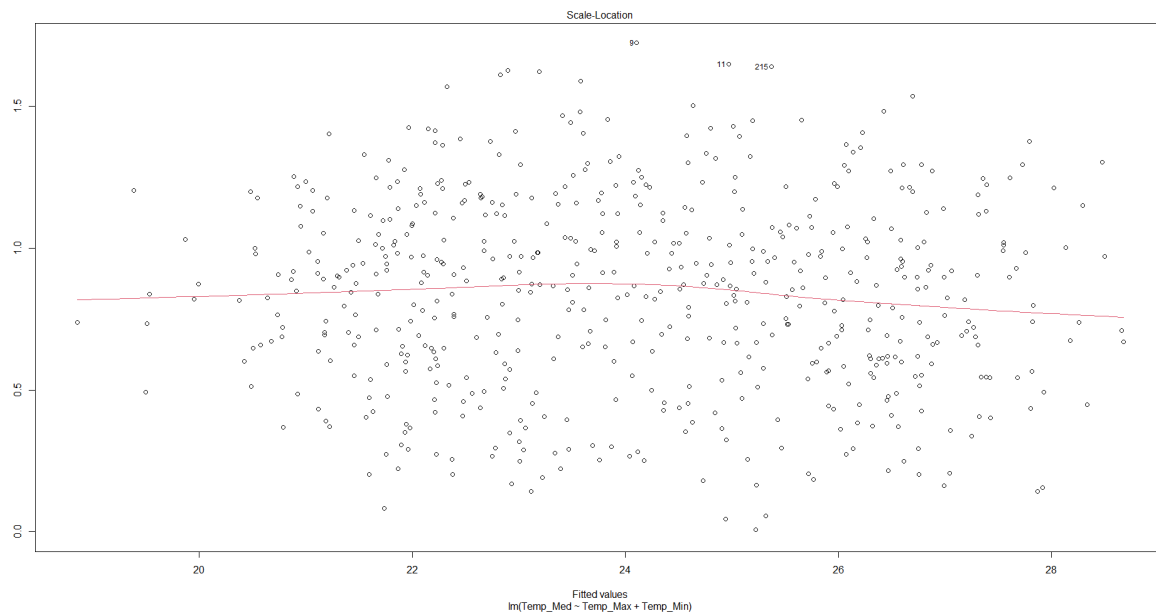


Fonte: Autores, 2023.

A suavidade com a qual os dados estão dispersos supõe a normalidade dos valores residuais, uma evidência crucial para a validade estatística das inferências associadas ao modelo, como intervalos de confiança e testes de hipóteses. Segundo Wilcox (2012), distribuição normal nos resíduos de um modelo possibilita construir intervalos de previsão ao redor das estimativas, o que ajuda a quantificar a incerteza das previsões.

Posteriormente a isso, construiu-se o gráfico da variância residual para observar se a dispersão dos resíduos é constante em todos os níveis das variáveis independentes, como exposto a seguir.

Gráfico 04 – Gráfico do Teste de Homocedasticidade da Regressão Linear Múltipla.

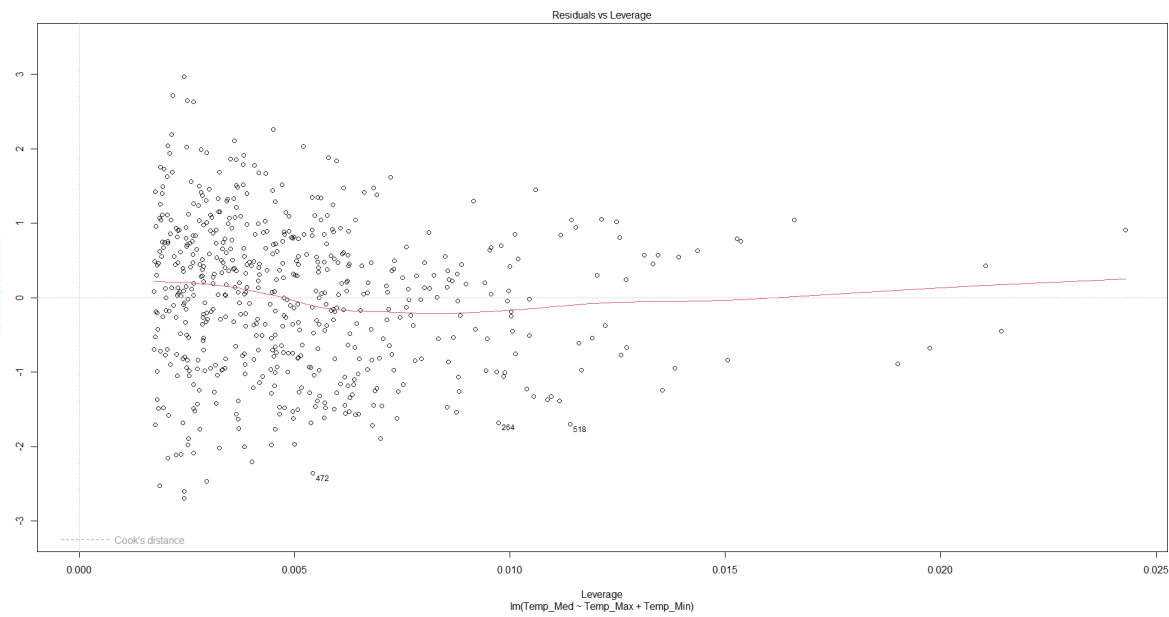


Fonte: Autores, 2023.

A imagem permite avaliar que a variabilidade dos resíduos é suficientemente constante em todos os níveis das variáveis independentes, como se fosse possível encaixar esta variabilidade em uma caixa retangular, isto é, a dispersão dos resíduos é a mesma ao longo do eixo X, não apresenta afunilamentos ou formas que indicariam um crescimento ou decrescimento na forma como estão dispersos, afirma-se, então, que a geometria da dispersão dos dados é homogênea. Esse resultado é crucial para se garantir a precisão das estimativas dos parâmetros, pois mostra que os mesmos não obtiveram maiores erros a partir de determinado ponto do problema estudado.

Ademais, como última análise do pré-processamento dos dados, investigou-se os pontos discrepantes e de alavancagem para avaliar a influência de cada observação no resultado geral do modelo.

Gráfico 05 – Gráfico de Cook para Observação de Pontos Discrepantes e de Alavancagem.



Fonte: Autores, 2023.

O Gráfico 05 evidencia que, considerando os escores entre 3 e -3, não existem pontos dos resíduos padronizados que fogem desse intervalo, negando que exista indícios de observações atípicas e valores discrepantes (Seber; Lee, 2003). Ainda nesse gráfico, é possível observar que não existem pontos de alavancagem nos resíduos, pois a influência dos resíduos padronizados não foi capaz de alterar o comportamento do gráfico em mais do que 0,3.

Uma vez aferida a qualidade da imputação realizado, a etapa de pré-processamento encerrou-se dividindo todo o histórico de 61 anos e 8 meses de dados em dois períodos: intervalo de treino, que compreende janeiro de 1960 até dezembro de 2019, e intervalo de teste, que contempla janeiro de 2020 até agosto de 2021. Acabada essa última etapa do tratamento, pôde-se dar início ao desenvolvimento da rede neural.

Dessa forma, munido dos dados tratados, iniciou-se a construção do modelo de rede neural determinando a arquitetura *feedforward* como estrutura organizacional das camadas, onde a informação flui invariavelmente da camada de entrada para a camada de saída, sem retroalimentação (Montavos; Orr; Müller, 2012). Explica-se que foi utilizada a técnica padrão de somar a quantidade de neurônios da camada de entrada e saída, dividido por dois, para se obter um valor

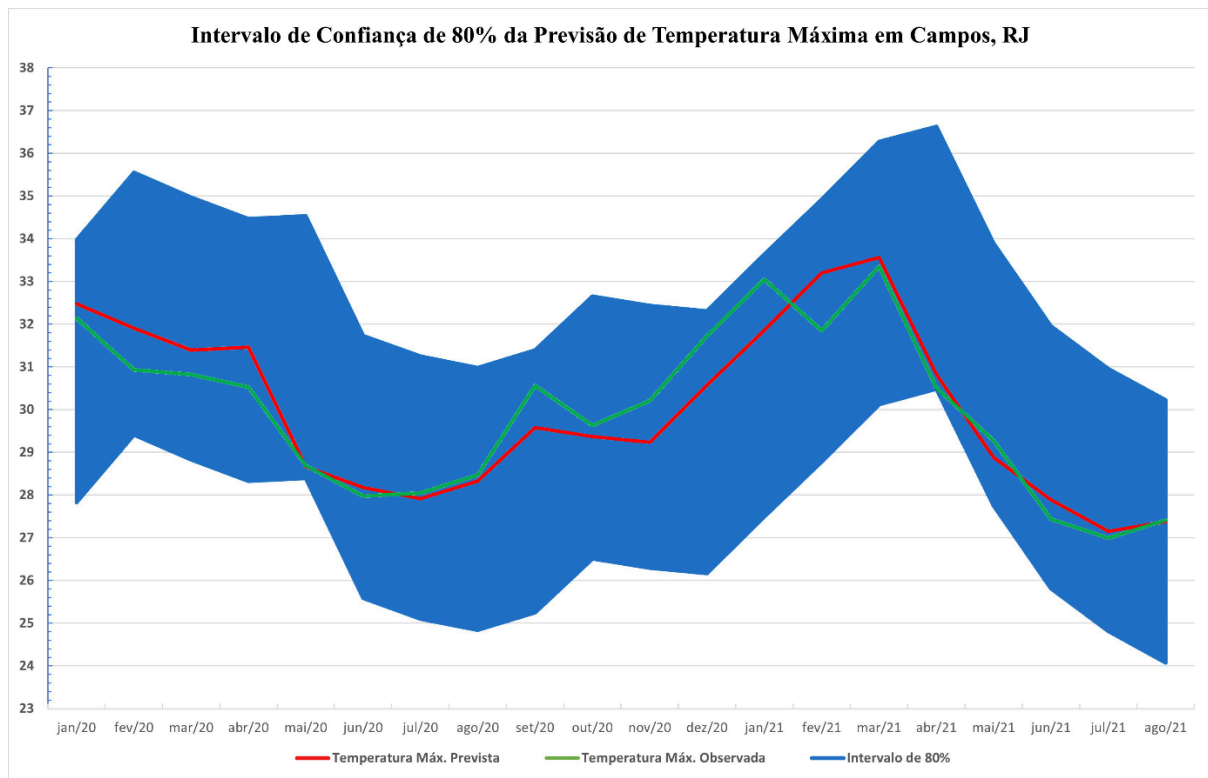
estimado de neurônios ideais para o problema, sem que haja complexidade desnecessária no modelo, isto é, evitando que ocorra um aumento de processamento sem uma devida evolução no coeficiente de determinação.

A arquitetura final da rede admite uma RNA FF com 29 neurônios na camada de entrada, 15 neurônios na camada oculta e 1 neurônio na camada de saída, sendo esta única saída, presumivelmente, a própria previsão climática em formato de série temporal, e todo o processo de treinamento foi realizado em 75 iterações com os dados que integram o intervalo de treino. Vale citar que é utilizada função de ativação ReLU na camada oculta, onde o valor de entrada permanecerá inalterado caso seja positivo e zerado caso seja negativo.

Além disso, alguns hiperparâmetros tiveram de ser otimizados experimentalmente ao longo do processo de construção do modelo, destacando-se a taxa de aprendizado e o tamanho do lote. A taxa de aprendizado ficou definida como 0,001 para que não desacelerasse o processo de aprendizado, mas contribuísse para evitar oscilações. As alterações no tamanho do lote, por outro lado, foram realizadas visando obter regularização implícita, evitando que o modelo se ajustasse demais aos detalhes específicos do conjunto de treinamento, o que dificultaria em generalizar para novos dados, essa sobreposição pôde ser contornada com o tamanho de lote 32.

Uma vez construído e treinado o modelo de previsão, realizou-se uma previsão de vinte meses, ou seja, um ano e oito meses, para ser comparada com o intervalo de teste, a fim de averiguar a qualidade dos valores supostos. Em posse dessa previsão e com o auxílio da Equação 01, emitiu-se uma imagem concatenando o intervalo de confiança de 80% da previsão junto aos valores reais observados no período, a visualização gráfica dessa comparação pode ser acompanhada na Gráfico 06.

Gráfico 06 – Intervalo de Confiança de 80% da Previsão Climática.

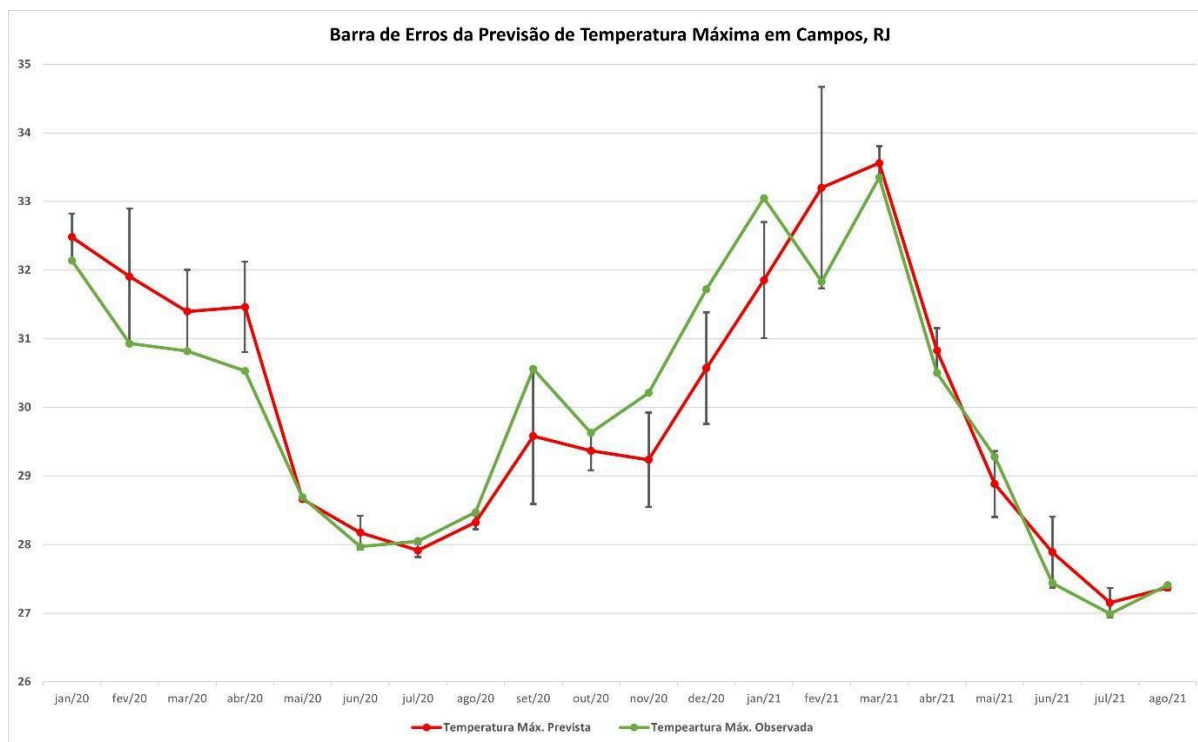


Fonte: Autores, 2023.

Evidencia-se que o intervalo de confiança de 80% contém os dados previstos e observados, não tendo sido necessária a emissão do intervalo de 95%, medida comumente utilizada em estudos do gênero (James *et al.*, 2017). Em termos práticos, a faixa do limite superior e inferior de 80% de chance da presença dos valores esperados já dispõe de todo o intervalo de teste, este resultado atribui significância a precisão estatística da previsão.

Por seu turno, como medida de incerteza suplementar para certificar se os resultados observados estão dentro das variações estimadas, gerou-se um gráfico de linhas com a barra de erros (Gráfico 07), calculada a partir do desvio padrão.

Gráfico 07 – Gráfico de Linhas com as Barras de Erro Padrão.



Fonte: Autores, 2023.

Nota-se que quatro pontos ficaram levemente fora das variações do erro padrão, apontando que a previsão possui momentos de maior imprecisão ao longo da sazonalidade da série. Esse indicativo revela a necessidade de cálculo das métricas de erro, para mensurar a dimensão dos erros da previsão, como executada e exposta abaixo.

Quadro 01 – Quadro das Métricas de Performance do Modelo Preditivo.

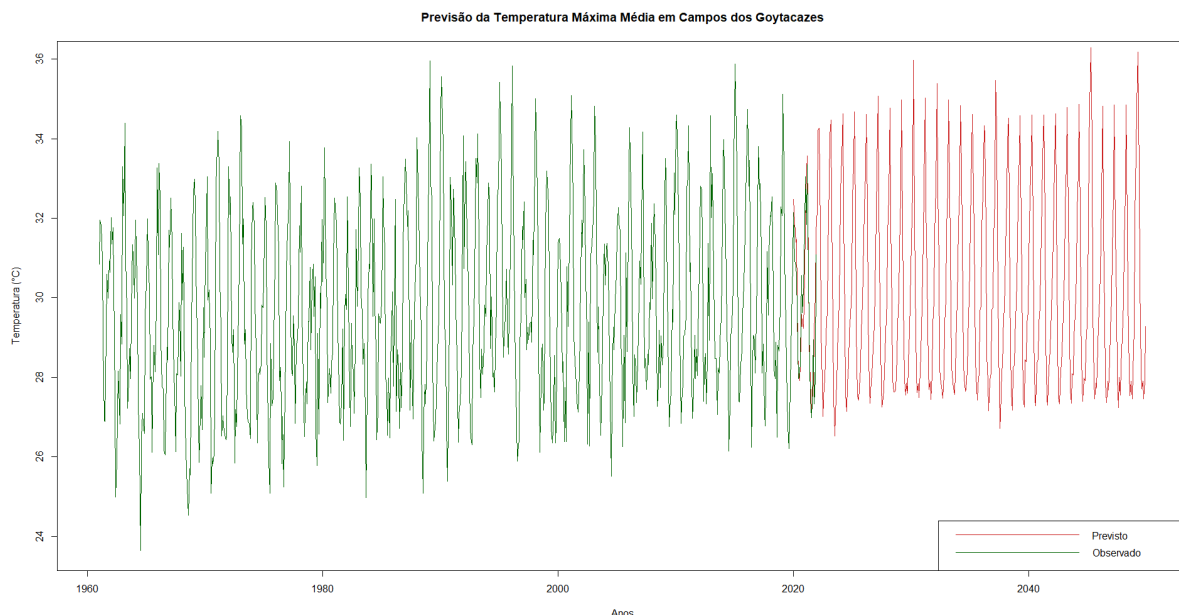
Métricas de Performance	Erros da Previsão do Modelo
MAE	0,6812833
MSE	0,7764269
RMSE	0,8811509
MAPE	2,281079

Fonte: Autores, 2023.

Apurados os resultados dos testes de performance, nota-se que o modelo apresenta um erro médio de $\pm 0,68$ °C para a maioria das previsões e de $\pm 0,88$ °C para as previsões menos precisas. Percentualmente, é esperado 2,28% de erro em cada ponto da série, para cima ou para baixo do valor real. Dessa forma, entende-se que o modelo preditivo construído cumpre bem o seu papel por possuir um resultado menor que 1 °C de erro, mesmo no pior cenário de caso de um ponto discrepante.

Devido as métricas de erro apresentarem resultados satisfatórios, realizou-se uma previsão de longo prazo com a rede neural desenvolvida, prevendo até dezembro de 2049. Tratando-se de uma previsão de longo prazo, optou-se por exibir o resultado em formato de série temporal, adicionando também as observações que integraram o intervalo de treino e teste, como visto abaixo.

Gráfico 08 – Série Temporal de Previsão Climática da Temperatura Máxima em Campos dos Goytacazes até dezembro de 2049.



Fonte: Autores, 2023.

Por meio do Gráfico 08, interpreta-se que a temperatura máxima média no município de Campos dos Goytacazes deve continuar a aumentar e a apresentar

eventos extremos pontuais, como picos de ondas de calor. O piso da temperatura apresenta um aumento, indicando que as mínimas da temperatura máxima média serão maiores. Esses indícios demonstram que alterações nos padrões climáticos podem ocorrer, como ao se observar na década de 60 em que a menor média da temperatura máxima chegou a menos de 24 °C e a maior média a pouco mais de 34 °C, enquanto na década de 40, segundo a previsão, a menor média estará em torno de 26,3 °C e a maior passará de 36 °C.

Dessa forma, os resultados obtidos dialogam com a discussão de Silva e Dereczynski (2014), que observam o aumento da temperatura máxima média na região do Norte/Noroeste Fluminense, onde se situa Campos dos Goytacazes, nas últimas décadas e apontam a região como a mais quente do Estado do Rio de Janeiro. Ressalta-se, ainda, que essa localidade é costeira e, por conta disso, apresenta uma menor variação de temperatura no ciclo de dia e noite e ao longo do ano, o que exige uma maior atenção e cuidado, uma vez que o aumento da temperatura máxima média tende a afetar todas as estações sazonais, favorecendo a ocorrência de eventos climáticos extremos, até mesmo com outras variáveis meteorológicas.

De modo inclusivo, o trabalho de Brasiliense *et al.* (2020) também reconhece o histórico ascendente da temperatura na região do Norte/Noroeste Fluminense e, somado aos indicativos de aumento na temperatura máxima média obtidos pelo modelo de rede neural, realça-se sobre como essas mudanças térmicas podem ocasionar aumentos na evaporação do Rio Paraíba do Sul, acarretando períodos mais longos de estiagem e impactando no abastecimento de água e na agricultura local, sobretudo para os pequenos produtores.

Vale citar, também, que o intenso desmatamento antrópico, promovido na região desde o século XIX para expansão agrícola, reduziu a cobertura florestal de Campos dos Goytacazes para 3% da representação original, contribuindo para a aceleração do aquecimento na cidade e entornos (Fabrício *et al.*, 2006). A região, que tem forte vocação agrícola, pode enfrentar desafios que impactam diretamente a sustentabilidade econômica e social, listam-se: 1 – A redução da produtividade devido as temperaturas mais altas afetarem o crescimento e o período de maturação agrícola; 2 – Condições climáticas extremas, mesmo que sazonais,

limitarem as espécies que podem ser cultivadas na região; 3 – As temperaturas mais altas provocarem um aumento na evaporação da água do solo, ocasionando uma maior carência das plantas e pressionando os recursos hídricos da região, como algumas das possíveis consequências desse aumento predito.

4. CONCLUSÃO

Foi construído um modelo preditivo de longo prazo para predição da temperatura máxima média na cidade de Campos dos Goytacazes, RJ. A rede neural desenvolvida se mostrou capaz de descrever as variações na temperatura máxima ao longo do período de teste, bem como continuou indicando pontuais eventos extremos durante as décadas preditas. Os valores obtidos nas métricas de desempenho apresentaram menos de 1 °C de variação para mais ou para menos, mesmo em caso de ponto de discrepância, e as medidas de incerteza demonstraram que o modelo atende a expectativa inicial.

A previsão realizada está de acordo com o que se encontra na literatura sobre a região, e dialoga sobre as consequências que deverão ser observadas caso esta se concretize. De acordo com os dados previstos, planejamentos acerca da adaptação e mitigação de danos na região de Campos dos Goytacazes podem ser necessários para garantir que o município continue com sua forte presença nos índices agrícolas do estado. Ademais, a administração pública poderá ter que se adaptar a essas mudanças para nível social e sanitário, uma vez que o calor é passível de aumentar a incidência de problemas cardiovasculares e desidratação, podendo impactar na saúde pública. Vale citar, além disso, que é esperado que as modificações supramencionadas se desdobrem dessa forma não apenas na realidade do município estudado, mas a nível mesoregional.

Para futuras pesquisas, sugere-se que seja explorada a construção de modelos de séries temporais, como o modelo auto-regressivo integrado de médias móveis, para a alimentação de uma futura rede neural, gerando, assim, um processo de modelagem híbrida que pode captar de forma ainda mais sensível os complexos padrões atmosféricos. Por fim, recomenda-se, também, que os *tipping points* (marcos de não-retorno) sejam mais explorados em análises porvindouras, estipulando novas interpretações ao se atingir esses pontos.

5. REFERÊNCIAS

- Box, G. E. P.; Jenkins, G. M.; Reinsel, G. C. (2008). Time Series Analysis: forecasting and control. Nova Jersey: Wiley, 4^a e.
- Brasil. Instituto Brasileiro de Geografia e Estatística. (2023). Ranking da Produção Agrícola no Estado do Rio de Janeiro. Rio de Janeiro: IBGE. Disponível em: <<https://cidades.ibge.gov.br/brasil/rj/campos-dos-goytacazes/pesquisa/14/10366?indicador=10246&tipo=ranking&ano=2022>>. Acesso em: 08 de dezembro de 2023.
- Brasiliense, C. S.; Dereczynski, C. P.; Satyamurty, P.; Chou, S. C.; Calado, R. N. (2020) Climatologias da Temperatura do Ar e da Precipitação na Bacia do Rio Paraíba do Sul, Região Sudeste do Brasil. Anuário do Instituto de Geociências – UFRJ, Rio de Janeiro, v. 43, n. 1, p. 35-365.
- Diniz, É. S.; Thiele, J. (2021). Modelos de Regressão em R. Clube de Autores: Joinville, 2^a e.
- Fabrizio, A. C.; João Marcelo, A. B.; José Manoel, L. G.; Josival, S. S.; Marcelo, T. N. (2006). Comunidade Arbórea de Uma Floresta de Baixada Aluvial no Município de Campos dos Goytacazes, RJ. CERNE, Lavras, v. 12, n. 2, p. 157-166.
- Fernandes, L. G L.; Navaux, P. O. A; Portugal, M. S. (1996). Previsão de Séries de Tempo: redes neurais artificiais e modelos estruturais. Pesquisa e Planejamento Econômico, v. 26, n. 2, p. 1-19.
- Guimarães, D. P.; Assad, E. D.; Landau, E. C.; Correa, T.; Costa, C. (2009). Uso de Redes Neurais para a Projeção Climática no Brasil: temperatura mínima. AINFO – Agência de Informação Embrapa e Acervo Documental, Brasília, p. 1-6.
- Haykin, S. (2000). Redes Neurais: princípios e prática. Tradução: Paulo Martins Engel. Porto Alegre: Bookman, 2^o e.
- James, G.; Witten, D.; Hastie, T.; Tibshirani, R. (2017). An Introduction to Statistical Learning: with Applications in R. Berlim: Springer, 7^a e.
- Montavon, G; Orr, G; Müller, K. R. (2012). Neural Networks: Tricks of the Trade. Berlim: Springer, 2^a e.

Pessoa, A. S. A.; Silva, J. D. S.; Júnior, H. C. (2006). Redução de Dados Meteorológicos Aplicados a Previsão Climática por Redes Neurais. MTC INPE – Memória Tecnológica e Científica do Instituto Nacional de Pesquisas Espaciais, p. 1-6.

Saucier, W. J. (1969). Princípios de Análise Meteorológica. Tradução: Douglas Strang. Rio de Janeiro: Ao Livro Técnico, 1ª e.

Schiavo, E. H. C.; Azevedo, M. P.; Xavier, L. de S.; Peixoto, J. A. A. (2004). Estudo comparativo entre redes neurais artificiais e análise de regressão múltipla na avaliação de bens, para pequenas amostragens. XXIV Encontro Nacional de Engenharia de Produção, Florianópolis, p. 3216-3222.

Seber, G. A. F.; Lee, A. J. (2003). Linear Regression Analysis. Nova Jersey: Wiley, 2ª e.

Silva, W. L.; Dereczynski, C. P. (2014). Caracterização Climatológica e Tendências Observadas em Extremos Climáticos no Estado do Rio de Janeiro. Anuário do instituto de Geociências – UFRJ, Rio de Janeiro, v. 37, n. 2, p. 123-138.

Williams, J. (2009). The AMS Weather Book: The Ultimate Guide to America's Weather. Chicago: University of Chicago Press, 1ª e.

Ynoue, R. Y.; Michelle, S. R.; Ambrizzi, T.; Silva, G. A. M. (2017). Meteorologia: noções básicas. São Paulo: Oficina de Textos, 1ª e.

Zheng, A. (2016). Evaluating Machine Learning Models. Sebastopol: O'Reilly Media, 1ª e.