

PREVISÃO DA SÉRIE TEMPORAL DE PREÇOS DO
BIODIESEL USANDO REDES NEURAIS

Ellizeu Rodrigues Sena

Projeto de Graduação apresentado ao Curso
de Engenharia Eletrônica e de Computação
da Escola Politécnica, Universidade Federal
do Rio de Janeiro, como parte dos requisitos
necessários à obtenção do título de Enge-
nheiro.

Orientador: José Manoel de Seixas

Rio de Janeiro
Novembro de 2024

PREVISÃO DA SÉRIE TEMPORAL DE PREÇOS DO
BIODIESEL USANDO REDES NEURAIS

Ellizeu Rodrigues Sena

PROJETO DE GRADUAÇÃO SUBMETIDO AO CORPO DOCENTE DO CURSO
DE ENGENHARIA ELETRÔNICA E DE COMPUTAÇÃO DA ESCOLA PO-
LITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE ENGENHEIRO ELETRÔNICO E DE COMPUTAÇÃO

Autor:



Ellizeu Rodrigues Sena

Orientador:

Prof. José Manoel de Seixas, D. Sc.

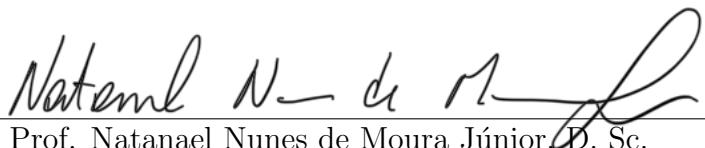
Examinador:

Prof. José Gabriel Rodriguez Carneiro Gomes, Ph. D.

Examinador:

Kriseida Carmen Portella Guedelha Alekseev, D. Sc.

Examinador:


Prof. Natanael Nunes de Moura Júnior, D. Sc.

Rio de Janeiro

Novembro de 2024

Declaração de Autoria e de Direitos

Eu, *Ellizeu Rodrigues Sena* CPF 058.510.307-04, autor da monografia *Previsão da Série Temporal de Preços do Biodiesel usando Redes Neurais*, subscrevo para os devidos fins, as seguintes informações:

1. O autor declara que o trabalho apresentado na disciplina de Projeto de Graduação da Escola Politécnica da UFRJ é de sua autoria, sendo original em forma e conteúdo.
2. Excetuam-se do item 1. eventuais transcrições de texto, figuras, tabelas, conceitos e idéias, que identifiquem claramente a fonte original, explicitando as autorizações obtidas dos respectivos proprietários, quando necessárias.
3. O autor permite que a UFRJ, por um prazo indeterminado, efetue em qualquer mídia de divulgação, a publicação do trabalho acadêmico em sua totalidade, ou em parte. Essa autorização não envolve ônus de qualquer natureza à UFRJ, ou aos seus representantes.
4. O autor pode, excepcionalmente, encaminhar à Comissão de Projeto de Graduação, a não divulgação do material, por um prazo máximo de 01 (um) ano, improrrogável, a contar da data de defesa, desde que o pedido seja justificado, e solicitado antecipadamente, por escrito, à Congregação da Escola Politécnica.
5. O autor declara, ainda, ter a capacidade jurídica para a prática do presente ato, assim como ter conhecimento do teor da presente Declaração, estando ciente das sanções e punições legais, no que tange a cópia parcial, ou total, de obra intelectual, o que se configura como violação do direito autoral previsto no Código Penal Brasileiro no art.184 e art.299, bem como na Lei 9.610.
6. O autor é o único responsável pelo conteúdo apresentado nos trabalhos acadêmicos publicados, não cabendo à UFRJ, aos seus representantes, ou ao(s) orientador(es), qualquer responsabilização/ indenização nesse sentido.
7. Por ser verdade, firmo a presente declaração.



Ellizeu Rodrigues Sena

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Escola Politécnica - Departamento de Eletrônica e de Computação
Centro de Tecnologia, bloco H, sala H-217, Cidade Universitária
Rio de Janeiro - RJ CEP 21949-900

Este exemplar é de propriedade da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazenar em computador, microfilmar ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es).

AGRADECIMENTO

Agradeço à Deus, em primeiro lugar, porque acredito que sem ele nada é possível.

Aos meus pais Helena Dias e Euzimar Sena, pelos conselhos e apoio em todos os momentos da minha vida.

À minha esposa, Mayara Sena, pela companhia nas idas pro Fundão, por todo o incentivo e por todo apoio no desenvolvimento do trabalho.

Aos meus amigos da UFRJ, Caroline Coelho, Pedro Lídio, João Pedro, Iuri Veloso e Breno Galves, por permanecerem comigo nos momentos bons e ruins.

Aos meus supervisores da EPE, Kriseida Alekseev e Dan Gadelman, por me proporcionarem a primeira experiência profissional e por todo o incentivo a seguir carreira acadêmica.

Ao meu orientador, José de Seixas, por aceitar me orientar e proporcionar a realização desse trabalho.

Aos professores que gentilmente aceitaram fazer parte da minha banca, Natanael Júnior, José Gabriel e Kriseida Alekseev.

RESUMO

O biodiesel é um biocombustível renovável produzido de fontes vegetais e animais, desempenha um papel importante na transição energética global. A previsão precisa de seu preço é essencial para otimizar sua produção e comercialização, além de apoiar decisões no mercado de biocombustíveis, como o percentual obrigatório de mistura ao diesel.

Este trabalho explora a aplicação de Redes Neurais Artificiais (RNAs) e modelos estatísticos para prever o preço do biodiesel, aliados a modelos de extração de características, como tendência, ciclos senoidais e sazonais, para melhorar o desempenho das previsões. A pesquisa investiga o uso de diferentes modelos para prever o preço do biodiesel, considerando variáveis como o preço do óleo de soja e os preços de biodiesel por regiões do Brasil. As RNAs foram escolhidas pela sua capacidade de modelar padrões não lineares e relações complexas entre essas variáveis.

Os resultados mostram que as RNAs têm bom desempenho na previsão do preço do biodiesel, obtendo medidas de desempenho melhores que os modelos estatísticos. Modelos menos profundos, como Linear Normalizado (do inglês *Normalized Linear*, N-Linear) e Perceptron Multicamadas (do inglês *Multilayer Perceptron*, MLP), apresentaram melhor resultado nas previsões em comparação aos mais profundos, sugerindo que abordagens mais simples podem ser mais eficazes para capturar os padrões dinâmicos dos preços do biodiesel. Modelos morfológicos, com Perceptron Morfológico Crescente (do inglês *Increasing Morphological Perceptron*, IMP) e Neurônio Crescente Decrescente Linear (do inglês *Increasing-Decreasing Linear Neuron*, IDLN), também foram testados, mas não apresentaram desempenho satisfatório nas previsões.

Palavras-Chave: previsão de séries temporais, biodiesel, redes neurais, inteligência artificial, biocombustíveis, sustentabilidade.

ABSTRACT

Biodiesel is a renewable biofuel produced from plant and animal sources and plays an important role in the global energy transition. Accurately predicting its price is essential to optimize its production and commercialization, in addition to supporting decisions in the biofuel market, such as the mandatory percentage of blending with diesel.

This work explores the application of Artificial Neural Networks (ANN) and statistical models to predict the price of biodiesel, combined with feature extraction models, such as trends, sinusoidal and seasonal cycles, to improve the performance of the predictions. The research investigates the use of different models to predict the price of biodiesel, considering variables such as the price of soybean oil and biodiesel prices by region of Brazil. ANN were chosen for their ability to model non-linear patterns and complex relationships between these variables.

The results show that ANNs perform well in predicting biodiesel prices, obtaining better performance measures than statistical models. Shallower models, such as Normalized Linear (**N-Linear**) and Multilayer Perceptron (**MLP**), performed better in predictions compared to deeper ones, suggesting that simpler approaches may be more effective in capturing the dynamic patterns of biodiesel prices. Morphological models, with Increasing-Decreasing Linear Neuron (**IDLN**) and Increasing Morphological Perceptron (**IMP**), were also tested, but did not show satisfactory performance in the predictions.

Key-words: time series forecasting, biodiesel, neural networks, artificial intelligence, biofuels, sustainability.

LISTA DE ABREVIATURAS

ANP Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

ARIMA Modelo Autoregressivo Integrado de Médias Móveis (do inglês *Autoregressive Integrated Moving Average Model*)

CUDA Arquitetura de Dispositivo de Computação Unificada (do inglês *Compute Unified Device Architecture*)

DA-RNN Rede Neural Recorrente Baseada em Atenção de Dois Estágios (do inglês *Dual-Stage Attention-Based Recurrent Neural Network*)

GPU Unidade de Processamento Gráfico (do inglês *Graphics Processing Unit*)

IDLN Neurônio Crescente Decrescente Linear (do inglês *Increasing-Decreasing Linear Neuron*)

IMP Perceptron Morfológico Crescente (do inglês *Increasing Morphological Perceptron*)

IPCA Índice Nacional de Preços ao Consumidor Amplo

LPS Laboratório de Processamento de Sinais

LSTM Rede Neural com Memória de Longo e Curto Prazo (do inglês *Long Short-Term Memory*)

MAE Erro Absoluto Médio (do inglês *Mean Absolute Error*)

MAPE Erro Percentual Absoluto Médio (do inglês *Mean Absolute Percentage Error*)

MLP Perceptron Multicamadas (do inglês *Multilayer Perceptron*)

MSE Erro Quadrático Médio (do inglês *Mean Squared Error*)

N-Linear Linear Normalizado (do inglês *Normalized Linear*)

NARX Rede Neural Autoregressiva Não-linear com Entradas Exógenas (do inglês *Nonlinear Autoregressive Exogenous Model*)

NHiTS Interpolação Neural Hierárquica para Previsão de Séries Temporais (do inglês *Neural Hierarchical Interpolation for Time Series Forecasting*)

NLP Processamento de Linguagem Natural (do inglês *Natural Language Processing*)

NLPCA Análise de Componentes Principais Não Lineares (do inglês *Nonlinear Principal Component Analysis*)

OSI Iniciativa de código aberto (do inglês *Open Source Initiative*)

PCA Análise de Componentes Principais (do inglês *Principal Component Analysis*)

PNPB Programa Nacional de Produção e Uso de Biodiesel

ReLU Unidade Linear Retificada (do inglês *Rectified Linear Unit*)

RMSE Raiz do Erro Quadrático Médio (do inglês *Root Mean Squared Error*)

RNA Rede Neural Artificial

RNN Rede Neural Recorrente (do inglês *Recurrent Neural Network*)

RW Passeio Aleatório (do inglês *Random Walk*)

SLE Erro Quadrado Logarítmico (do inglês *Squared Logarithmic Error*)

THEIL Estatística U de Theil

Sumário

1 Introdução	1
1.1 Tema e Delimitação	2
1.2 Motivação e Justificativa	2
1.3 Objetivos	3
1.4 Método	4
1.5 Materiais	5
1.6 Estrutura do Trabalho	6
2 O Biodiesel	8
2.1 Vantagens em Relação ao Óleo Diesel	8
2.2 Produção de Biodiesel no Brasil	9
2.3 Regulação e Início do Mercado	9
2.3.1 Percentuais de Adição no Diesel	10
3 Redes Neurais Artificiais	13
3.1 Perceptron de Rosenblatt	14
3.1.1 Limitações e Extensões	14
3.2 Treinamento de Redes Neurais	15
3.2.1 Retropropagação	16
3.3 Funções de Ativação	17
3.4 Arquiteturas	18
4 Previsão de Séries Temporais	20
4.1 Métodos de Decomposição e Normalização	21
4.1.1 Decomposição dos Dados	21
4.1.2 Normalização dos Dados	23

4.2 Janelas Temporais	24
4.2.1 Janela Deslizante	25
4.2.2 Janela de Takens	25
4.3 Medidas de Avaliação do Modelo	26
4.3.1 Erro Percentual Absoluto Médio (do inglês <i>Mean Absolute Percentage Error</i>, MAPE)	26
4.3.2 Erro Quadrático Médio (do inglês <i>Mean Squared Error</i>, MSE)	26
4.3.3 Raiz do Erro Quadrático Médio (do inglês <i>Root Mean Squared Error</i>, RMSE)	27
4.3.4 Erro Absoluto Médio (do inglês <i>Mean Absolute Error</i>, MAE)	27
4.3.5 Estatística U de Theil (THEIL)	27
4.4 Coeficiente de Hurst	28
4.4.1 Definição e Cálculo	29
5 Descrição do Método	30
5.1 Coleta de Dados	30
5.2 Divisão do Conjunto de Dados	31
5.3 Pré-processamento dos Dados	32
5.4 Treinamento e Avaliação dos Modelos	35
6 Resultados	37
6.1 Janela Deslizante	37
6.1.1 NARX	38
6.1.2 DA-RNN	38
6.1.3 MLP	44
6.1.4 IMP	45
6.1.5 IDLN	47
6.1.6 N-Linear	50
6.1.7 NHITS	53
6.2 Janela de Takens	56
6.2.1 NARX	61
6.2.2 DA-RNN	62
6.2.3 MLP	65

6.2.4	IMP	68
6.2.5	IDLN	68
6.2.6	N-Linear	71
6.2.7	NHiTS	74
7	Conclusões e Trabalhos Futuros	90
Bibliografia		93
A Arquitetura dos Modelos para Previsão de Séries Temporais		99
A.1	Rede Neural Autoregressiva Não-linear com Entradas Exógenas (do inglês <i>Nonlinear Autoregressive Exogenous Model</i> , NARX)	100
A.2	Rede Neural Recorrente Baseada em Atenção de Dois Estágios (do inglês <i>Dual-Stage Attention-Based Recurrent Neural Network</i> , DA-RNN)	101
A.3	Perceptron Multicamadas (do inglês <i>Multilayer Perceptron</i> , MLP)	102
A.4	Perceptron Morfológico Crescente (do inglês <i>Increasing Morphological Perceptron</i> , IMP)	103
A.5	Neurônio Crescente Decrescente Linear (do inglês <i>Increasing-Decreasing Linear Neuron</i> , IDLN)	105
A.6	Linear Normalizado (do inglês <i>Normalized Linear</i> , N-Linear)	105
A.7	Interpolação Neural Hierárquica para Previsão de Séries Temporais (do inglês <i>Neural Hierarchical Interpolation for Time Series Forecasting</i> , NHiTS)	106

Lista de Figuras

2.1 Série de Preços do Biodiesel no Brasil (R\$/litro) a partir de 2022	10
3.1 Diagrama da MLP nas fases de retropropagação	16
5.1 Fluxo de Processamento	30
5.2 Matriz de Correlação entre as Séries Candidatas (01/2022 - 07/2023)	32
5.3 Linha do Tempo da Divisão dos Conjuntos	32
5.4 Fluxo de Pré-Processamento	33
5.5 Autocorrelação da Série de Preços do Biodiesel	34
5.6 Frequências de Fourier da Série de Preços do Biodiesel	34
5.7 Linha do Tempo dos Períodos de Disponibilidade das Séries de Bio- diesel e Óleo de Soja	35
5.8 Validação Cruzada com 3 dobras	36
6.1 Resultados do modelo NARX no conjunto de teste	39
6.2 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo NARX no conjunto de teste	40
6.3 Histograma dos erros do modelo NARX no conjunto de teste	41
6.4 Resultados do modelo DA-RNN no conjunto de teste	42
6.5 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo DA-RNN no conjunto de teste	43
6.6 Histograma dos erros do modelo DA-RNN no conjunto de teste	44
6.7 Resultados do modelo MLP no conjunto de teste	45
6.8 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo MLP no conjunto de teste	46
6.9 Histograma dos erros do modelo MLP no conjunto de teste	47
6.10 Resultados do modelo IMP no conjunto de teste	48

6.11 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo IMP no conjunto de teste	49
6.12 Histograma dos erros do modelo IMP no conjunto de teste	50
6.13 Resultados do modelo IDLN no conjunto de teste	51
6.14 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo IDLN no conjunto de teste	52
6.15 Histograma dos erros do modelo IDLN no conjunto de teste	53
6.16 Resultados do modelo N-Linear no conjunto de teste	54
6.17 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo N-Linear no conjunto de teste	55
6.18 Histograma dos erros do modelo N-Linear no conjunto de teste	56
6.19 Resultados do modelo NHITS no conjunto de teste	57
6.20 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo NHITS no conjunto de teste	58
6.21 Histograma dos erros do modelo NHITS no conjunto de teste	59
6.22 Resultados da Validação Cruzada treinado com Biodiesel Nacional usando Janela Deslizante	59
6.23 Resultados da Validação Cruzada treinado com Biodiesel Nacional e Óleo de Soja usando Janela Deslizante	60
6.24 Resultados do Conjunto de Teste treinado com Biodiesel Nacional usando Janela Deslizante	60
6.25 Resultados do Conjunto de Teste treinado com Biodiesel Nacional e Óleo de Soja usando Janela Deslizante	61
6.26 Resultados do modelo NARX Takens no conjunto de teste	63
6.27 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo NARX Takens no conjunto de teste	64
6.28 Histograma dos erros do modelo NARX Takens no conjunto de teste	65
6.29 Resultados do modelo DA-RNN Takens no conjunto de teste	66
6.30 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo DA-RNN Takens no conjunto de teste	67
6.31 Histograma dos erros do modelo DA-RNN Takens no conjunto de teste	68
6.32 Resultados do modelo MLP Takens no conjunto de teste	69

6.33 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo MLP Takens no conjunto de teste	70
6.34 Histograma dos erros do modelo MLP Takens no conjunto de teste	71
6.35 Resultados do modelo IMP Takens no conjunto de teste	72
6.36 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo IMP Takens no conjunto de teste	73
6.37 Histograma dos erros do modelo IMP Takens no conjunto de teste	74
6.38 Resultados do modelo IDLN Takens no conjunto de teste	75
6.39 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo IDLN Takens no conjunto de teste	76
6.40 Histograma dos erros do modelo IDLN Takens no conjunto de teste	77
6.41 Resultados do modelo N-Linear Takens no conjunto de teste	78
6.42 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo N-Linear Takens no conjunto de teste	79
6.43 Histograma dos erros do modelo N-Linear Takens no conjunto de teste	80
6.44 Resultados do modelo NHITS Takens no conjunto de teste	83
6.45 Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo NHITS Takens no conjunto de teste	84
6.46 Histograma dos erros do modelo NHITS Takens no conjunto de teste	85
6.47 Resultados da Validação Cruzada treinado com Biodiesel Nacional usando Janela de Takens	85
6.48 Resultados da Validação Cruzada treinado com Biodiesel Nacional e Óleo de Soja usando Janela de Takens	87
6.49 Resultados do Conjunto de Teste treinado com Biodiesel Nacional usando Janela de Takens	87
6.50 Resultados do Conjunto de Teste treinado com Biodiesel Nacional e Óleo de Soja usando Janela de Takens	89
A.1 Diagrama da NARX	101
A.2 Diagrama da DA-RNN	102
A.3 Diagrama da MLP	103
A.4 Diagrama da IMP	104
A.5 Diagrama da IDLN	105

A.6 Diagrama da N-Linear	106
A.7 Diagrama da NHITS	107

Lista de Tabelas

2.1 Percentual de Mistura de Biodiesel no Brasil ao Longo do Tempo	12
6.1 Resultados da Validação Cruzada usando Janela Deslizante	81
6.2 Resultados do Conjunto de Teste usando Janela Deslizante	82
6.3 Resultados da Validação Cruzada usando Janela de Takens	86
6.4 Resultados do Conjunto de Teste usando Janela de Takens	88

Capítulo 1

Introdução

As séries temporais são conjuntos de observações coletadas em intervalos regulares e com dependências adjacentes, refletindo o comportamento histórico de um evento ao longo do tempo [1]. A previsão dessas séries é importante para antecipar eventos futuros, como preços de ações e demanda de energia [2, 3]. No entanto, muitas séries temporais não são sistemas invariantes no tempo, sendo influenciadas por mudanças estruturais e fatores externos [4]. Por isso, modelos avançados são necessários para lidar com essas variações e fatores dinâmicos, buscando uma maior consistência e acurácia nas previsões.

A previsão de preços de biodiesel é desafiadora devido às influências econômicas, ambientais e políticas que afetam esse mercado emergente, tornando essencial o uso de técnicas avançadas, como modelos estatísticos e redes neurais, para obter previsões com baixo erro [5]. Além disso, a importância das energias renováveis e biocombustíveis, como o biodiesel, está crescendo na busca por soluções sustentáveis e na desaceleração das mudanças climáticas [6]. A adoção dessas fontes de energia reduz a dependência de combustíveis fósseis e promove uma economia de baixo carbono, destacando a necessidade de previsões de preços confiáveis para apoiar a tomada de decisões no setor.

Redes neurais são eficazes na previsão de séries temporais porque podem identificar relações não lineares e lineares [7], adaptando-se a novos dados e mudanças no mercado. Esse potencial permite uma tomada de decisão mais informada para diversas partes interessadas, incluindo produtores e consumidores, contribuindo para

uma gestão mais eficiente de estoques e planejamento estratégico.

Desta forma, a aplicação de modelos avançados de redes neurais na previsão de preços de biodiesel é uma abordagem promissora para melhorar a precisão das previsões. Neste contexto, este trabalho propõe o desenvolvimento de um modelo de previsão de preços de biodiesel utilizando redes neurais, de modo a fornecer uma ferramenta confiável e precisa para antecipar as variações de preços nesse mercado em constante evolução.

1.1 Tema e Delimitação

O tema deste trabalho é a previsão da série temporal de Preços do Biodiesel [8].

Neste sentido, o problema a ser resolvido é o treinamento do modelo com uma série temporal curta de preços de Biodiesel e as séries correlatas do Óleo de Soja, a fim de prever os preços futuros com precisão e confiabilidade.

Este trabalho se limitará a uma previsão por um período curto do preço do biodiesel no Brasil, aproximadamente 3 meses. Por se tratar de um mercado em desenvolvimento ainda não há dados suficientes para uma previsão mais longa. Serão utilizados conjuntos de dados públicos do Biodiesel no Brasil [8] e do Óleo de Soja no Banco Mundial [9].

1.2 Motivação e Justificativa

O aumento das concentrações de CO_2 na atmosfera é um dos principais fatores responsáveis pelas mudanças climáticas e pelo aquecimento global. As emissões de gases de efeito estufa provenientes da queima de combustíveis fósseis têm impactos graves no meio ambiente, incluindo o derretimento das calotas polares, o aumento do nível do mar e eventos climáticos extremos. Diante desse cenário, a busca por alternativas sustentáveis e menos poluentes, como os biocombustíveis, torna-se essencial para mitigar os efeitos adversos dessas emissões e promover uma transição para fontes de energia mais limpas.

O uso de um biocombustível pode contribuir para a redução das emissões de CO_2 . A previsão de preços de biodiesel apresenta um desafio complexo devido ao seu comportamento dinâmico e às influências de fatores econômicos, ambientais e políticos. O mercado de biodiesel, ainda em desenvolvimento, é sujeito a variáveis que afetam tanto a oferta quanto a demanda, além de possuir poucos dados para treinamento, tornando as previsões de preços uma tarefa complexa.

As redes neurais oferecem uma abordagem eficaz e robusta para enfrentar esses problemas. Ao analisar grandes volumes de dados e identificar padrões complexos, essas técnicas avançadas podem fornecer previsões mais precisas e confiáveis. A utilização de redes neurais no desenvolvimento de modelos preditivos para o mercado de biodiesel pode facilitar a tomada de decisões estratégicas, como a otimização de estoques e o planejamento de investimentos, além de ajudar na adaptação às variações do mercado.

Em resumo, a aplicação de redes neurais na previsão de preços de biodiesel não só contribui para uma análise mais detalhada e precisa das variações de preços, mas também apoia a transição para fontes de energia mais sustentáveis. Dessa forma, a pesquisa contribui para uma melhor compreensão do mercado de biodiesel e para o desenvolvimento de estratégias mais eficientes que apoiem a redução das emissões de CO_2 e a promoção de alternativas energéticas sustentáveis.

1.3 Objetivos

O objetivo deste trabalho é desenvolver um modelo de previsão de preços de biodiesel usando redes neurais para fornecer uma ferramenta confiável e precisa. Essa abordagem visa antecipar variações de preços em um mercado em evolução, ajudando na otimização de decisões estratégicas e na melhoria da gestão no setor de energia renovável.

O objetivo geral é, então, desenvolver um modelo de previsão de preços de biodiesel utilizando redes neurais em um período limitado. Desta forma, tem-se como objetivos específicos: (1) realizar a coleta e análise de dados relevantes, como histórico

de preços, indicadores econômicos e valores de matérias primas; (2) desenvolver um modelo de redes neurais capaz de processar esses dados e identificar padrões complexos que influenciam os preços; (3) avaliar e aprimorar o desempenho do modelo por meio de técnicas de otimização de hiperparâmetros e medidas de desempenho; (4) aplicar o modelo desenvolvido para prever os preços futuros.

1.4 Método

O método adotado neste trabalho consiste em utilizar as séries de preços regionais do biodiesel, juntamente com outras variáveis correlatas. Essas séries temporais passam por uma análise de pré-processamento e, em seguida, são aplicados modelos de Machine Learning e Redes Neurais para realizar a modelagem e previsão dos preços nacionais do biodiesel.

Na análise de séries temporais considera-se que a série é composta por componentes, como heterocedasticidades, tendências, ciclos senoidais, sazonalidades e resíduos, sendo esses componentes considerados independentes uns dos outros. Assim, através da Análise de Componentes Independentes [10], é realizada uma etapa de pré-processamento dos dados, que inclui a identificação e separação desses componentes. Além disso, é realizada a verificação de outliers [11], ou seja, valores discrepantes que possam afetar a qualidade e a interpretação dos dados. Essa etapa tem como objetivo obter uma série temporal mais limpa e preparar para as próximas etapas da análise e modelagem, garantindo a confiabilidade e a integridade dos dados utilizados.

Após o pré-processamento, os resíduos da série são isolados, os quais geralmente contêm ruídos e relações não lineares entre as variáveis. Para prever esses resíduos, são empregadas redes neurais e outros modelos de Machine Learning, como Passo Aleatório (do inglês *Random Walk*, RW), Modelo Autoregressivo Integrado de Médias Móveis (do inglês *Autoregressive Integrated Moving Average Model*, ARIMA), MLP [7], Rede Neural Autoregressiva Não-linear com Entradas Exógenas (do inglês *Nonlinear Autoregressive Exogenous Model*, NARX) [12], Rede Neural Recorrente Baseada em Atenção de Dois Estágios (do inglês *Dual-Stage Attention-Based Re-*

current Neural Network, DA-RNN) [13], IMP [14], IDLN [15], N-Linear [16] e Interpolação Neural Hierárquica para Previsão de Séries Temporais (do inglês *Neural Hierarchical Interpolation for Time Series Forecasting*, NHITS) [17].

Para o ajuste das redes neurais, são utilizadas medidas de desempenho como Erro Percentual Absoluto Médio (do inglês *Mean Absolute Percentage Error*, MAPE), Erro Quadrado Logarítmico (do inglês *Squared Logarithmic Error*, SLE), Erro Absoluto Médio (do inglês *Mean Absolute Error*, MAE), Erro Quadrático Médio (do inglês *Mean Squared Error*, MSE) e Raiz do Erro Quadrático Médio (do inglês *Root Mean Squared Error*, RMSE), em conjunto com algoritmos de otimização e ajuste de hiperparâmetros, como GridSearch, RayTune [18] e Optuna [19], para determinar a melhor arquitetura da rede e obter o melhor ajuste possível. Também são aplicados algoritmos para escolha das principais variáveis, como Matriz de Correlação.

O êxito deste trabalho está centrado no pré-processamento dos dados e na definição da melhor arquitetura para a rede neural, com o objetivo de minimizar os erros e o custo computacional envolvido. As séries utilizadas no treinamento da rede neural são obtidas a partir de bancos de dados de domínio público, garantindo a disponibilidade e confiabilidade das informações utilizadas.

1.5 Materiais

- Cluster do Laboratório de Processamento de Sinais (LPS), preferencialmente com Unidade de Processamento Gráfico (do inglês *Graphics Processing Unit*, GPU);
- Python 3.12.4 (Iniciativa de código aberto (do inglês *Open Source Initiative*, OSI)):
 - JupyterLab 4.2.2;
 - Arquitetura de Dispositivo de Computação Unificada (do inglês *Compute Unified Device Architecture*, CUDA) 12.2;
 - Pytorch 2.3.1;
 - Darts 0.30.0.

- Docker ([OSI](#));
- Séries temporais do histórico de preços:
 - Biodiesel nacional e regionais;
 - Óleo de Soja atual e futuro.

1.6 Estrutura do Trabalho

O presente trabalho está organizado em 7 capítulos, conforme a seguir:

Capítulo 1 - Introdução: Reponsável por contextualizar o tema, delimitar o problema, justificar a pesquisa, apresentar os objetivos e o método adotado.

Capítulo 2 - O Biodiesel: Apresenta uma visão geral sobre o biodiesel, sua importância e impacto no mercado de energia, bem como os desafios e oportunidades associados a esse biocombustível.

Capítulo 3 - Redes Neurais Artificiais: Introduz as redes neurais artificiais, explicando seus fundamentos teóricos, arquiteturas mais comuns, funcionamento e o processo de aprendizagem, destacando seu uso em problemas de predição e classificação.

Capítulo 4 - Previsão de Séries Temporais: Discute os conceitos e técnicas de previsão de séries temporais, destacando a importância da análise de dados e modelagem para antecipar eventos futuros.

Capítulo 5 - Descrição do Método: Apresenta o método adotado para desenvolver o modelo de previsão de preços de biodiesel, incluindo a coleta e análise de dados, o pré-processamento das séries temporais, a modelagem com redes neurais e a avaliação do desempenho do modelo.

Capítulo 6 - Resultados : Apresenta os resultados obtidos com a aplicação do modelo de previsão de preços de biodiesel, incluindo a análise dos dados, a comparação com outros modelos e a avaliação da precisão das previsões.

Capítulo 7 - Conclusões e Trabalhos Futuros: Apresenta as conclusões gerais da pesquisa, ressaltando a contribuição do estudo para a área de previsão de preços de biocombustíveis e redes neurais artificiais. Propõe possíveis melhorias no modelo e sugere direções para pesquisas futuras.

Capítulo 2

O Biodiesel

Este capítulo tem como objetivo apresentar uma breve introdução sobre o Biodiesel. Além disso, será descrito seu desenvolvimento no mercado brasileiro.

O Biodiesel é um biocombustível produzido a partir de fontes renováveis como óleos vegetais e gorduras animais que reagem por meio de transesterificação com um álcool primário. Embora o biodiesel possa ser produzido a partir de várias matérias-primas, no Brasil o óleo de soja é a principal fonte de matéria-prima para a produção de biodiesel [20], respondendo por mais de 70% da produção.

No Brasil, a introdução do biodiesel na matriz energética tem sido uma resposta tanto às questões ambientais quanto à necessidade de diversificação das fontes de energia e colocam o Brasil como referência no ramo de biocombustíveis [21].

2.1 Vantagens em Relação ao Óleo Diesel

A substituição do Óleo Diesel pelo Biodiesel apresenta uma série de vantagens. No ponto de vista ambiental, a queima do Biodiesel B100 não libera Óxidos de Enxofre (SO_x), que são agentes causadores da chuva ácida, além de ser biodegradável, se decompondo mais rápido que combustíveis fósseis. No que diz respeito à saúde da população, sua queima gera emissões com baixa concentração de Monóxido de Carbono (CO) [22], que pode causar sérios problemas respiratórios e cardiovasculares. Além disso, pode ser usado em motores a diesel convencionais sem a necessidade de grandes modificações [23].

O biodiesel apresenta uma vantagem significativa em relação à pegada de carbono devido ao seu ciclo de CO_2 relativamente curto. O dióxido de carbono emitido durante a queima do biodiesel é compensado pela fotossíntese das plantas que fornecem as matérias-primas [22], o que contribui para a minimização dos impactos no aquecimento global.

Em contrapartida, sua produção depende do uso de terras agrícolas [23] e de produtos da cesta básica, como o óleo de soja [24], provocando aumento nos preços. O biodiesel também é mais corrosivo e sujeito a degradação do que os combustíveis fósseis, o que pode provocar problemas em motores mais antigos que ainda possuem componentes feitos de materiais como plásticos, metais e borrachas de origem natural [23].

2.2 Produção de Biodiesel no Brasil

A produção de biodiesel no Brasil começou a ganhar impulso no início dos anos 2000, com a implementação do Programa Nacional de Produção e Uso de Biodiesel (PNPB).

Segundo a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), a produção de biodiesel no Brasil alcançou cerca de 7,5 bilhões de litros em 2023 [25], em 2024 espera-se produzir 10 bilhões de litros [26].

2.3 Regulação e Início do Mercado

O PNPB, criado em 2004, foi o marco inicial para a regulamentação do biodiesel no Brasil. Este programa foi essencial para estabelecer as bases legais e econômicas para a produção e comercialização do biodiesel. Em 2005, foi autorizada a mistura de biodiesel ao diesel, até 2%. Em 2008, a mistura se tornou obrigatória, em 2% (B2), aumentando gradualmente ao longo dos anos [20].

Inicialmente, as vendas eram exclusivamente por meio de leilões realizados pela ANP. A partir de 2021, foi autorizada a compra diretamente com os fornecedores

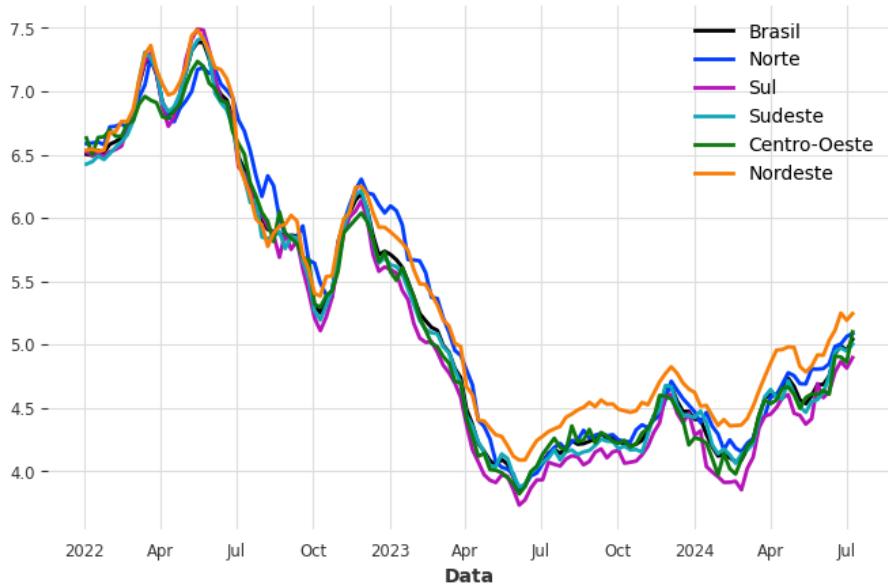


Figura 2.1: Série de Preços do Biodiesel no Brasil (R\$/litro) a partir de 2022

Fonte: Elaborado pelo autor, baseado em ANP [8]

com venda bimestral regulada pela ANP [27].

2.3.1 Percentuais de Adição no Diesel

A mistura obrigatória de biodiesel no diesel tem sido aumentada progressivamente desde sua criação, se adaptando a oferta das matérias primas e ao desenvolvimento das tecnologias de produção.

Devido à crise de oferta de soja em 2020 aliada ao aumento do Índice Nacional de Preços ao Consumidor Amplo (IPCA), o percentual de mistura foi reduzido de 12% para 10% em setembro de 2020 para reduzir os efeitos da crise. Em 2023, o percentual de mistura de biodiesel no diesel volta a subir [28] de acordo com o planejamento da ANP. Além disso, a recente Lei 14.993, de 2024 [29], sancionada pelo presidente da república, estabelece novos percentuais crescentes para a mistura de biodiesel no diesel, como parte da estratégia nacional para promover os chamados “combustíveis do futuro” [30].

Atualmente, especialistas acreditam que a produção é capaz de atender a demanda de uma mistura de 20% de biodiesel no óleo diesel e com potencial para chegar a 25% [31], ainda longe das projeções da ANP, como mostrado na Tabela 2.1.

2.3.1.1 Biodiesel 100% (B100)

A ANP regula o uso experimental do biodiesel B100 através de resoluções específicas. Para utilizar B100 experimentalmente, é necessário obter a prévia anuênciça da ANP, apresentando um projeto detalhado que especifica o objetivo do experimento, os volumes a serem utilizados, a logística envolvida, os riscos e as medidas de mitigação [32].

Tabela 2.1: Percentual de Mistura de Biodiesel no Brasil ao Longo do Tempo

Período	Percentual de Mistura (%)
01/2005 - 12/2007*	2
01/2008 - 06/2008	2
07/2008 - 06/2009	3
07/2009 - 12/2009	4
01/2010 - 06/2014	5
07/2014 - 10/2014	6
11/2014 - 02/2017	7
03/2017 - 02/2018	8
03/2018 - 08/2019	10
09/2019 - 02/2020	11
03/2020 - 08/2020	12
09/2020 - 10/2020**	10
11/2020 - 12/2020	11
01/2021 - 02/2021	12
03/2021 - 04/2021	13
05/2021 - 08/2021**	10
09/2021 - 10/2021	12
11/2021 - 03/2023**	10
04/2023 - 03/2024	12
04/2024 - 02/2025***	13
03/2025 - 02/2026***	15
03/2026 - 02/2027***	16
03/2027 - 02/2028***	17
03/2028 - 02/2029***	18
03/2029 - 02/2030***	19
03/2030***	20

* Percentual de mistura de biodiesel facultativo.

** Queda no percentual obrigatório devido à escassez de soja [28].

*** Projeção.

Fonte: ANP [21] e Lei 14.993, de 2024 [29]

Capítulo 3

Redes Neurais Artificiais

RNA é uma classe de modelos de aprendizado de máquina inspirados no funcionamento do cérebro humano, sendo capazes de aprender e tomar decisões com base em amostras apresentados durante o processo de treinamento [12]. A partir dessas amostras, as redes neurais generalizam o aprendizado para dados previamente desconhecidos, permitindo que realizem inferências sobre novos conjuntos de dados.

As **RNAs** são geralmente estruturadas em três tipos de camadas: camada de entrada, camadas ocultas (ou de processamento) e camada de saída. Essas camadas podem ser organizadas de diversas maneiras, formando topologias específicas que variam de acordo com a aplicação. À medida que os dados de entrada percorrem as camadas da rede, eles são transformados progressivamente por meio de funções de ativação e ajustes sinápticos, que determinam a força das conexões entre os neurônios.

O ajuste dos pesos sinápticos ocorre durante o processo de treinamento, em que algoritmos de otimização, como o método do gradiente descendente, são utilizados para minimizar o erro entre a saída prevista e a saída desejada. A retropropagação do erro é uma técnica amplamente utilizada para este fim, ajustando os pesos de forma iterativa e eficiente, garantindo a convergência para uma solução ótima.

Quando uma rede neural recebe um vetor de entrada x , ela realiza cálculos que geram uma saída y , de acordo com a Equação 3.1:

$$y = f(Wx + b) \quad (3.1)$$

Nesta equação, W representa a matriz de pesos sinápticos da rede, b corresponde ao vetor de viés ajustado durante o treinamento, e f é a função de ativação, responsável por introduzir não linearidades no modelo. A saída y é, portanto, o resultado final do processamento, obtido após a aplicação dessas transformações.

3.1 Perceptron de Rosenblatt

O Perceptron, introduzido por Rosenblatt, 1958 [33], foi um dos primeiros modelos de rede neural e representa um marco no desenvolvimento da inteligência artificial. Ele foi concebido para realizar tarefas de classificação linear, sendo capaz de separar classes de dados que são linearmente separáveis. O modelo Perceptron é composto por um único neurônio, que processa um conjunto de entradas, cada uma associada a um peso sináptico, e produz uma saída binária com base em uma função de ativação [33].

O funcionamento do Perceptron é descrito da seguinte forma:

$$y = \begin{cases} 1, & \text{se } \sum_{i=1}^n w_i x_i + b > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (3.2)$$

Sendo x_i as entradas do modelo, w_i os pesos sinápticos associados a cada entrada, b o viés (bias) do modelo, ajustável durante o treinamento, e y a saída binária do Perceptron.

3.1.1 Limitações e Extensões

Embora o Perceptron tenha sido um avanço significativo, ele é limitado a problemas que podem ser separados por uma linha reta (ou hiperplano no caso de dimensões superiores), ou seja, dados que são linearmente separáveis. Em 1969, Minsky e Papert [34] demonstraram que o Perceptron era incapaz de resolver problemas como o XOR, que não são linearmente separáveis. Essa limitação levou à

introdução de arquiteturas mais complexas, como o [MLP](#), que supera essa deficiência ao incorporar camadas ocultas e ativação não linear [\[34\]](#).

O Perceptron de Rosenblatt foi fundamental para o desenvolvimento posterior de redes neurais e permanece como um conceito básico no aprendizado de máquina.

O treinamento do Perceptron utiliza uma regra de aprendizado simples e eficiente, conhecida como regra do Perceptron [\[33\]](#). Para cada amostra de treinamento, os pesos são ajustados de acordo com o erro entre a saída prevista e a saída desejada:

$$w_i(t+1) = w_i(t) - \eta \frac{\partial E}{\partial w_i(t)} \quad (3.3)$$

Sendo $w_i(t)$ o peso sináptico do neurônio i no instante t , η a taxa de aprendizado, que controla o tamanho do passo de ajuste $e \frac{\partial E}{\partial w_i(t)}$ o gradiente do erro em relação ao peso.

Essa regra de atualização ajusta os pesos para reduzir o erro de classificação em cada iteração.

3.2 Treinamento de Redes Neurais

O treinamento das redes neurais consiste na apresentação dos dados de entrada, acompanhados das respectivas saídas desejadas (supervisionado), ou simplesmente dos dados de entrada no caso de aprendizagem não supervisionada. Durante o treinamento, o objetivo principal é minimizar a função de perda, que avalia as saídas previstas pela rede com base nas saídas reais.

O algoritmo mais comum para esse ajuste é o gradiente descendente, que atualiza os pesos sinápticos conforme a direção de maior declínio do erro. Métodos como o gradiente estocástico, Adam e RMSprop são variações amplamente utilizadas, aprimorando o desempenho em diferentes cenários.

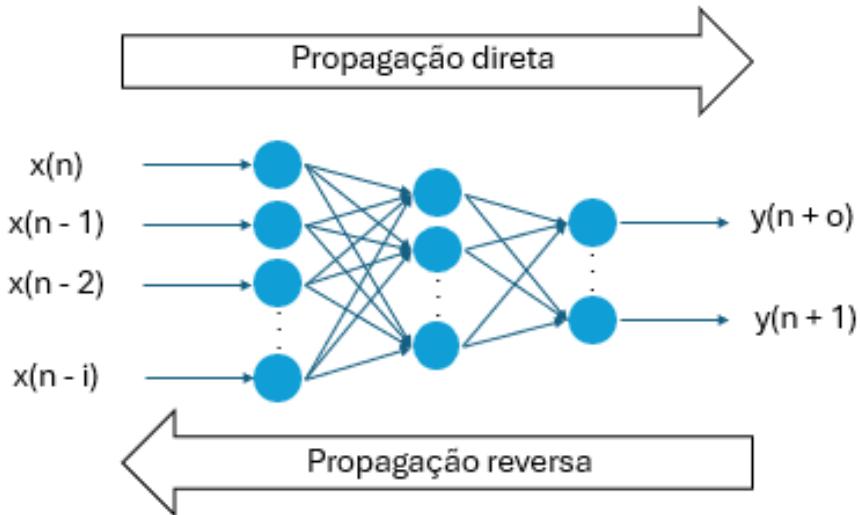


Figura 3.1: Diagrama da **MLP** nas fases de retropropagação

Fonte: Elaborado pelo autor, baseado em [12, 35]

3.2.1 Retropropagação

A retropropagação é o algoritmo mais utilizado para o treinamento de redes neurais supervisionadas. Introduzido por Rumelhart et al, 1986 [35], esse algoritmo ajusta os pesos sinápticos da rede para minimizar o erro entre a saída prevista e a saída desejada, utilizando o gradiente descendente.

A retropropagação ocorre em duas fases principais, mostradas na Figura 3.1 e detalhadas a seguir:

- **Fase de propagação direta:** Os dados de entrada são propagados pela rede, camada por camada, até que a saída seja calculada. Esta saída é então comparada com o valor esperado, e o erro é calculado com base em uma função de perda, como o **MSE** ou a entropia cruzada.
- **Fase de propagação reversa:** O erro calculado é retropropagado da camada de saída para as camadas anteriores, ajustando os pesos sinápticos de acordo com o gradiente do erro em relação a cada peso. O objetivo é reduzir o erro a cada iteração, tornando o modelo mais preciso.

A atualização dos pesos é dada pela seguinte equação:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial F_o}{\partial w_{ij}(t)} \quad (3.4)$$

Sendo $w_{ij}(t)$ o peso sináptico entre o neurônio i e o neurônio j no instante t , η a taxa de aprendizado, que controla o tamanho do passo de ajuste e $\frac{\partial F_o}{\partial w_{ij}(t)}$ o gradiente da função objetivo em relação ao peso.

A retropropagação é eficaz em problemas complexos, permitindo o ajuste eficiente dos parâmetros, mas pode enfrentar desafios como o problema do gradiente desaparecendo, que pode ser mitigado com técnicas como funções de ativação como a Unidade Linear Retificada (do inglês *Rectified Linear Unit*, **ReLU**) e o uso de redes profundas com inicializações adequadas de peso.

3.3 Funções de Ativação

As funções de ativação são elementos essenciais nas redes neurais, pois introduzem não linearidade no modelo, permitindo que ele aprenda padrões complexos. Entre as funções de ativação mais utilizadas estão:

- **Sigmoide:**

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.5)$$

É utilizada principalmente em saídas binárias, transformando valores reais em probabilidades entre 0 e 1.

- **ReLU:**

$$f(x) = \max(0, x) \quad (3.6)$$

Amplamente utilizada em redes profundas por sua simplicidade e eficiência computacional.

- **Tanh:**

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.7)$$

A função tangente hiperbólica, semelhante à sigmoide, mas com saída no intervalo $[-1, 1]$.

- **Softmax:**

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (3.8)$$

Usada em problemas de classificação multiclasse, transformando um vetor de valores em probabilidades.

3.4 Arquiteturas

As redes neurais podem assumir diferentes arquiteturas, de acordo com o problema que se busca resolver. A seguir, algumas das principais arquiteturas:

- **Rede Neural Autoregressiva Não-linear com Entradas Exógenas** (do inglês *Nonlinear Autoregressive Exogenous Model*, **NARX**): A rede **NARX** é uma arquitetura específica de redes neurais utilizada para modelagem de séries temporais, onde a saída depende não apenas das entradas passadas, mas também de saídas anteriores. Ela é amplamente aplicada em sistemas dinâmicos.
- **Rede Neural Recorrente** (do inglês *Recurrent Neural Network*, **RNN**): As **RNNs** possuem conexões entre os neurônios que permitem o processamento sequencial de dados, tornando-as adequadas para tarefas como modelagem de séries temporais, reconhecimento de fala e tradução automática.
- **Perceptron Multicamadas** (do inglês *Multilayer Perceptron*, **MLP**): O **MLP** é uma das arquiteturas mais clássicas e simples de redes neurais, composta por múltiplas camadas de neurônios com conexões completas entre si. É utilizado principalmente para problemas de classificação e regressão.
- **Redes Morfológicas**: As redes neurais morfológicas utilizam operadores matemáticos morfológicos no lugar de funções de ativação tradicionais. Elas são especialmente úteis para tarefas de processamento de imagens, onde operações como dilatação e erosão são comuns.

- **Redes Profundas:** As Redes Neurais Profundas são caracterizadas por terem múltiplas camadas ocultas, permitindo a aprendizagem de representações mais complexas e abstratas dos dados. Elas são amplamente utilizadas em áreas como reconhecimento de imagens e processamento de linguagem natural. A profundidade da rede permite a extração de características de alto nível, o que é fundamental para o sucesso em tarefas como a classificação de imagens e tradução automática.

Capítulo 4

Previsão de Séries Temporais

Séries temporais, por definição, são conjuntos ordenados de observações, onde cada observação é o valor mensurado de uma variável de interesse em intervalos de tempo específicos [1]. Essas observações são coletadas sequencialmente ao longo do tempo, e a análise de séries temporais envolve a modelagem e previsão dessas observações com base em padrões históricos. A estrutura fundamental de uma série temporal pode ser representada como:

$$X = \{x_t \in \mathbb{R} \mid t = 1, 2, 3, \dots, N\} \quad (4.1)$$

onde x_t representa o valor observado no tempo t , e N é o número total de observações. As séries temporais são amplamente utilizadas para identificar tendências, sazonalidades e ciclos senoidais nos dados, bem como para fazer previsões futuras com base nesses padrões identificados. A análise de séries temporais pode envolver diversas técnicas estatísticas e de aprendizado de máquina, como modelos ARIMA, modelos de suavização exponencial, e métodos baseados em redes neurais. O objetivo é entender a dinâmica dos dados ao longo do tempo e aplicar esse entendimento para prever comportamentos futuros com o maior grau de precisão possível.

4.1 Métodos de Decomposição e Normalização

4.1.1 Decomposição dos Dados

A decomposição é uma técnica fundamental no processamento de séries temporais, usada para analisar e entender as diferentes componentes que constituem os dados. A decomposição permite separar uma série temporal em componentes distintas [10], facilitando a análise e a modelagem. A compreensão dessas componentes pode melhorar a precisão dos modelos preditivos.

A série temporal pode ser definida como:

$$x_t = T_t + C_t + S_t + R_t \quad (4.2)$$

onde x_t representa o valor observado da série temporal no tempo t . A componente T_t refere-se à tendência, que captura a direção geral ou o padrão de longo prazo nos dados. A tendência pode ser crescente, decrescente ou constante ao longo do tempo e mostra o padrão de longo prazo dos dados. A componente C_t descreve os ciclos senoidais, que se refere às variações que ocorrem em períodos de tempo mais longos, como ciclos econômicos ou de mercado que não se repetem em intervalos fixos, mas são identificáveis ao longo de vários anos. A componente S_t corresponde à sazonalidade, que representa variações regulares e previsíveis que ocorrem em intervalos de tempo específicos e fixos, como meses ou estações do ano, refletindo padrões que se repetem regularmente e são previsíveis. Por fim, R_t é o resíduo, também conhecido como erro ou ruído, que é a parte da série temporal que não é explicada pelas outras componentes. Inclui variações aleatórias e imprevisíveis que não podem ser atribuídas à tendência, ciclos senoidais ou sazonais.

Compreender essas componentes é essencial para uma análise eficaz de séries temporais, pois permite identificar padrões, prever tendências futuras e ajustar modelos para melhorar sua precisão. Este capítulo explorará detalhadamente esses métodos, suas vantagens e limitações, proporcionando uma base sólida para a aplicação prática da decomposição em séries temporais.

4.1.1.1 Remoção de Tendência Linear

A remoção de tendência linear (T_t) é uma técnica utilizada para eliminar a componente linear de um conjunto de dados, permitindo que outros padrões ou características dos dados se destaquem mais claramente. Isso é particularmente útil em séries temporais, onde tendências podem mascarar variações sazonais ou outros comportamentos.

Um dos métodos aplicados para remover a tendência linear é a regressão linear. Nesse processo, ajusta-se um modelo linear aos dados, representado pela equação:

$$T = \alpha + \beta X \quad (4.3)$$

onde T é o valor observado, X é a variável independente (por exemplo, o tempo), α é o coeficiente linear, e β é o coeficiente angular que representa a inclinação da tendência. Após ajustar o modelo, a tendência linear é subtraída dos dados originais, resultando em uma série de dados sem tendência, que pode ser analisada ou processada de forma mais eficaz.

A remoção da tendência é útil em diversas aplicações, como na análise econômica e na meteorologia, onde se busca entender variações mais sutis sem a influência de uma tendência subjacente de longo prazo.

4.1.1.2 Remoção de Ciclos Senoidais

A remoção de ciclos senoidais (C_t) é uma técnica utilizada para eliminar a componente cíclica de um conjunto de dados, eliminando flutuações que ocorrem em intervalos irregulares e sem uma frequência fixa ao longo do tempo. Essas oscilações são causadas por eventos externos. Para remover a componente cíclica, são utilizadas técnicas como a decomposição clássica e Análise Espectral [3].

4.1.1.3 Remoção de Sazonalidade

A remoção da sazonalidade (S_t) consiste em remover flutuações regulares e repetitivas nos dados, com padrões que se repetem em determinados intervalos,

como mensal, trimestral e anual [3]. A sazonalidade é normalmente encontrada em combustíveis de acordo com os ciclos de oferta de matéria prima, variação da demanda e outras séries. Esta componente pode ser explicada como a variação do comportamento do consumidor e da oferta ao longo do ano.

4.1.1.4 Estimação dos Resíduos

A estimação dos resíduos (R_t) é uma etapa complexa na previsão de séries temporais. Essa componente é composta por variações não capturadas pelas remoções de tendência, ciclos senoidais e sazonalidade [3]. Essa componente necessita de métodos mais avançados para estimação, para isso são usados modelos estatísticos como o **ARIMA** e os modelos neurais.

4.1.2 Normalização dos Dados

A normalização é um processo fundamental no pré-processamento dos dados. Os métodos de normalização ajustam os valores de um conjunto de dados para que eles sigam uma escala ou distribuição específica, o que pode melhorar o desempenho de algoritmos e facilitar a interpretação dos resultados. Este capítulo mostrará alguns dos métodos de normalização disponíveis na literatura.

4.1.2.1 Escalonamento Padrão

O escalonamento padrão é uma técnica de normalização que ajusta os valores dos dados para que tenham uma média de 0 e um desvio padrão de 1. Essa abordagem é útil quando se deseja centralizar os dados e ajustar sua dispersão, o que pode melhorar o desempenho de algoritmos de aprendizado de máquina que assumem que os dados são distribuídos de forma normal [36]. O processo de escalonamento é realizado pela Equação 4.4, onde $X_{escalado}$ representa o valor escalonado, X é o valor original, μ é a média do conjunto de dados, e σ é o desvio padrão.

$$X_{escalado} = \frac{X - \mu}{\sigma} \quad (4.4)$$

Essa técnica é particularmente útil em algoritmos de aprendizado de máquina como máquinas de vetores de suporte e regressão logística, onde a escala dos da-

dos pode impactar a eficácia do modelo e a velocidade de convergência durante o treinamento.

4.1.2.2 Normalização Min-Max

A normalização Min-Max é uma técnica utilizada para ajustar os valores dos dados para um intervalo específico, geralmente entre 0 e 1. Essa abordagem é útil quando se deseja manter a relação proporcional entre os valores de entrada, mas reduzir a variação dos dados para um intervalo controlado [37]. O processo de normalização é feito pela Equação 4.5, onde $X_{escalado}$ representa o valor normalizado, X é o valor original, e $\min(X)$ e $\max(X)$ são os valores mínimo e máximo do conjunto de dados, respectivamente.

$$X_{escalado} = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (4.5)$$

Essa técnica é útil em métodos de aprendizado de máquina onde a escala dos dados pode influenciar o desempenho do modelo, como em redes neurais, onde variáveis em diferentes escalas podem levar a uma convergência mais lenta durante o treinamento.

4.2 Janelas Temporais

Na maioria dos casos, as séries temporais não são estacionárias em todo o período, fatores externos podem alterar o comportamento da série ao longo do tempo, o que torna a extração de características uma tarefa complexa. De forma a minimizar o impacto dessas variações, é adotada a abordagem de janelas temporais.

A janela temporal é um subconjunto de observações da série temporal defasadas no tempo, possibilitando a extração de características necessárias para a previsão do fenômeno gerador. Diferentes tipos de janelas são utilizados, como TAKENS, SAVIT e GREEN, PI e PETERSON, TANAKA. Essas técnicas buscam identificar os retardos temporais mais relevantes para a previsão dos próximos pontos, definidas pela Equação 4.6.

$$\hat{y}_{t:t+h} = f(X(t, w)) + r_t \quad (4.6)$$

onde $\hat{y}_{t:t+h}$ são os valores, para o intervalo de tempo $[t+1 : t+h]$, f é a função de previsão, e w é o tamanho da janela temporal. A função f representa o mapeamento entre os pontos do passado $X(t, w)$. O termo r_t representa um termo de ruído, que tende a diminuir à medida que w aumenta [3]. Desta forma, a escolha adequada do formato e do tamanho da janela temporal é uma etapa importante do processo de previsão, pois influencia diretamente na qualidade das previsões.

4.2.1 Janela Deslizante

Nesta abordagem, a série temporal é subdividida em janelas de comprimento fixo (J , com $J < N$). As janelas são deslocadas ao longo da série com um horizonte de previsão (H) que define o intervalo entre o início de uma janela e a próxima, conforme ilustrado pela Equação 4.7.

$$j_i = \{x_t \mid t = i, i+1, i+2, \dots, i+J-1\} \text{ para } i = 1, 1+H, 1+2H, \dots, N-J+1 \quad (4.7)$$

Nesta fórmula, j_i representa a i -ésima janela, que contém J observações consecutivas da série temporal, começando na posição i . O parâmetro H define o deslocamento entre o início de uma janela e a próxima, e as janelas são geradas até o final da série, garantindo que cada janela tenha exatamente J observações.

4.2.2 Janela de Takens

Nesta abordagem, a série temporal é subdividida em janelas com comprimento fixo (J , com $J < N$). Cada janela é formada a partir de um atraso de tempo τ entre as observações, conforme ilustrado pela Equação 4.8.

$$t_i = \{x_{i+k\tau} \mid k = 0, 1, 2, \dots, J-1\} \text{ para } i = 1, 2, \dots, N - J\tau \quad (4.8)$$

Nesta fórmula, t_i representa a i -ésima janela, que contém J observações da série temporal, começando na posição i e com um atraso de τ entre as observações subsequentes. O parâmetro i define a posição inicial da janela e é variado até garantir que cada janela tenha exatamente J observações, com a condição $i \leq N - J\tau$.

4.3 Medidas de Avaliação do Modelo

As medidas de avaliação de modelo são utilizadas para quantificar a precisão e a qualidade das previsões feitas por um modelo em relação aos valores reais observados. Essas medidas são essenciais para comparar o desempenho de diferentes modelos e para entender como eles se comportam em diferentes condições de dados.

4.3.1 Erro Percentual Absoluto Médio (do inglês *Mean Absolute Percentage Error*, **MAPE**)

O **MAPE** é uma medida que expressa o erro médio absoluto entre os valores previstos (p_t) e os valores reais (x_t) em termos percentuais [38]. Ele é definido pela Equação 4.9, onde N é o número total de observações. O **MAPE** é útil em contextos onde a interpretação do erro em termos de percentual é mais relevante.

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{x_t - p_t}{x_t} \right| \quad (4.9)$$

Uma vantagem do **MAPE** é que ele é fácil de interpretar e comparar entre diferentes modelos ou datasets. No entanto, ele pode ser influenciado por valores reais (x_t) próximos de zero, o que pode resultar em valores extremamente altos para o erro percentual [39].

4.3.2 Erro Quadrático Médio (do inglês *Mean Squared Error*, **MSE**)

O **MSE** mede o erro médio ao quadrado entre os valores previstos (p_t) e os valores reais (x_t). A fórmula é apresentada na Equação 4.10. O **MSE** é amplamente utilizado porque penaliza erros maiores de forma mais significativa, uma vez que os erros são elevados ao quadrado [40].

$$MSE = \frac{1}{N} \sum_{t=1}^N (x_t - p_t)^2 \quad (4.10)$$

Essa medida é útil quando se deseja enfatizar grandes desvios nas previsões. No entanto, ela também pode ser sensível a outliers, pois estes contribuem de forma

desproporcional para o valor final do **MSE**.

4.3.3 Raiz do Erro Quadrático Médio (do inglês *Root Mean Squared Error*, **RMSE**)

O **RMSE** é a raiz quadrada do **MSE**, conforme a Equação 4.11. Ele mantém as propriedades do **MSE**, mas é expresso na mesma unidade dos valores originais, o que facilita a interpretação [40].

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - p_t)^2} \quad (4.11)$$

O **RMSE** é preferido em situações onde grandes erros são particularmente indesejáveis, e sua interpretação direta em termos das unidades dos dados originais é uma vantagem adicional.

4.3.4 Erro Absoluto Médio (do inglês *Mean Absolute Error*, **MAE**)

O **MAE** é a média dos erros absolutos entre as previsões (p_t) e os valores reais (x_t), conforme mostrado na Equação 4.12. O **MAE** é uma medida intuitiva que mede o erro médio sem considerar a direção do erro (positivo ou negativo) [40].

$$MAE = \frac{1}{N} \sum_{t=1}^N |x_t - p_t| \quad (4.12)$$

Diferente do **MSE** e **RMSE**, o **MAE** não penaliza erros grandes de forma desproporcional, o que o torna uma escolha adequada quando se deseja uma métrica de erro mais equilibrada.

4.3.5 Estatística U de Theil (**THEIL**)

O índice de Theil é uma medida utilizada para avaliar a precisão das previsões de um modelo em comparação com um modelo de referência. Ele é particularmente útil para medir a qualidade das previsões em séries temporais e pode ser dividido em dois componentes: U_1 e U_2 [41].

4.3.5.1 U_1 de Theil

A estatística U_1 do Índice de Theil avalia a precisão das previsões comparadas com um modelo de referência. A fórmula é dada por:

$$U_1 = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - p_t)^2}}{\sqrt{\frac{1}{N} \sum_{t=1}^N x_t^2} + \sqrt{\frac{1}{N} \sum_{t=1}^N p_t^2}} \quad (4.13)$$

Quanto mais precisas as previsões, menor o valor da estatística U_1 . A estatística U_1 é limitada entre 0 e 1, com valores mais próximos de 0 indicando maior precisão de previsão [41].

4.3.5.2 U_2 de Theil

A estatística U_2 do Índice de Theil avalia o desempenho do modelo em relação ao modelo ingênuo [41]. A fórmula é dada pela equação:

$$U_2 = \frac{\sqrt{\sum_{t=1}^{N-1} \left(\frac{P_{t+1}-X_{t+1}}{x_t} \right)^2}}{\sqrt{\sum_{t=1}^{N-1} \left(\frac{X_{t+1}-X_t}{x_t} \right)^2}} \quad (4.14)$$

Os valores de U_2 são classificados em três intervalos:

- $U_2 \in [0, 1)$, o modelo é melhor que o modelo ingênuo
- $U_2 \in (1, \infty)$, o modelo é pior que o modelo ingênuo
- $U_2 = 1$, o modelo é igual ao modelo ingênuo

4.4 Coeficiente de Hurst

O Coeficiente de Hurst, denotado por H , é uma medida estatística utilizada para analisar a long-range dependence (dependência de longo alcance) em séries temporais. Originalmente desenvolvido por Harold Edwin Hurst na década de 1950, o coeficiente é uma ferramenta fundamental na análise de dados financeiros, climáticos e muitos outros campos onde o comportamento temporal é relevante.

4.4.1 Definição e Cálculo

O coeficiente de Hurst varia entre 0 e 1 e fornece informações sobre a tendência e a persistência da série temporal [42]:

- **H < 0.5:** A série temporal exibe um comportamento de média reversão, indicando que tendências altas são seguidas por tendências baixas e vice-versa.
- **H = 0.5:** A série é considerada um Passeio Aleatório, onde não há dependência temporal.
- **H > 0.5:** A série exibe um comportamento de persistência, onde altas tendências tendem a ser seguidas por altas e baixas por baixas.

Uma forma de calcular o Coeficiente de Hurst é por meio da análise de resíduo de variância (R/S). O procedimento básico envolve os seguintes passos:

1. Calcular a média da série temporal $X(t)$:

$$\bar{X} = \frac{1}{N} \sum_{t=1}^N X(t)$$

onde N é o número total de observações.

2. Calcular os desvios acumulados a partir da média:

$$Y(t) = \sum_{i=1}^t (X(i) - \bar{X})$$

3. Calcular o desvio padrão dos desvios acumulados em várias escalas n :

$$R(n) = \max(Y(t)) - \min(Y(t))$$

$$S(n) = \sqrt{\frac{1}{n} \sum_{t=1}^n (X(t) - \bar{X})^2}$$

4. Calcular o R/S:

$$R/S = \frac{R(n)}{S(n)} \quad (4.15)$$

Capítulo 5

Descrição do Método

O método aplicado a este trabalho pode ser dividido em 5 etapas principais, se dividindo entre desenvolvimento e teste. Inicia-se pela coleta de dados, a partir de fontes abertas e confiáveis, seguido pela etapa de divisão do conjunto de dados em desenvolvimento e teste, etapas de pré-processamento, onde características principais são extraídas antes do treinamento de um modelo neural e por fim, a avaliação dos modelos no conjunto de desenvolvimento, para a escolha de hiperparâmetros e no conjunto de teste, para avaliar a capacidade de generalização. O diagrama de processamento do método é mostrado na Figura 5.1 e o método detalhado nas seções a seguir.

5.1 Coleta de Dados

A primeira etapa deste trabalho é a coleta de dados, que é uma etapa importante para a construção de um modelo de previsão. A qualidade dos dados coletados influencia diretamente a qualidade do modelo. Para este trabalho, os dados foram coletados de fontes públicas e confiáveis, como a ANP (em conformidade com a Lei

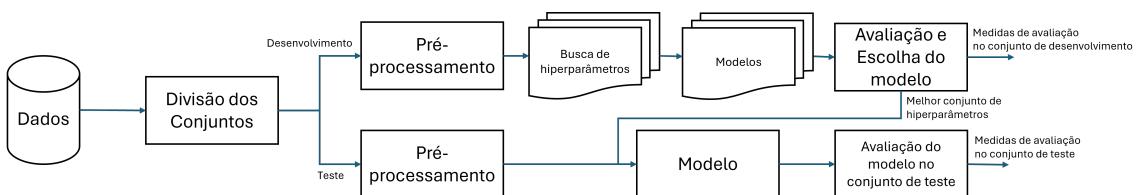


Figura 5.1: Fluxo de Processamento

de Acesso a Informação [43]), o Banco Mundial e o Index Mundi. O intervalo de tempo escolhido para a análise compreende o período entre janeiro de 2022 e julho de 2023, permitindo a avaliação da série em 2 períodos sazonais para uma previsão no curto prazo.

Além disso, serão usados como covariáveis passadas as séries de preços do Óleo de Soja Internacional e do Óleo de Soja no Mercado Futuro. Para possibilitar uma análise de um período de tempo maior, as séries de preços do Óleo de Soja Internacional e do Óleo de Soja no Mercado Futuro foram coletadas desde 1960 e 1994, respectivamente.

A seleção das séries temporais foi baseada na disponibilidade de dados históricos e na correlação observada entre elas durante o período de análise, conforme ilustrado na Figura 5.2. A alta correlação entre as séries regionais de preços do biodiesel sugere que a inclusão dessas séries no modelo não contribui com informações adicionais relevantes.

De maneira semelhante, as covariáveis foram escolhidas com base em critérios de correlação. A série de preços do óleo de soja internacional demonstra uma alta correlação com as séries de preços do biodiesel, tornando-a uma candidata promissora para o treinamento do modelo devido à maior disponibilidade de dados. Por outro lado, a série de preços do óleo de soja no mercado futuro também apresenta alta correlação com as séries de preços do biodiesel. No entanto, como essa série tem uma correlação perfeita com a série de preços do óleo de soja internacional, sua inclusão no modelo não adiciona valor informativo adicional.

Portanto, as séries escolhidas para treinamento do modelo são: Preço do Biodiesel Nacional e Preço do Óleo de Soja Internacional.

5.2 Divisão do Conjunto de Dados

Para uma modelagem e avaliação mais eficaz do modelo, os dados foram divididos em dois conjuntos: desenvolvimento e teste, conforme mostrado na Figura 5.3.

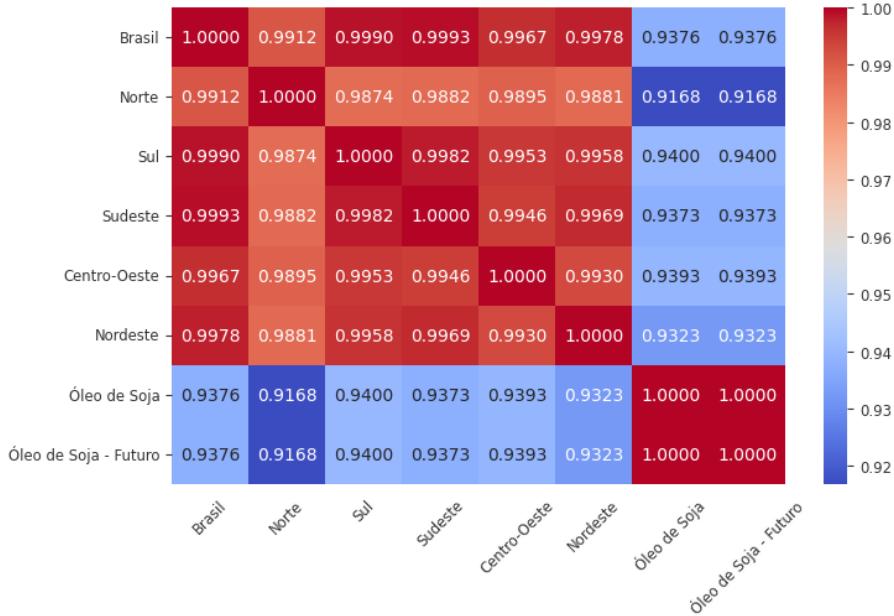


Figura 5.2: Matriz de Correlação entre as Séries Candidatas (01/2022 - 07/2023)

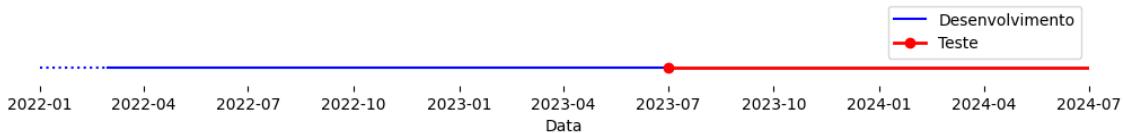


Figura 5.3: Linha do Tempo da Divisão dos Conjuntos

O período até julho de 2023 foi alocado para o conjunto de desenvolvimento, que é utilizado para ajustar os hiperparâmetros do modelo e validar sua eficácia. Esse conjunto inclui dados suficientes para observar a dinâmica sazonal e garantir a capacidade de generalização do modelo. Os dados restantes são usados como conjunto de teste, fornecendo uma avaliação final não tendenciosa do modelo com dados que não foram vistos anteriormente. Este modelo de divisão é uma prática recomendada pela literatura para garantir confiabilidade e robustez do modelo de previsão [44].

5.3 Pré-processamento dos Dados

Após a divisão dos conjuntos, é aplicado o processo de pré-processamento. Que pode ser subdividido em 4 partes principais: Processo de descorrelação, divisão

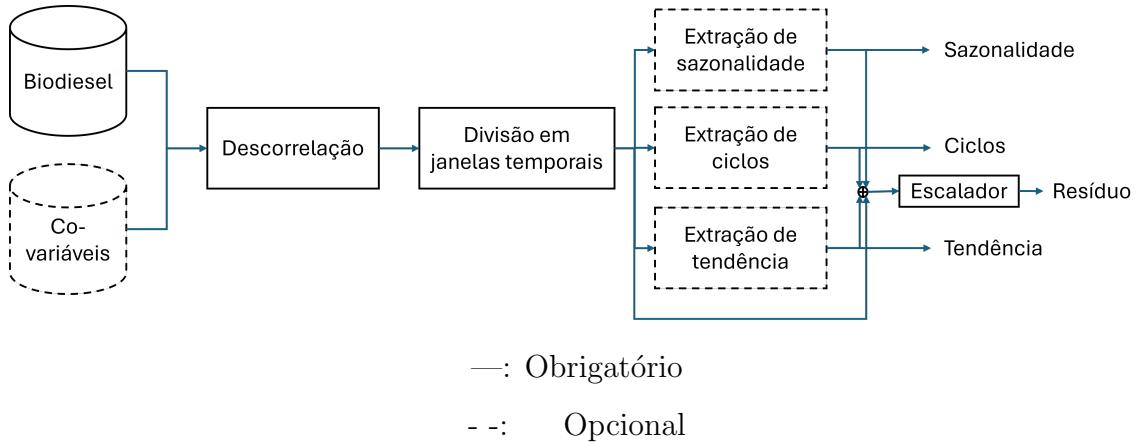


Figura 5.4: Fluxo de Pré-Processamento

em janelas temporais, extração de características e escalamento do resíduo, como mostrado na Figura 5.4.

Inicialmente, o conjunto de desenvolvimento foi submetido a uma análise para identificar padrões de sazonalidade e ciclos senoidais. Caso sejam identificados, o tamanho da janela temporal deve ser escolhido de modo a capturar pelo menos 2 observações de cada padrão.

A verificação da sazonalidade foi realizada por meio da autocorrelação da série temporal. No entanto, a análise não revelou ciclos sazonais completos, conforme ilustrado na Figura 5.5. Além disso, para identificar ciclos senoidais, foi empregada a Transformada Rápida de Fourier (FFT). No entanto, a análise de FFT também não conseguiu identificar ciclos senoidais significativos na série temporal, como mostrado na Figura 5.6. Esses resultados indicam a ausência de padrões sazonais e cíclicos evidentes nos dados analisados.

A verificação da persistência foi realizada utilizando o Coeficiente de Hurst, conforme descrito na Equação 4.15. Essa abordagem permite analisar a série temporal para verificar se ela exibe um comportamento de persistência, o que significa que os valores futuros podem ser previstos a partir dos dados passados.

Com base nessa análise, obteve-se um Coeficiente de Hurst igual a 0,73, constatou-se que a série apresenta um comportamento de persistência, o que implica uma

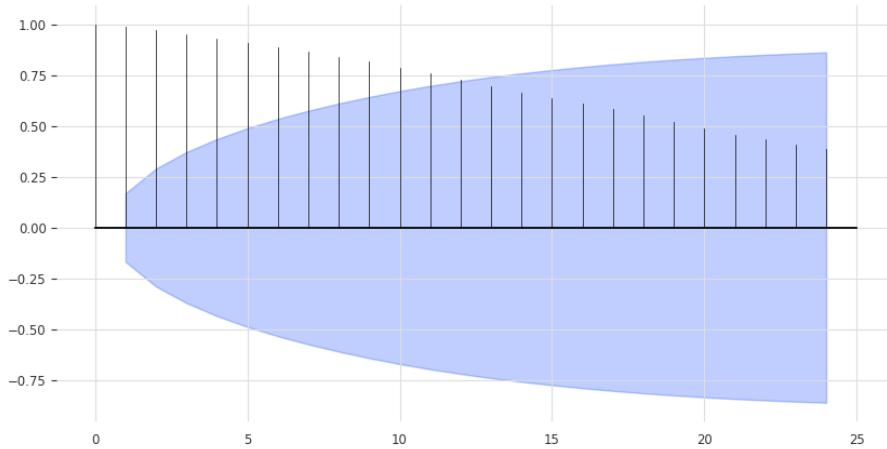


Figura 5.5: Autocorrelação da Série de Preços do Biodiesel

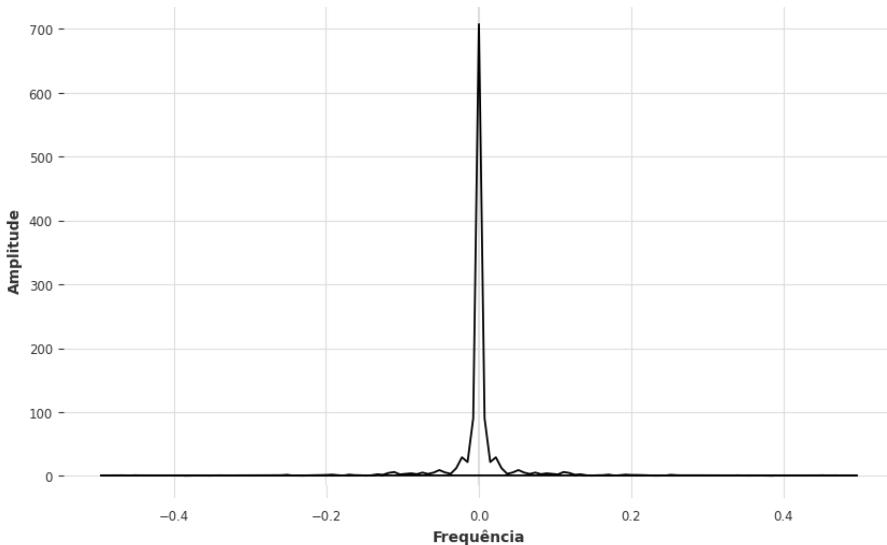


Figura 5.6: Frequências de Fourier da Série de Preços do Biodiesel

correlação positiva entre os valores atuais e os valores passados. Em outras palavras, quando a série tende a aumentar, as tendências passadas indicam uma probabilidade de continuar subindo, e quando tende a cair, as tendências passadas sugerem uma continuidade desse movimento. Isso reforça a possibilidade de previsão a partir de dados históricos, validando a hipótese de dependência de longo prazo na série temporal.

Os dados selecionados, o preço do biodiesel nacional e do óleo de soja internacional, foram submetidos a um processo de descorrelação para remover dependências lineares entre as variáveis, como mostrado na matriz de correlação da Figura 5.2,

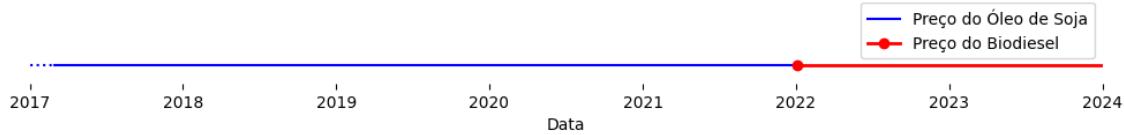


Figura 5.7: Linha do Tempo dos Períodos de Disponibilidade das Séries de Biodiesel e Óleo de Soja

removendo os dados da série do óleo de soja internacional no período em que há dados das duas séries, como mostrado na Figura 5.7. Em seguida, os dados foram divididos em dois experimentos distintos. O primeiro experimento considera apenas o preço do biodiesel nacional, enquanto o segundo experimento incorpora tanto o preço do biodiesel quanto o preço do óleo de soja, permitindo a análise do impacto combinado dessas variáveis nos modelos de previsão.

Posteriormente, os dados foram segmentados em janelas temporais, aplicando duas abordagens previamente descritas na fundamentação teórica: (1) Janela Deslizante com sobreposição, que permite capturar a continuidade temporal ao criar séries sobrepostas; e (2) Janela de Takens, que reconstrói o espaço de estados da série temporal a partir de um único sinal, utilizando atrasos específicos para capturar as dinâmicas subjacentes.

Foram aplicadas duas abordagens distintas para a normalização e decomposição dos dados. Na primeira abordagem, os dados brutos foram normalizados diretamente utilizando a Equação 4.5. Na segunda abordagem, os dados foram primeiramente processados para remover a tendência linear, ajustada pela Equação 4.3, resultando em dados residuais. Esses dados residuais foram então normalizados utilizando novamente a Equação 4.5.

5.4 Treinamento e Avaliação dos Modelos

Em seguida, foi realizada a validação cruzada para avaliar o desempenho do modelo de forma mais robusta, assegurando sua eficácia e confiabilidade em diferentes subconjuntos de dados.

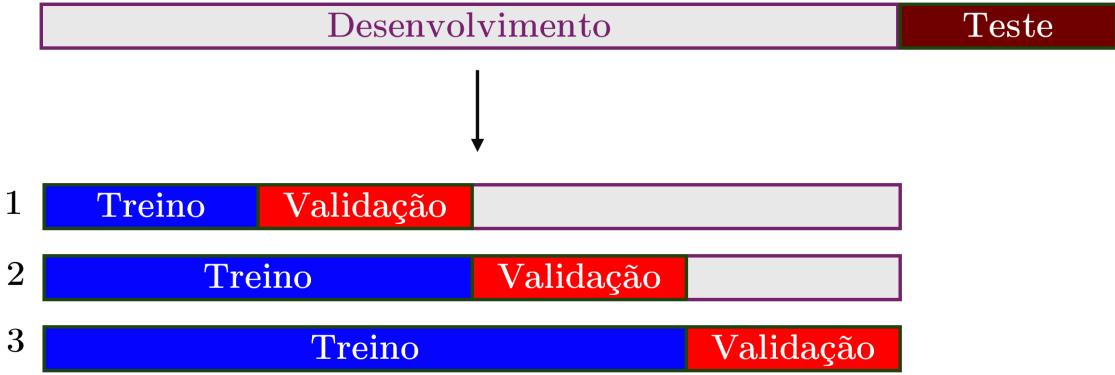


Figura 5.8: Validação Cruzada com 3 dobras

Fonte: Elaborado pelo autor, baseado em [2]

O conjunto de desenvolvimento é dividido em 2 subconjuntos. O conjunto de treinamento, usado para apresentar amostras de pares entrada e saída para o modelo e o conjunto de validação usado para avaliar o modelo em dados que não foram usados para o ajuste de parâmetros, evitando sobreajuste e garantindo a capacidade de generalização. Os conjuntos são divididos com tamanhos diferentes em cada dobra da validação cruzada, conforme mostrado na Figura 5.8.

Após a subdivisão, os modelos serão treinados usando o conjunto de treinamento e seu desempenho será avaliado pelas medidas de desempenho: MAPE, SLE, MSE, RMSE, U_1 e U_2 de cada conjunto de validação e com os resultados serão calculados média e desvio padrão para cada medida. O melhor conjunto de hiperparâmetros será escolhido pelo algoritmo Optuna [19] e o ajuste fino será feito pela GridSearch.

Após a definição dos hiperparâmetros com base nas medidas de desempenho, o modelo será treinado usando todo o conjunto de desenvolvimento e, em seguida, avaliado no conjunto de teste. Isso permitirá verificar sua capacidade de generalização para dados que não foram vistos anteriormente.

Capítulo 6

Resultados

Neste capítulo serão apresentados os resultados do modelo, avaliado pelas medidas de desempenho descritas anteriormente, junto com os gráficos da previsão do conjunto de teste.

Os resultados serão apresentados para cada modelo, com os dois conjuntos de dados disponíveis: Biodiesel Nacional e Óleo de Soja e as duas abordagens de pré-processamento: remoção de tendência linear e normalização min-max. Os diferentes tipos de janelamento: Janela deslizante e Janela de Takens serão considerados modelos distintos.

6.1 Janela Deslizante

Os resultados obtidos pelo modelo utilizando a abordagem de janela deslizante são apresentados nas Tabelas 6.1 e 6.2, que detalham as métricas de desempenho para cada cenário avaliado, tanto na validação cruzada quanto no conjunto de teste. Além disso, as Figuras 6.22, 6.23, 6.24 e 6.25 ilustram essa comparação de forma gráfica, representando os mesmos dados das tabelas em gráficos de pontos.

As subseções a seguir discutem os resultados de cada modelo, abordando os gráficos de previsão e a distribuição dos resíduos. Serão apresentadas as previsões comparadas aos valores reais, além da análise da distribuição dos resíduos por meio de histogramas e gráficos de dispersão, com o objetivo de compreender o comportamento do modelo e avaliar sua adequação aos dados.

6.1.1 NARX

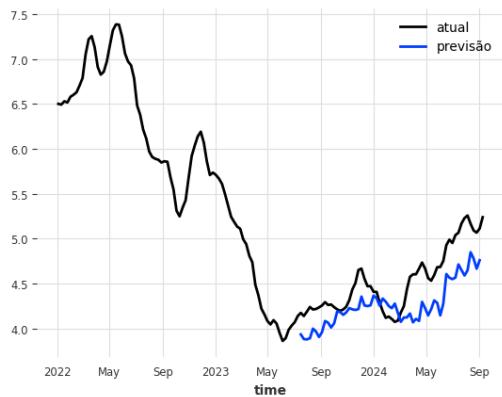
Conforme apresentado na Figura 6.1 destaca-se que o modelo (b) foi o único capaz de mapear os dados reais de forma satisfatória, apresentando baixos valores de erro no conjunto de teste e mantendo essa performance de forma consistente no conjunto de desenvolvimento. Os resultados obtidos indicam que o desvio padrão dos erros foi inferior a 10% das respectivas medidas, como detalhado nas Tabelas 6.1 e 6.2. Apesar do desempenho insatisfatório dos modelos (c) e (d), observa-se que ambos mantiveram consistência ao longo das validações.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura 6.2, observou-se que o modelo (b) exibe uma dispersão reduzida, evidenciando uma correlação positiva clara entre os valores previstos e observados. Em contraste, os modelos (a) e (c) demonstram maior dispersão para valores menores e uma redução dessa dispersão à medida que os valores aumentam, com uma linha de tendência também positiva. O modelo (d), por sua vez, apresenta uma dispersão considerável ao longo de todo o intervalo de valores, sem uma clara tendência.

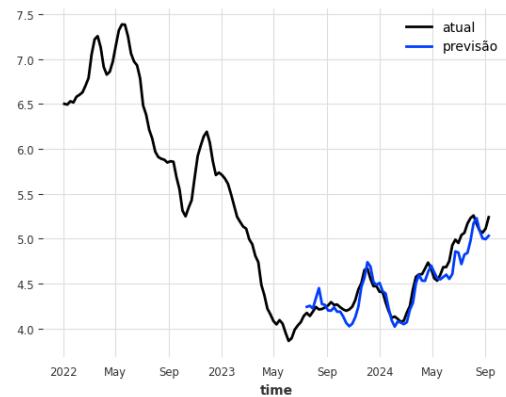
Por fim, os histogramas dos resíduos, ilustrados na Figura 6.3, revelam que o modelo (b) gera uma distribuição dos resíduos mais próxima de uma distribuição normal, embora com alguns valores discrepantes próximos a 0,2 e 0,3, que deslocam a média para a direita. Em contrapartida, os modelos (a), (c) e (d) exibem distribuições assimétricas, com a presença de valores discrepantes mais pronunciados, indicando um comportamento menos ideal na previsão dos resíduos.

6.1.2 DA-RNN

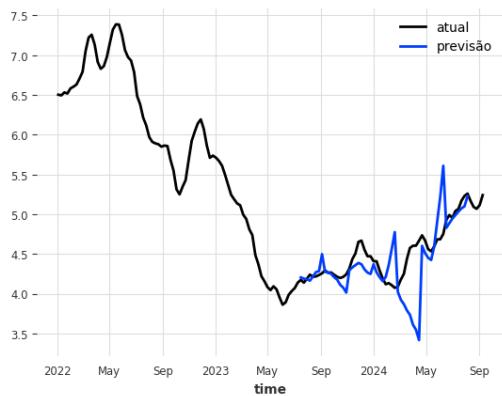
Como ilustrado na Figura 6.4, o modelo (c) apresentou uma capacidade eficaz de mapeamento, reproduzindo de forma precisa os dados reais e alcançando baixos valores de erro no conjunto de teste, conforme exposto na Tabela 6.2. Contudo, os resultados obtidos no conjunto de desenvolvimento revelaram-se inconsistentes, sugerindo um possível superdimensionamento do modelo para este conjunto de dados, conforme indicado pela Tabela 6.1. Ao comparar os resultados das Tabelas 6.1 e 6.2, nota-se uma inversão no desempenho: enquanto o modelo (c) obteve o melhor



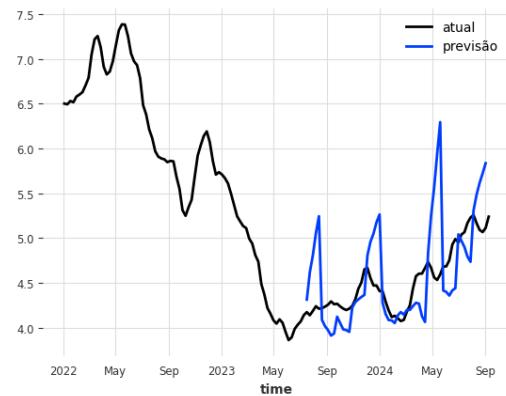
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

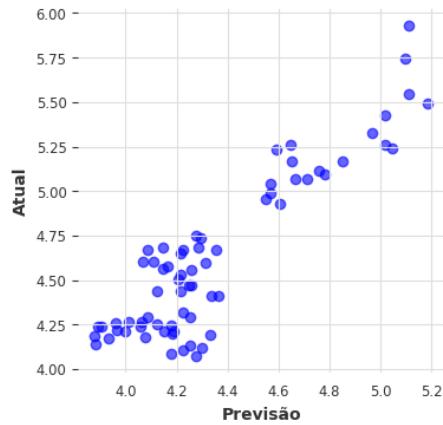


(c) Treinado com Biodiesel Nacional sem tendência

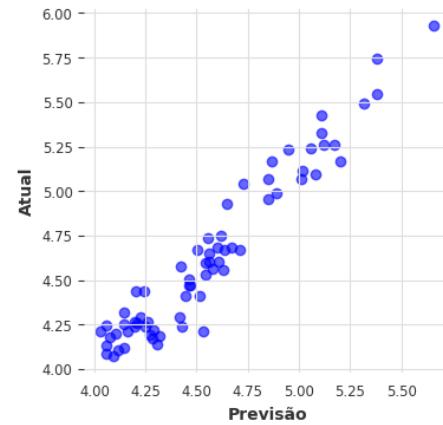


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

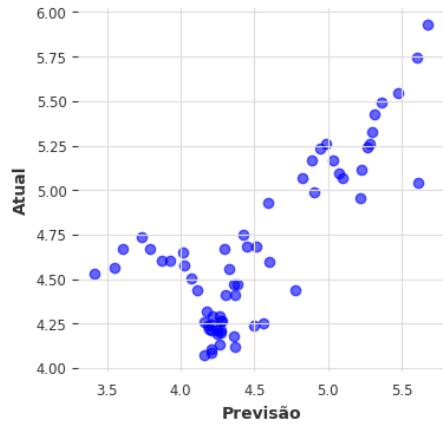
Figura 6.1: Resultados do modelo NARX no conjunto de teste



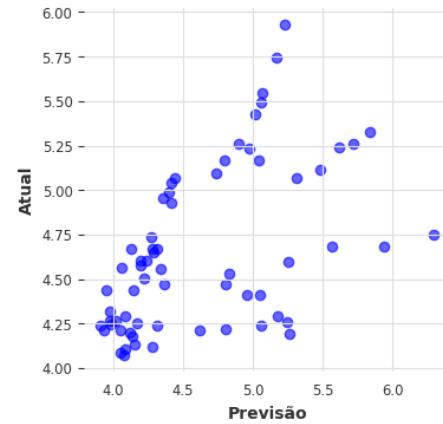
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

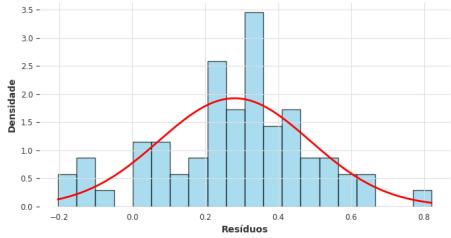


(c) Treinado com Biodiesel Nacional
sem tendência

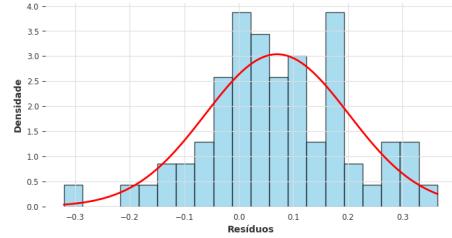


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

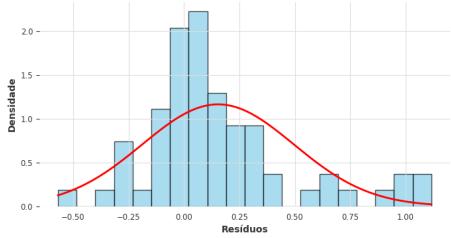
Figura 6.2: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo **NARX** no conjunto de teste



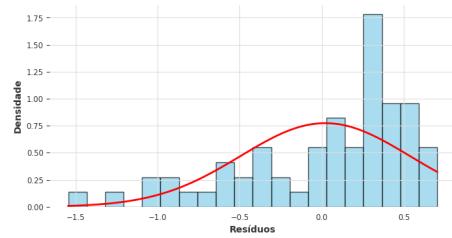
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



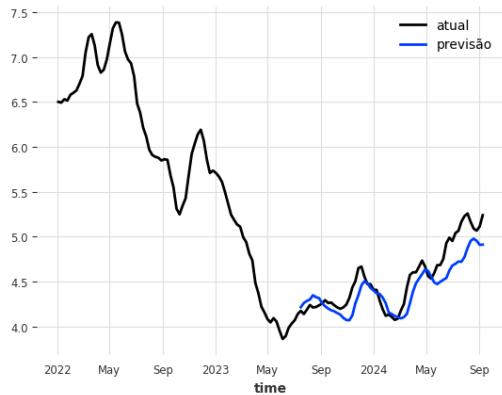
(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.3: Histograma dos erros do modelo NARX no conjunto de teste

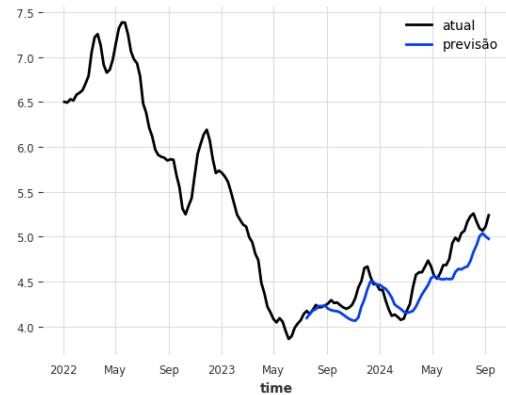
resultado no conjunto de teste, apresentou o pior desempenho no conjunto de desenvolvimento, ao passo que o modelo (d) demonstrou o comportamento oposto, que novamente sugere um superdimensionamento do modelo.

Adicionalmente, ao examinar os gráficos de dispersão apresentados na Figura 6.5, observou-se que o modelo (c) apresenta uma dispersão reduzida, o que indica uma correlação positiva clara entre os valores previstos e observados. Em contraste, os modelos (a), (b) e (d) exibem uma maior dispersão para valores mais baixos, com uma diminuição dessa dispersão à medida que os valores aumentam, sendo que a linha de tendência, embora também positiva, possui uma inclinação ligeiramente inferior a 1, sugerindo um aumento do erro conforme os valores reais se elevam.

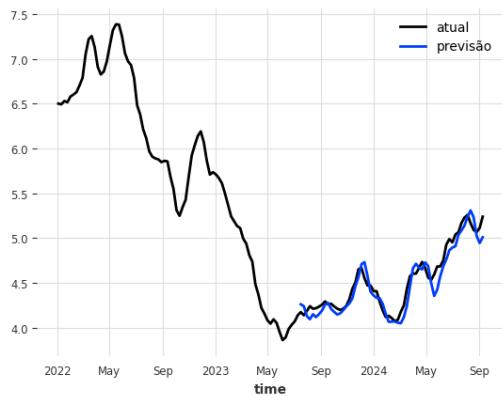
Por último, conforme ilustrado na Figura 6.6, os histogramas dos resíduos indicam que o modelo (c) gera uma distribuição dos resíduos que se aproxima de uma distribuição normal, com um leve desvio para a direita. Em contrapartida, os modelos (a), (b) e (d) apresentam distribuições com caudas pesadas, sugerindo maior variabilidade nos erros de previsão.



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

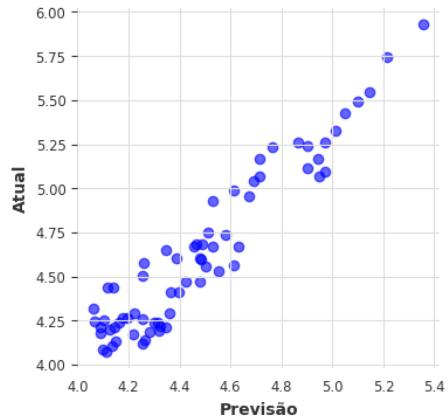


(c) Treinado com Biodiesel Nacional sem tendência

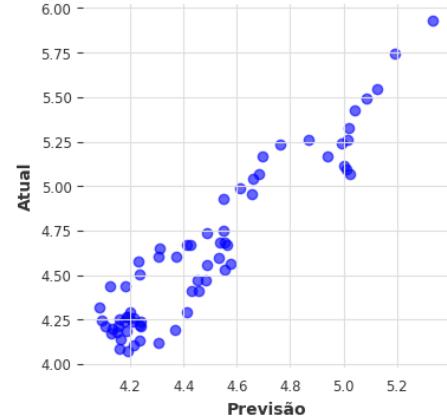


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

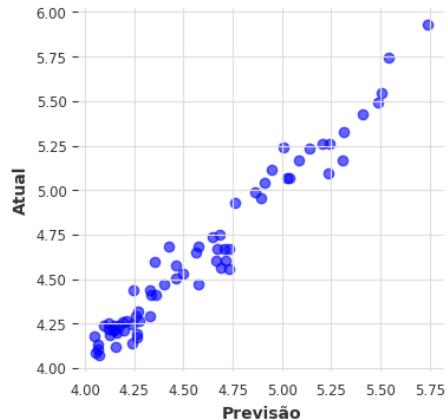
Figura 6.4: Resultados do modelo DA-RNN no conjunto de teste



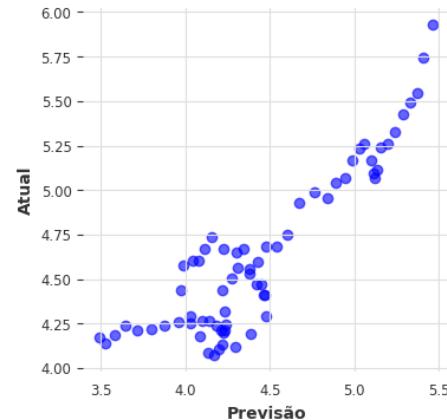
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

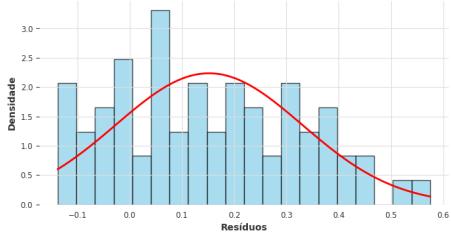


(c) Treinado com Biodiesel Nacional
sem tendência

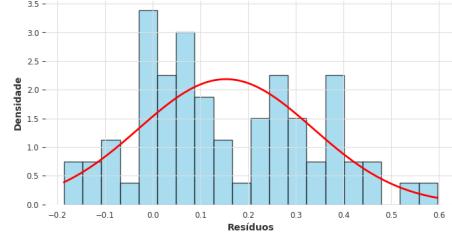


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

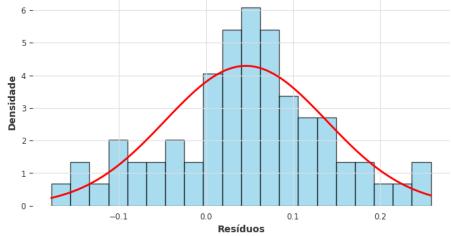
Figura 6.5: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo DA-RNN no conjunto de teste



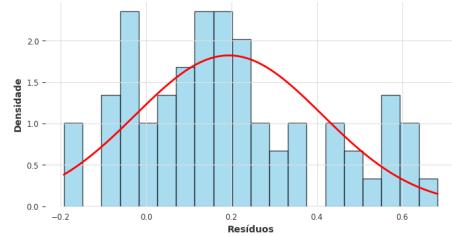
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

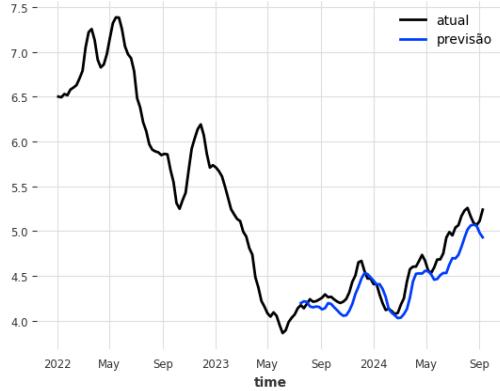
Figura 6.6: Histograma dos erros do modelo DA-RNN no conjunto de teste

6.1.3 MLP

Como mostrado na Tabela 6.2 e na Figura 6.7, os modelos (c) e (d) apresentaram os menores valores de erro entre todos os modelos baseados em janela deslizante. No entanto, apenas os modelos (b) e (d), que utilizaram um conjunto de treinamento mais extenso, mostraram resultados consistentes no conjunto de desenvolvimento, como evidenciado na Tabela 6.1. Esse desempenho sugere que o tamanho do conjunto de treinamento desempenha um papel importante na capacidade de generalização do modelo.

Adicionalmente, a análise dos gráficos de dispersão apresentados na Figura 6.8 revelou que os modelos (c) e (d) exibem uma dispersão reduzida, com correlação próxima de 1, indicando um bom ajuste entre os valores previstos e observados. Em contraste, os modelos (a) e (b) apresentaram maior dispersão, com as linhas de tendência mostrando uma inclinação ligeiramente superior a 1, sugerindo um aumento no erro conforme os valores reais aumentam.

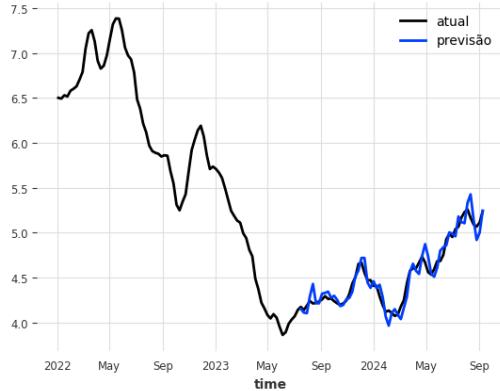
Por fim, ao examinar os histogramas dos resíduos, apresentados na Figura 6.9, observou-se que os modelos (a), (b) e (d) exibem distribuições que se aproximam de uma distribuição normal, embora com a presença de múltiplos picos. Em contrapartida, o modelo (c) demonstrou uma distribuição assimétrica, com maior concentração de resíduos em torno de zero.



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência

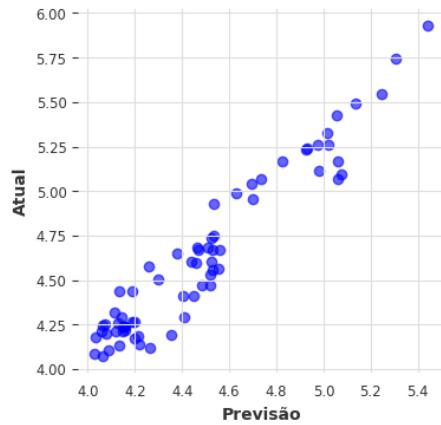


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

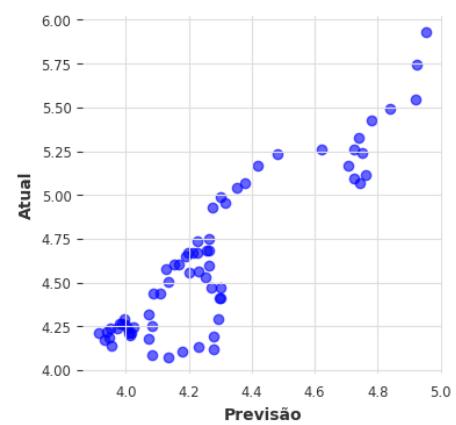
Figura 6.7: Resultados do modelo MLP no conjunto de teste

6.1.4 **IMP**

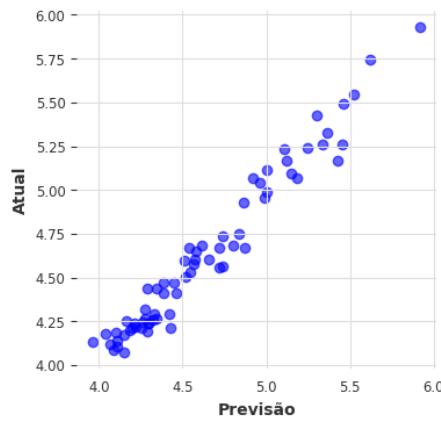
O modelo **IMP** apresentou resultados inconsistentes em todos os cenários analisados e não obteve um desempenho satisfatório no conjunto de teste, conforme ilustrado na Figura 6.10.



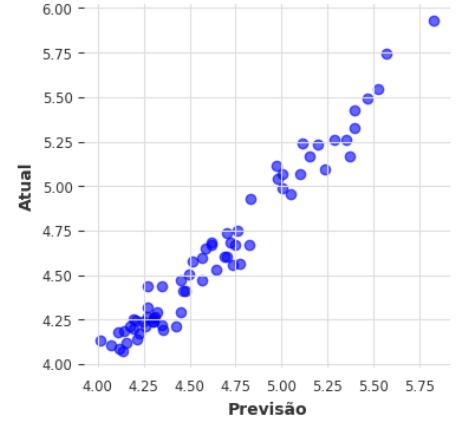
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

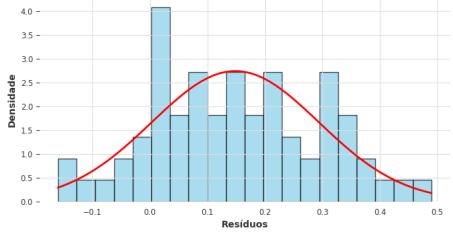


(c) Treinado com Biodiesel Nacional
sem tendência

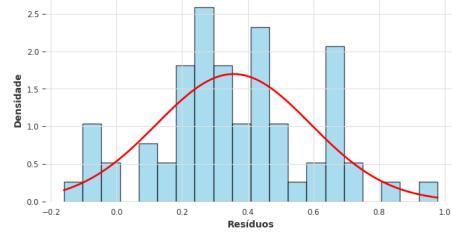


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

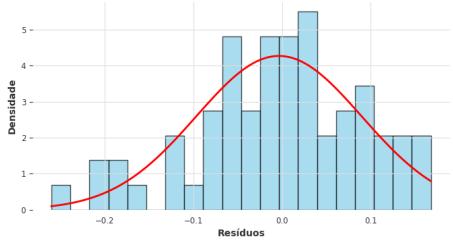
Figura 6.8: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo MLP no conjunto de teste



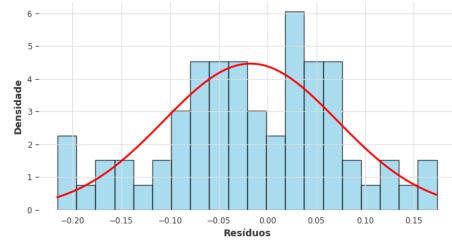
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.9: Histograma dos erros do modelo MLP no conjunto de teste

Adicionalmente, ao examinar os gráficos de dispersão apresentados na Figura 6.11, observou-se uma alta dispersão nos valores previstos por todos os modelos, indicando uma correlação fraca entre os valores previstos e observados. Os histogramas dos resíduos, apresentados na Figura 6.12, revelaram distribuições assimétricas e bimodais, sugerindo uma variabilidade incomum nos erros de previsão.

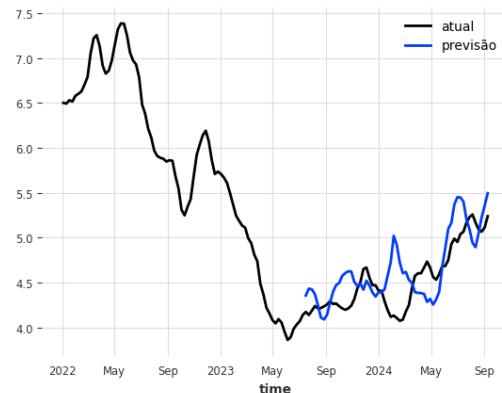
6.1.5 IDLN

Conforme mostrado na Figura 6.13, os modelos (a) e (c) apresentaram bons resultados no conjunto de teste, embora tenham demonstrado inconsistência no conjunto de desenvolvimento, como evidenciado na Tabela 6.2. Essa discrepância pode indicar que, apesar de seu bom desempenho em condições controladas, esses modelos enfrentam dificuldades em generalizar adequadamente para dados não vistos durante o treinamento.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura 6.14, observou-se que os modelos (a) e (c) exibem uma dispersão reduzida, com uma



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

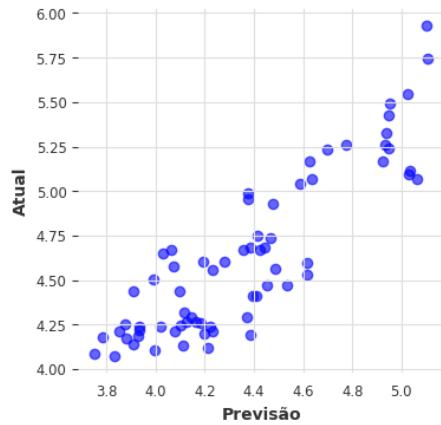


(c) Treinado com Biodiesel Nacional sem tendência

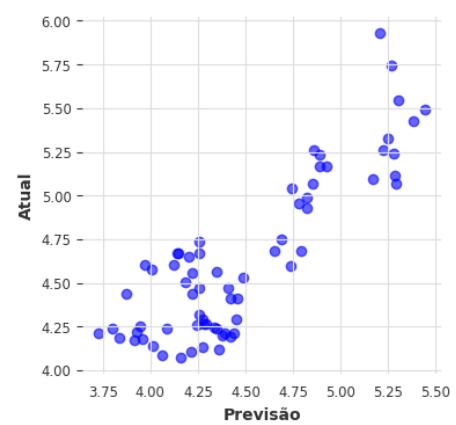


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

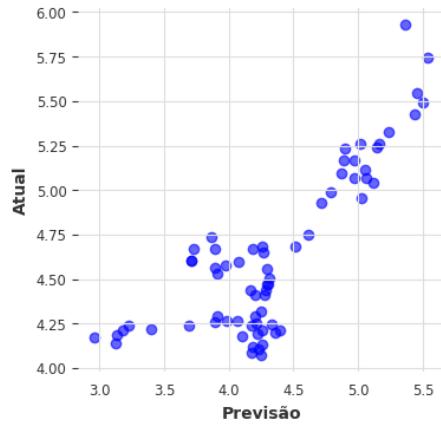
Figura 6.10: Resultados do modelo IMP no conjunto de teste



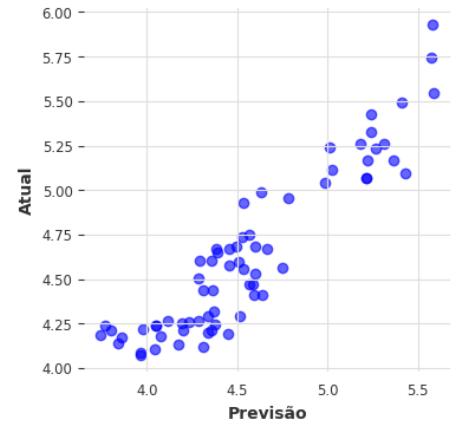
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

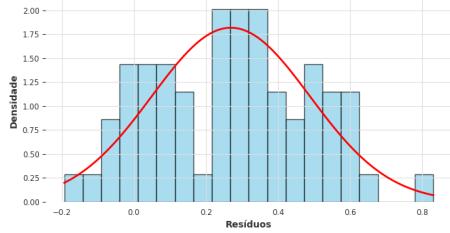


(c) Treinado com Biodiesel Nacional
sem tendência

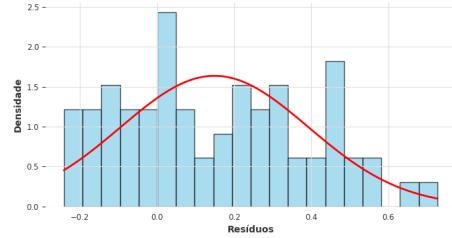


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

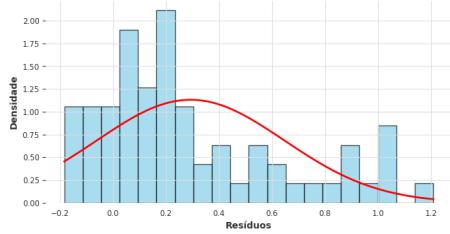
Figura 6.11: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo IMP no conjunto de teste



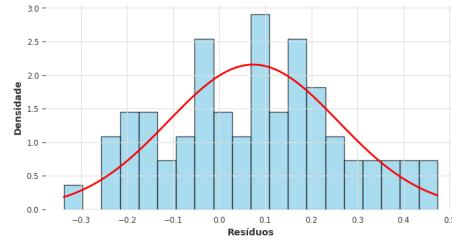
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

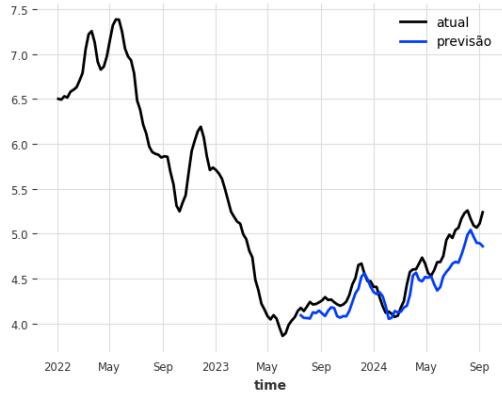
Figura 6.12: Histograma dos erros do modelo IMP no conjunto de teste

correlação positiva clara entre os valores previstos e observados, embora o modelo (a) apresente uma correlação superior a 1. Em contraste, os modelos (b) e (d) apresentam uma dispersão maior, com uma correlação aparentemente não linear. O modelo (b), em particular, demonstrou uma alta dispersão.

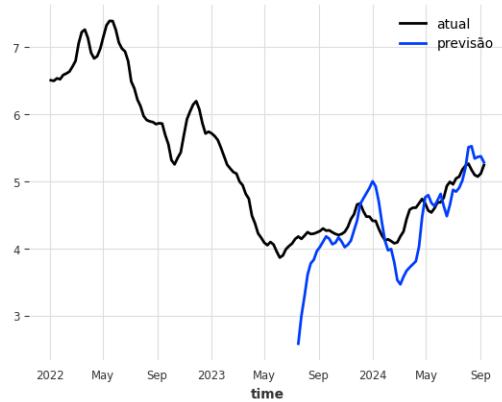
Por fim, ao examinar os histogramas apresentados na Figura 6.15, observou-se que o modelo (c) apresenta uma distribuição dos resíduos próxima a uma distribuição normal, enquanto os demais modelos exibem distribuições irregulares, com caudas pesadas e múltiplas modas.

6.1.6 N-Linear

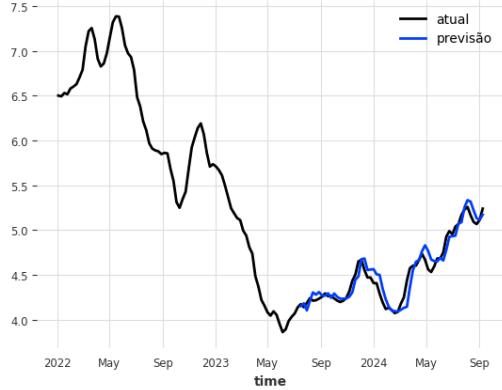
Conforme mostrado na Figura 6.16, os modelos (a) e (c) alcançaram um mapeamento próximo aos valores reais, destacando-se como os melhores no conjunto de teste. No entanto, de acordo com a Tabela 6.1, esses resultados mostraram-se inconsistentes ao longo dos diferentes conjuntos de validação cruzada, sugerindo uma possível falta de robustez na generalização dos modelos para diferentes subconjuntos



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

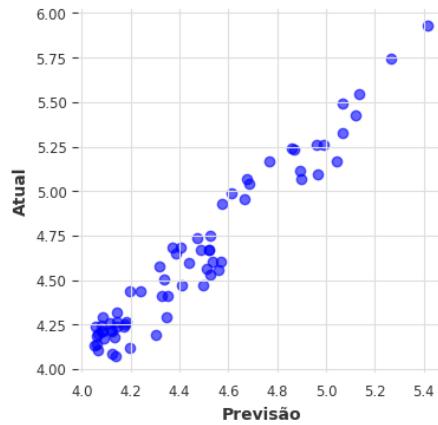


(c) Treinado com Biodiesel Nacional sem tendência

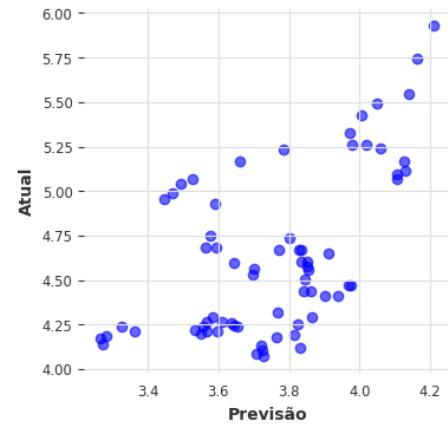


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

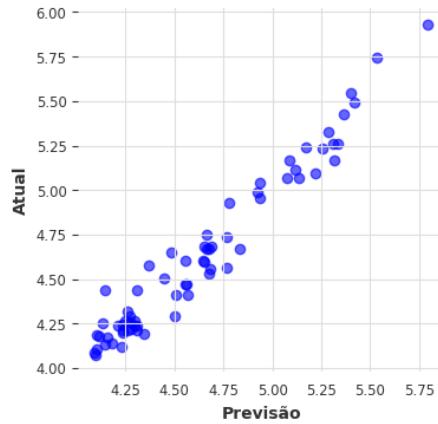
Figura 6.13: Resultados do modelo IDLN no conjunto de teste



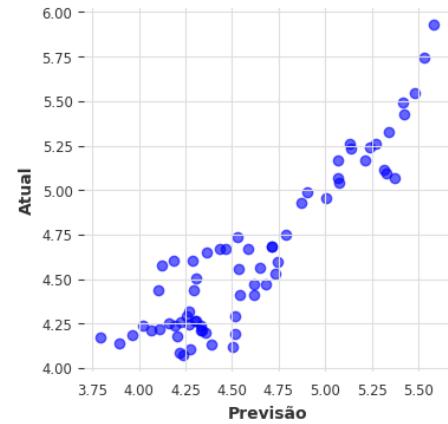
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

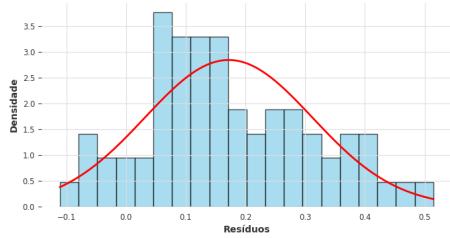


(c) Treinado com Biodiesel Nacional
sem tendência

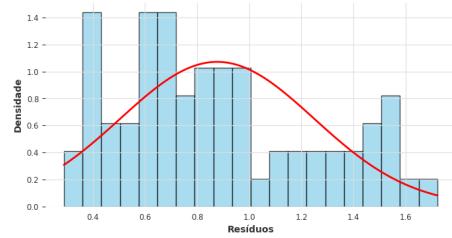


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

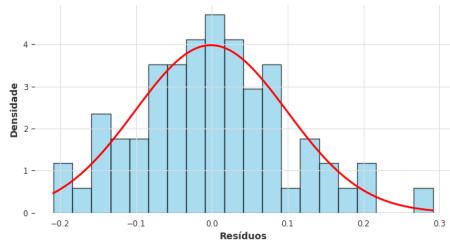
Figura 6.14: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo **IDLN** no conjunto de teste



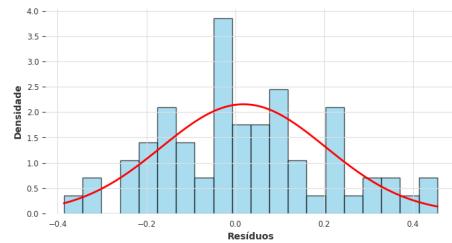
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.15: Histograma dos erros do modelo IDLN no conjunto de teste

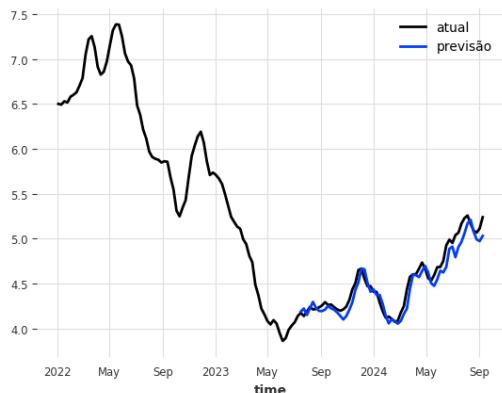
de dados.

Além disso, a análise dos gráficos de dispersão apresentados na Figura 6.17 revela que os modelos (a) e (c) apresentam baixa dispersão e uma clara correlação entre as previsões e os dados reais. Por outro lado, os modelos (b) e (d) exibem maior dispersão, embora ainda seja possível observar uma correlação positiva entre os valores previstos e observados.

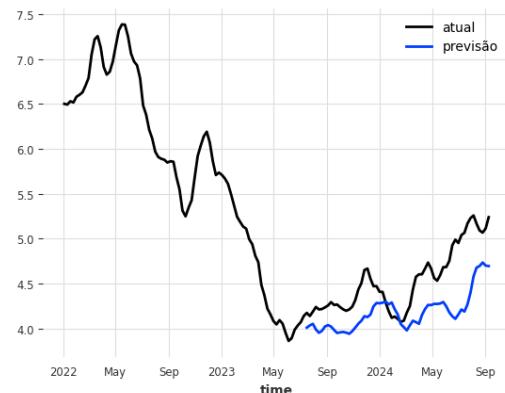
Por fim, ao examinar os histogramas apresentados na Figura 6.18, observou-se que o modelo (b) se aproxima de uma distribuição normal, enquanto os demais modelos apresentam um comportamento multimodal, com várias modas nas distribuições dos resíduos.

6.1.7 NHITS

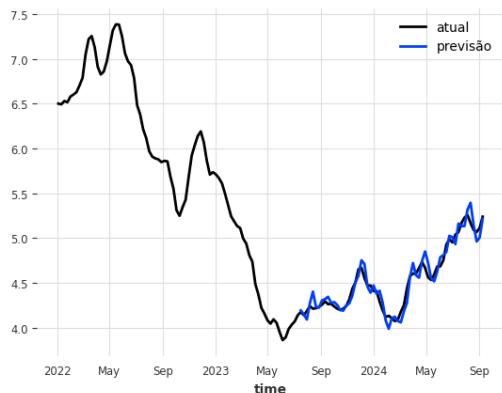
Conforme ilustrado na Figura 6.19, o modelo (c) conseguiu mapear de forma próxima os valores reais, embora, segundo a Tabela 6.1, esses resultados tenham



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

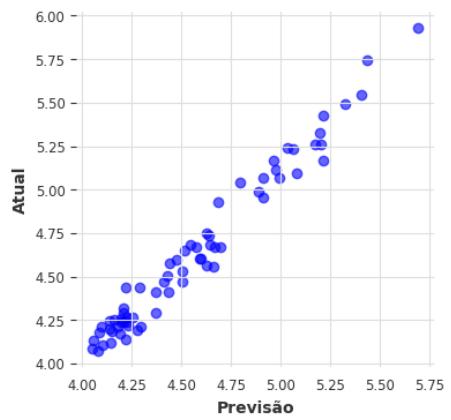


(c) Treinado com Biodiesel Nacional sem tendência

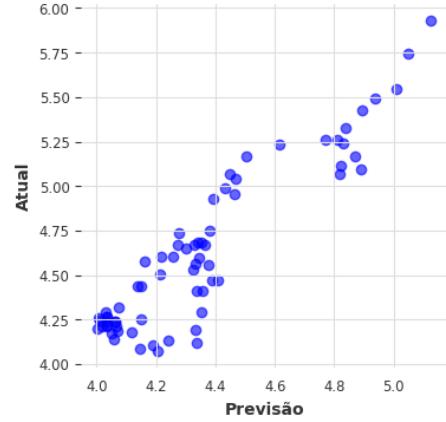


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

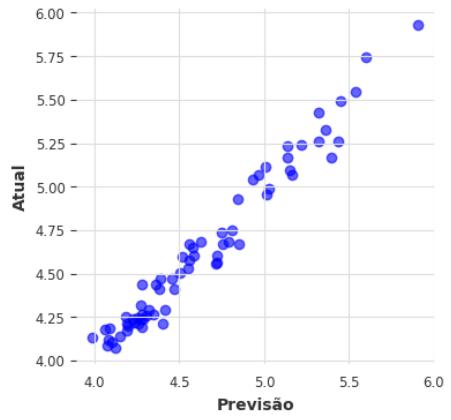
Figura 6.16: Resultados do modelo N-Linear no conjunto de teste



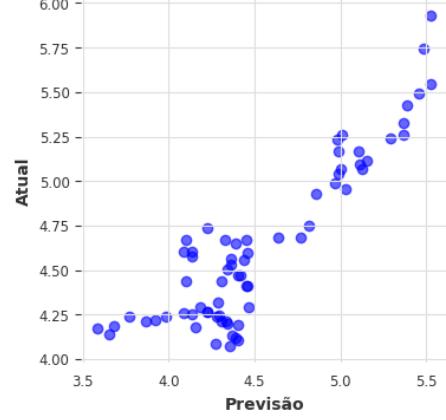
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

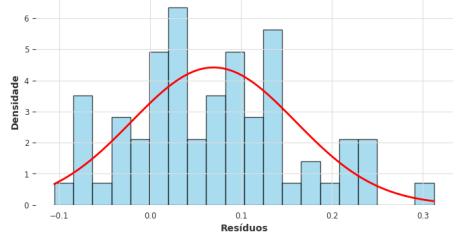


(c) Treinado com Biodiesel Nacional
sem tendência

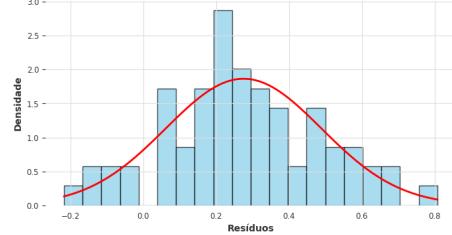


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

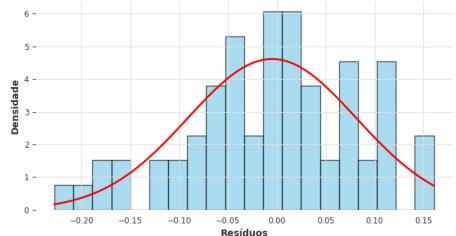
Figura 6.17: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo **N-Linear** no conjunto de teste



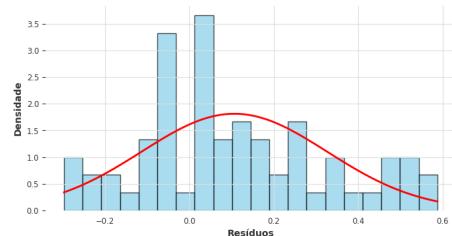
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.18: Histograma dos erros do modelo N-Linear no conjunto de teste

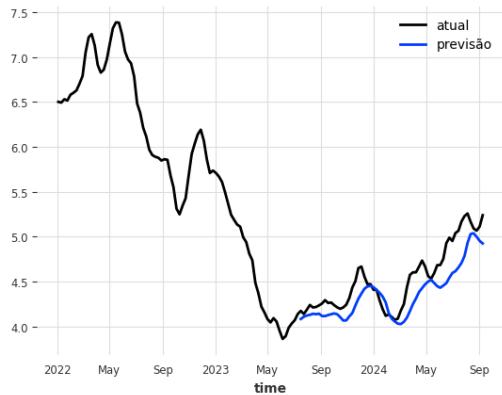
se mostrado inconsistentes. Os modelos (b) e (c) demonstraram consistência ao longo da validação cruzada, conforme evidenciado na Tabela 6.1, no entanto, ambos apresentaram desempenho inferior no conjunto de teste. Essa diferença entre os resultados de validação e teste sugere uma possível limitação na capacidade de generalização desses modelos quando expostos a dados não previamente vistos.

Além disso, a análise dos gráficos de dispersão apresentados na Figura 6.20 revela uma clara correlação positiva em todos os modelos, embora com uma maior dispersão para valores menores.

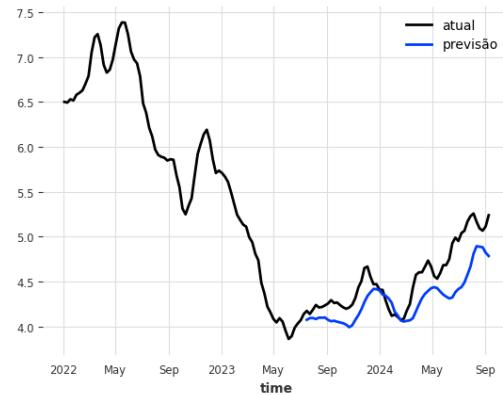
Ao examinar os histogramas da Figura 6.21, observa-se que todos os modelos exibem mais de uma moda, indicando uma distribuição multimodal dos resíduos.

6.2 Janela de Takens

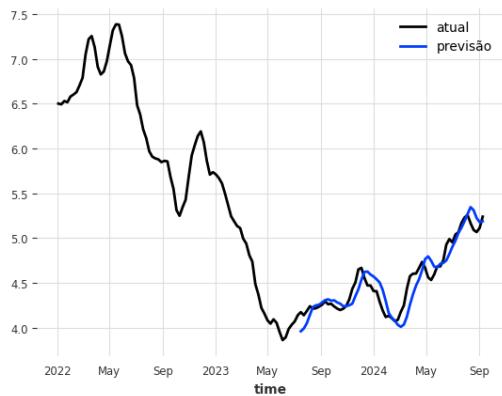
Os resultados obtidos pelo modelo utilizando a abordagem de janela de takens são apresentados nas Tabelas 6.3 e 6.4, que detalham as métricas de desempenho



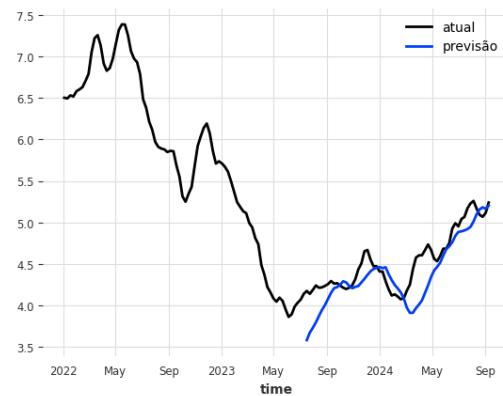
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

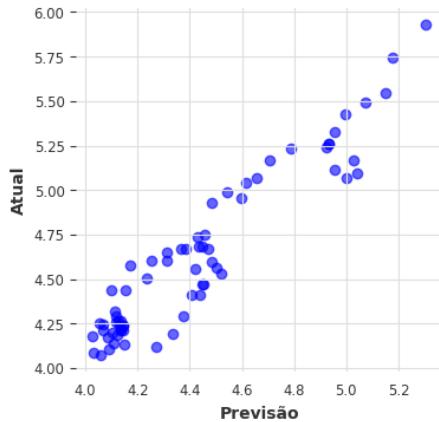


(c) Treinado com Biodiesel Nacional sem tendência

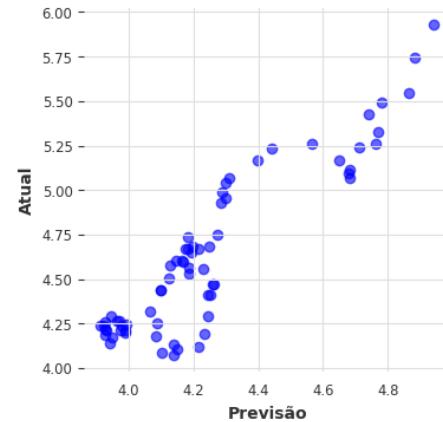


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

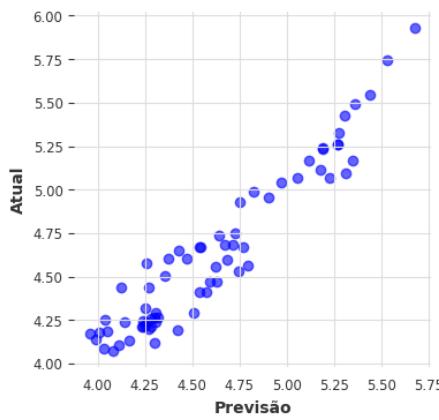
Figura 6.19: Resultados do modelo NHITS no conjunto de teste



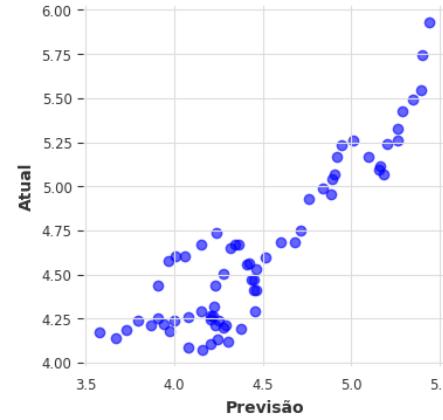
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja



(c) Treinado com Biodiesel Nacional
sem tendência



(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

Figura 6.20: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo NHITS no conjunto de teste

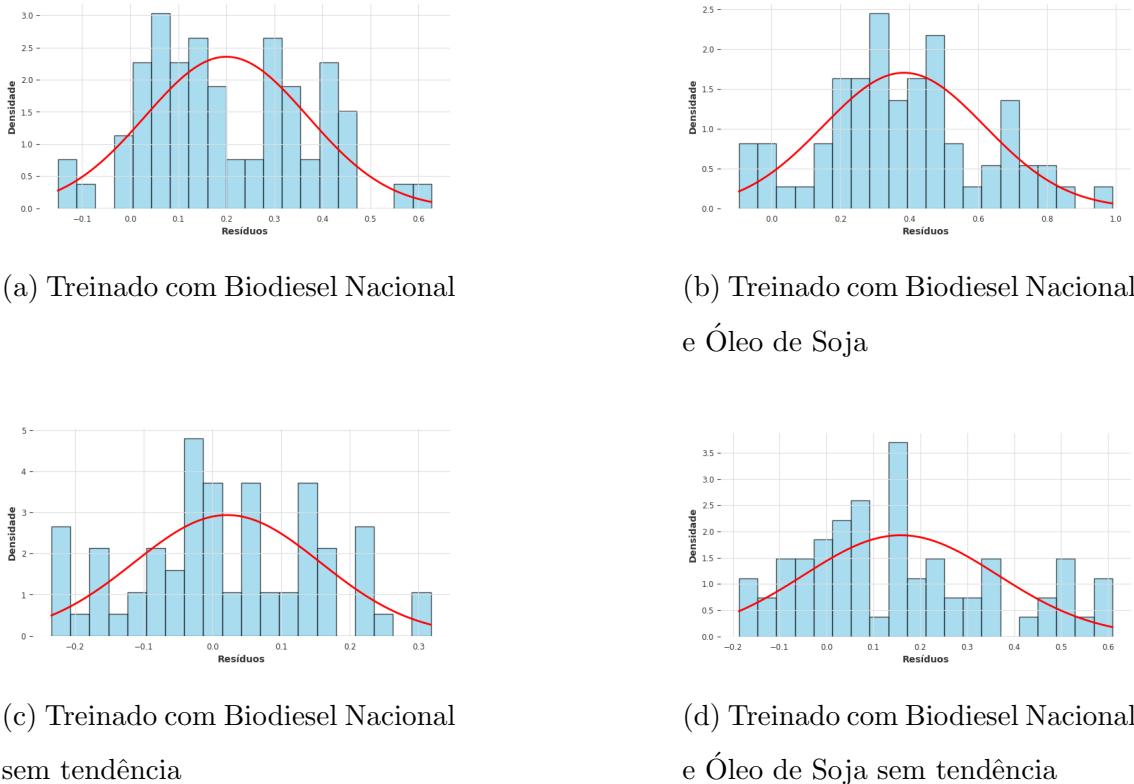
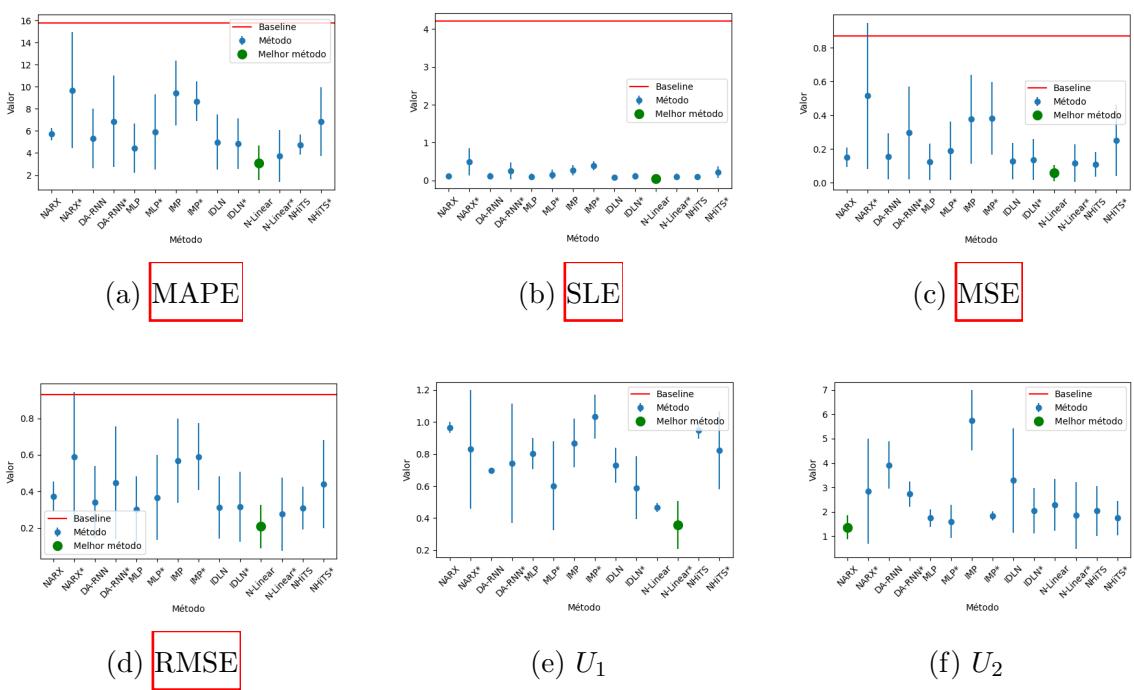
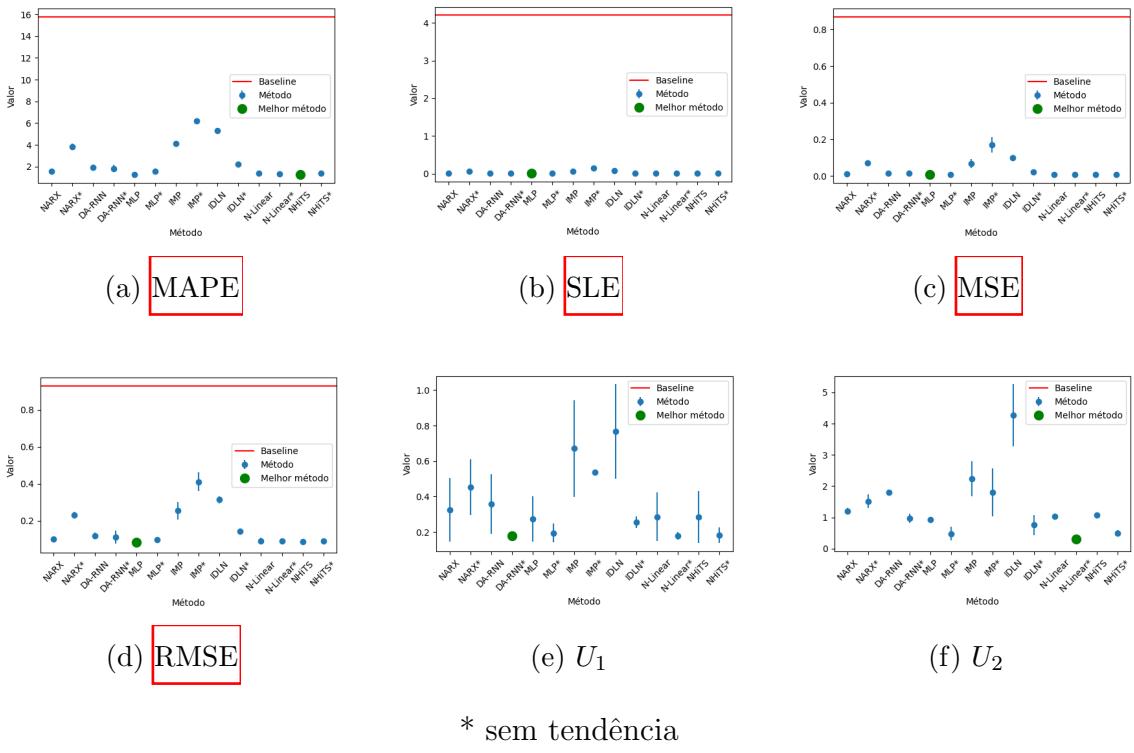


Figura 6.21: Histograma dos erros do modelo NHITS no conjunto de teste



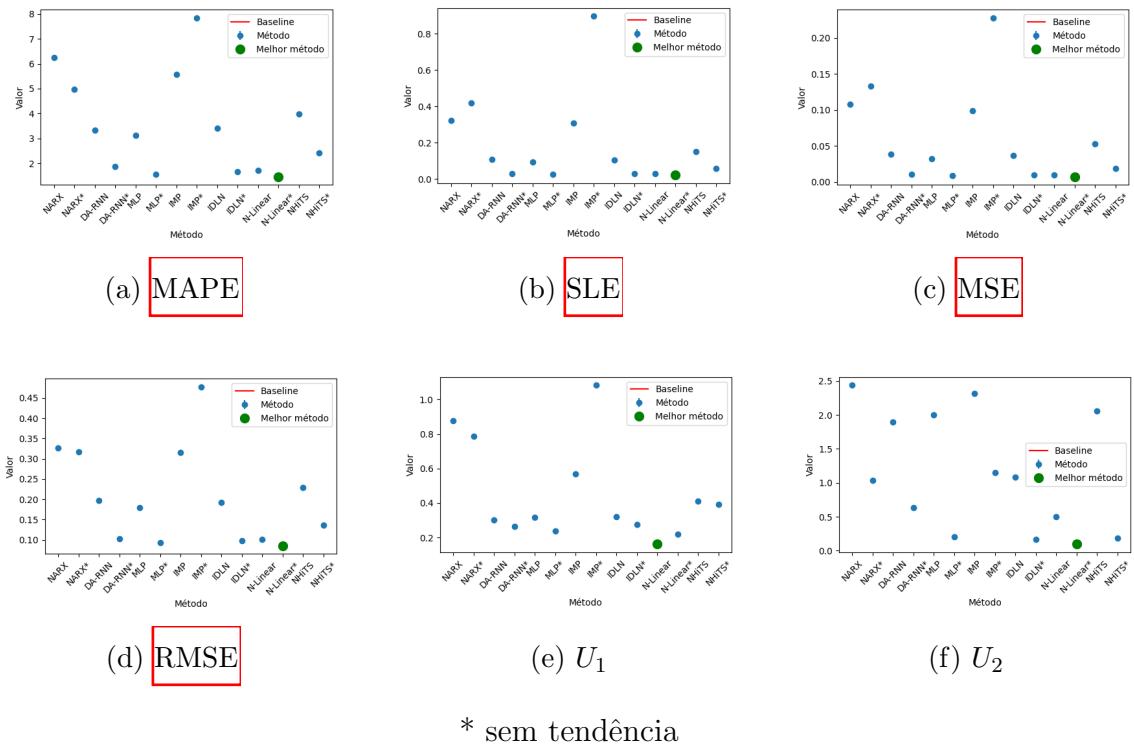
* sem tendência

Figura 6.22: Resultados da Validação Cruzada treinado com Biodiesel Nacional usando Janela Deslizante



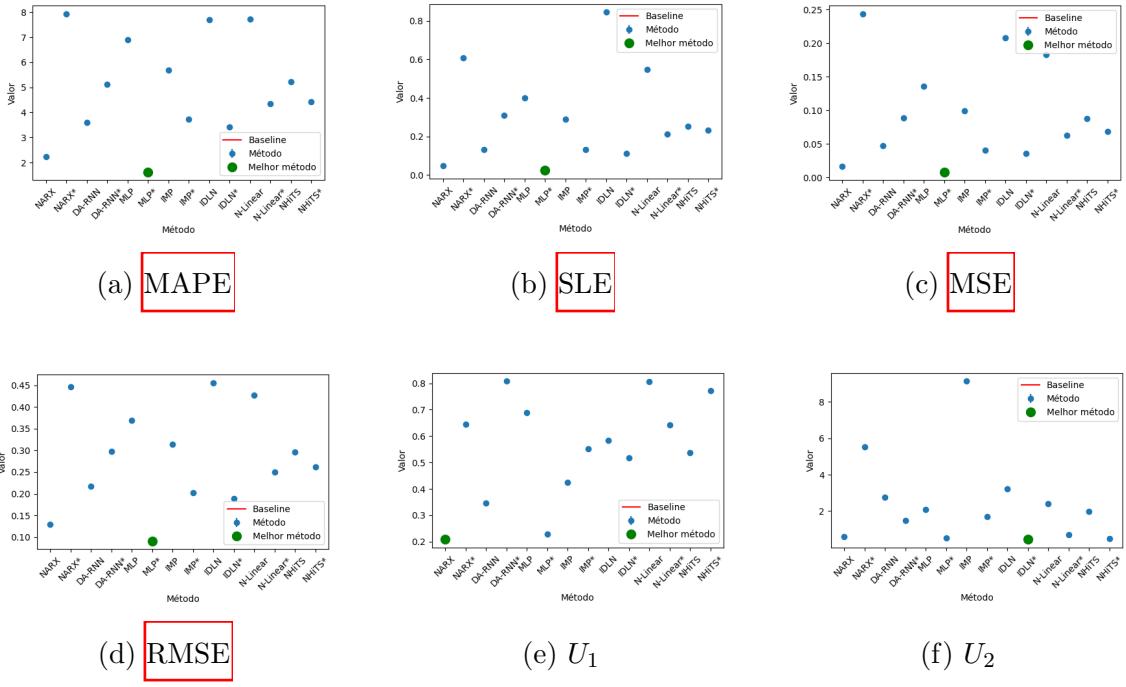
* sem tendência

Figura 6.23: Resultados da Validação Cruzada treinado com Biodiesel Nacional e Óleo de Soja usando Janela Deslizante



* sem tendência

Figura 6.24: Resultados do Conjunto de Teste treinado com Biodiesel Nacional usando Janela Deslizante



* sem tendência

Figura 6.25: Resultados do Conjunto de Teste treinado com Biodiesel Nacional e Óleo de Soja usando Janela Deslizante

para cada cenário avaliado, tanto na validação cruzada quanto no conjunto de teste. Além disso, as Figuras 6.47, 6.48, 6.49 e 6.50 ilustram essa comparação de forma gráfica, representando os mesmos dados das tabelas em gráficos de pontos.

As subseções a seguir discutem os resultados de cada modelo, abordando os gráficos de previsão e a distribuição dos resíduos. Serão apresentadas as previsões comparadas aos valores reais, além da análise da distribuição dos resíduos por meio de histogramas e gráficos de dispersão, com o objetivo de compreender o comportamento do modelo e avaliar sua adequação aos dados.

6.2.1 **NARX**

Conforme apresentado na Figura 6.26 mostra-se que o modelo (b) foi o único capaz de mapear os dados reais de forma satisfatória, apresentando baixos valores de erro no conjunto de teste e mantendo essa performance de forma consistente no conjunto de desenvolvimento. Os resultados obtidos indicam que o desvio padrão dos erros foi inferior a 10% das respectivas medidas, como detalhado nas Tabelas

6.3 e **6.4**. Os demais modelos apresentam resultados inconsistentes no conjunto de validação e altos valores de erro no conjunto de teste.

Além disso, ao avaliar os gráficos de dispersão apresentados na Figura **6.27**, observa-se que o modelo (b) apresenta uma correlação positiva clara entre os valores previstos e observados, com baixa dispersão. Em contrapartida, os modelos (a), (c) e (d) exibem maior dispersão, com uma correlação mais fraca entre os valores previstos e observados.

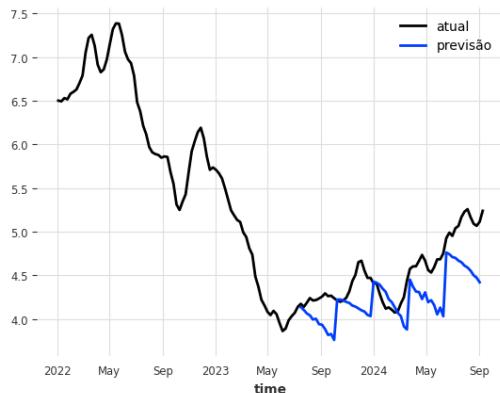
Por fim, ao analisar os histogramas da Figura **6.28**, observa-se que os modelos (a), (c) e (d) apresentam distribuições dos resíduos que se aproximam de uma distribuição normal, enquanto o modelo (b) exibe uma distribuição com cauda longa, sugerindo a presença de valores extremos nos resíduos.

6.2.2 DA-RNN

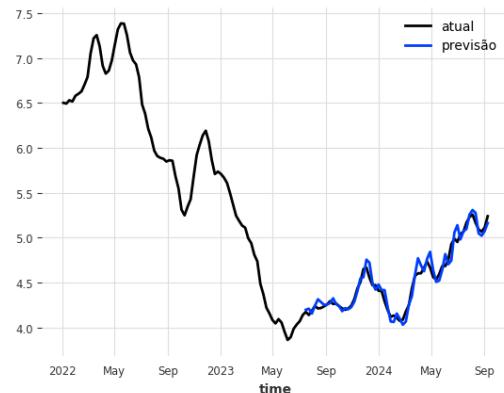
Conforme mostrado na Tabela **6.4** e ilustrado na Figura **6.29**, os modelos (b) e (c) obtiveram melhor resultado no conjunto de teste. Além disso, o modelo (b) também apresentou consistência na validação cruzada, como mostra a Tabela **6.3**, indicando boa capacidade de generalização para diferentes conjuntos de dados.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura **6.30**, observa-se que todos os modelos apresentam correlação positiva, embora os modelos (a) e (d) exibam maior dispersão, indicando uma maior variabilidade nos valores previstos em comparação aos demais.

Por fim, ao avaliar os histogramas da Figura **6.31**, observa-se que o modelo (c) apresenta uma distribuição dos resíduos que se aproxima de uma distribuição normal, sugerindo uma maior consistência nos erros de previsão. Em contraste, os modelos (a), (b) e (d) exibem distribuições com cauda longa, o que indica a presença de valores extremos nos resíduos.



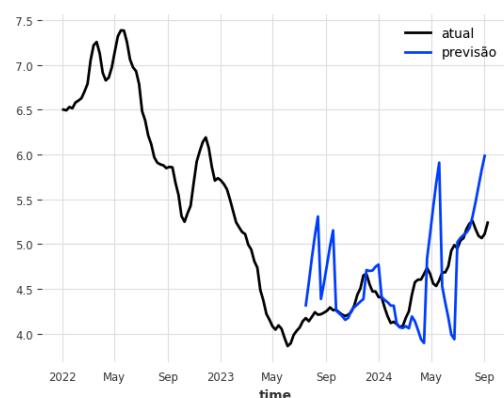
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

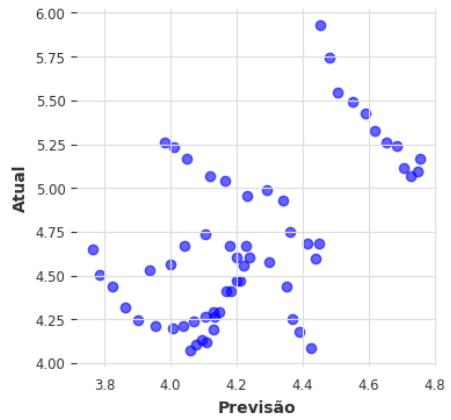


(c) Treinado com Biodiesel Nacional sem tendência

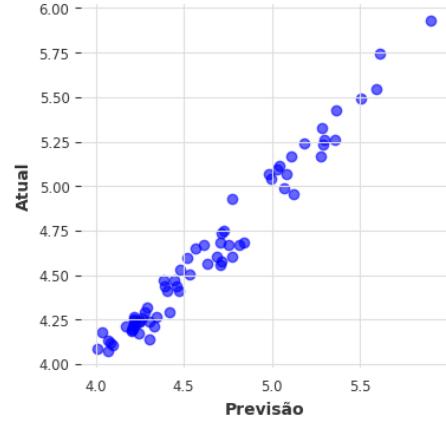


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

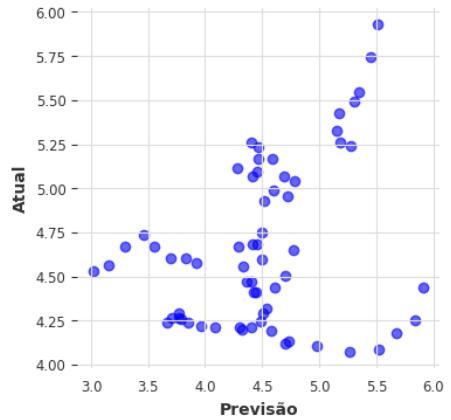
Figura 6.26: Resultados do modelo NARX Takens no conjunto de teste



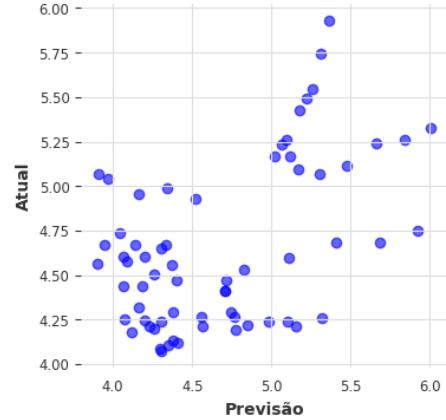
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

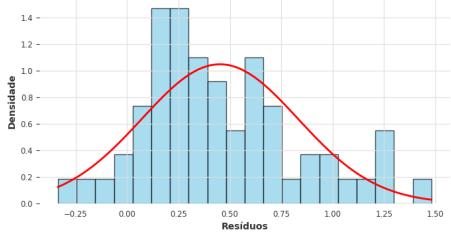


(c) Treinado com Biodiesel Nacional
sem tendência

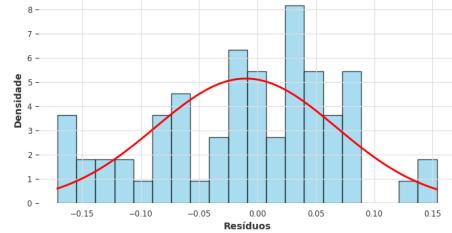


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

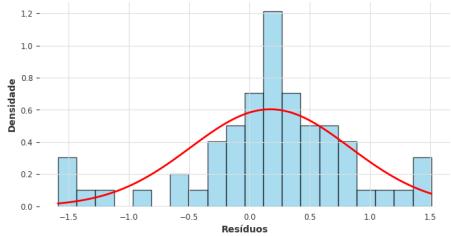
Figura 6.27: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo **NARX** Takens no conjunto de teste



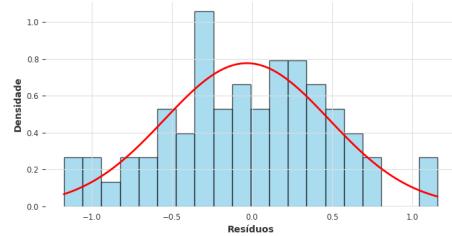
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

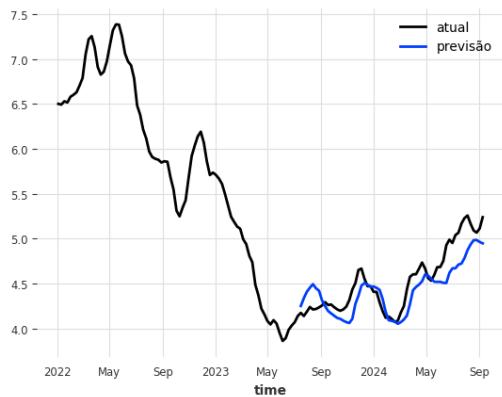
Figura 6.28: Histograma dos erros do modelo NARX Takens no conjunto de teste

6.2.3 MLP

O modelo MLP apresentou bons resultados em todos os casos, com o modelo (b) possuindo o menor erro entre todos os outros da janela de takens, como mostra a Tabela 6.4 e na Figura 6.32. Além disso, possui um dos melhores erros de validação cruzada com a melhor consistência, conforme a Tabela 6.3. Esses resultados indicam uma capacidade robusta de generalização e precisão do modelo nesse contexto.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura 6.33, observa-se que todos os modelos apresentam correlação positiva, embora o modelo (a) apresente maior dispersão, indicando maior variabilidade nos valores previstos.

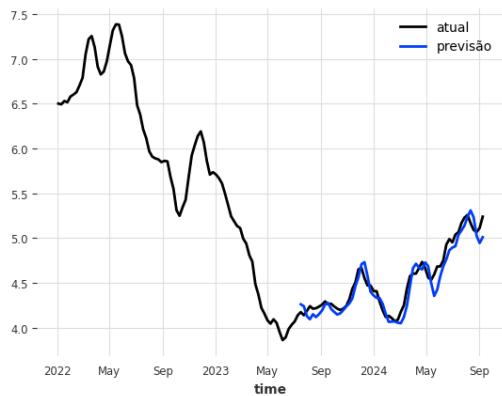
Por fim, ao examinar os histogramas da Figura 6.34, observa-se que os modelos (b) e (c) apresentam distribuições dos resíduos que se aproximam de uma distribuição normal. Em contraste, o modelo (a) exibe uma distribuição irregular, enquanto o modelo (b) apresenta uma distribuição com múltiplos picos, sugerindo um comportamento multimodal nos resíduos.



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

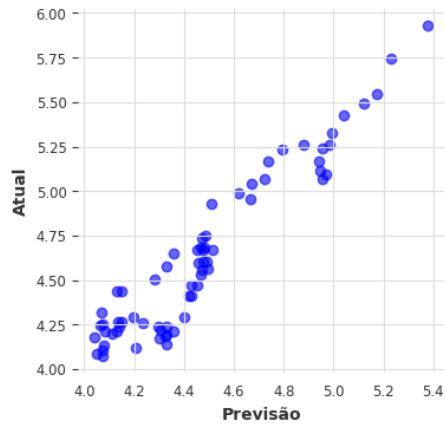


(c) Treinado com Biodiesel Nacional sem tendência

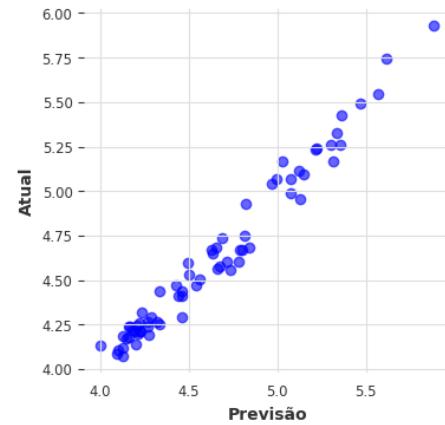


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

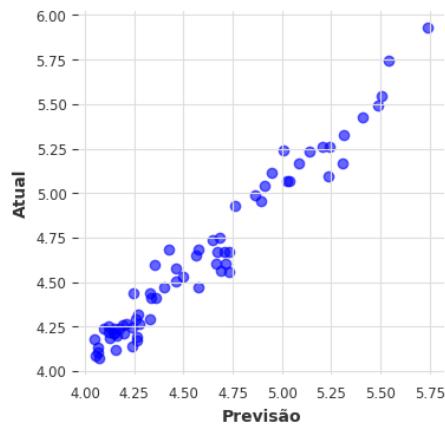
Figura 6.29: Resultados do modelo DA-RNN Takens no conjunto de teste



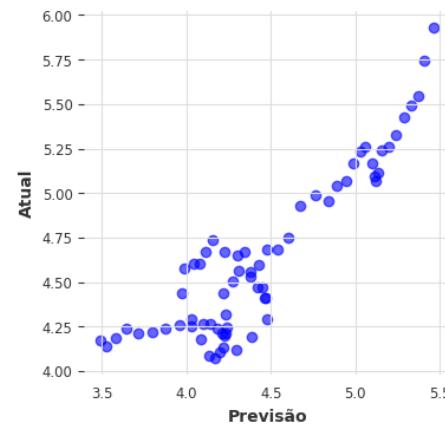
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

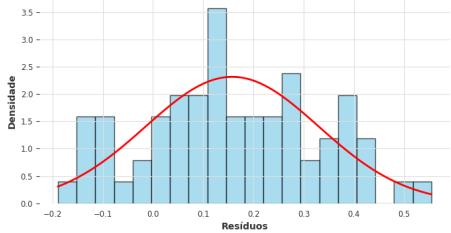


(c) Treinado com Biodiesel Nacional
sem tendência

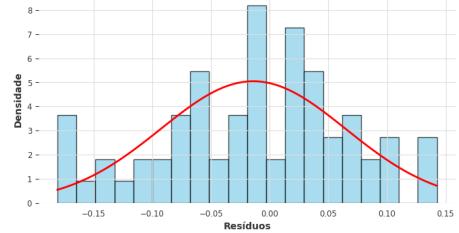


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

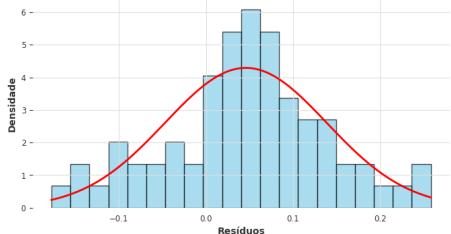
Figura 6.30: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo DA-RNN Takens no conjunto de teste



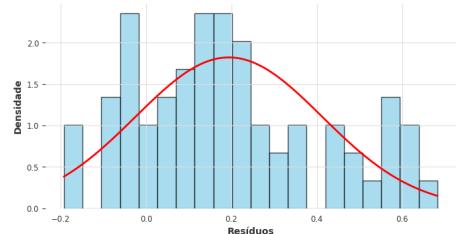
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.31: Histograma dos erros do modelo DA-RNN Takens no conjunto de teste

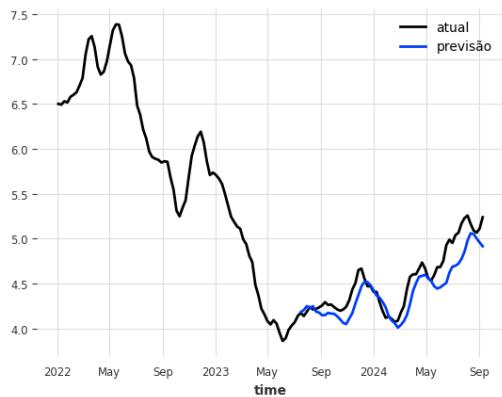
6.2.4 **IMP**

Os modelos **IMP** apresentaram erros elevados, quando comparados a outros modelos. Além disso, o modelo apresenta resultados inconsistentes para cada conjunto de treinamento, como mostra a Tabela 6.3 e ilustrado pela Figura 6.35, indicando dificuldade da generalização dos dados de treinamento.

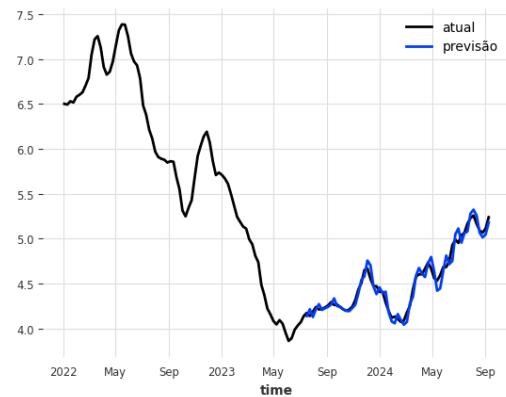
Adicionalmente, ao analisar os gráficos de dispersão apresentados na Figura 6.36, observa-se que, embora os modelos exibam correlação positiva, há uma grande dispersão, indicando maior variabilidade nos valores previstos. Os histogramas mostrados na Figura 6.37 revelam distribuições multimodais e com cauda longa, sugerindo a presença de valores extremos nos resíduos e uma estrutura complexa dos erros de previsão.

6.2.5 **IDLN**

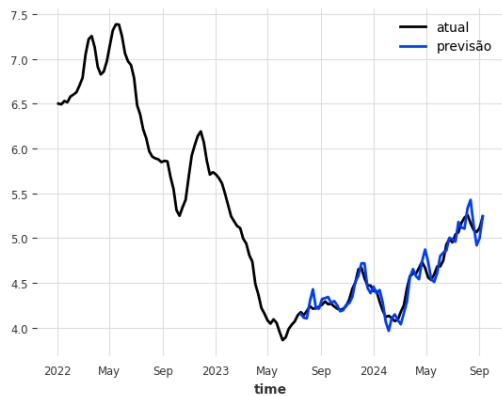
O modelo (c) apresentou resultado satisfatório no conjunto de teste, como mostrado na Figura 6.38. Apesar disso, o desempenho se mostra inconsistente, como



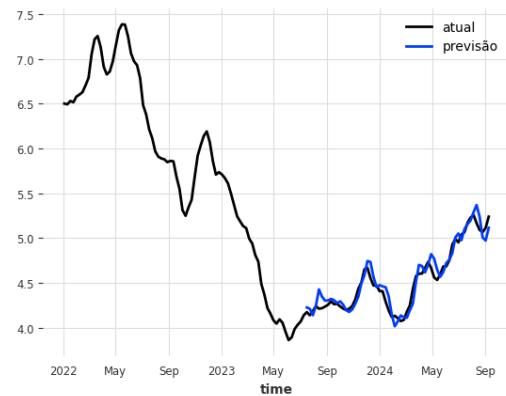
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

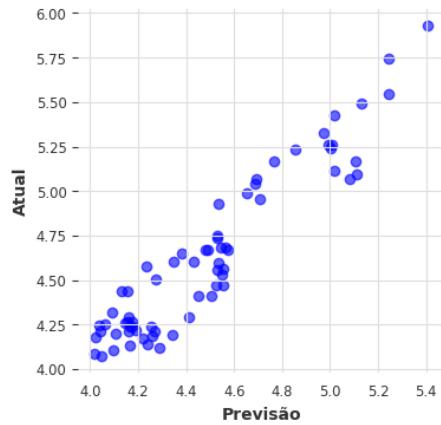


(c) Treinado com Biodiesel Nacional sem tendência

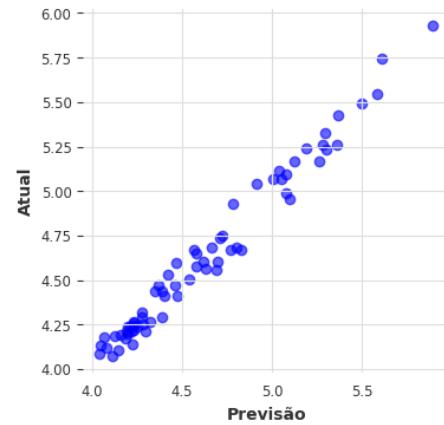


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

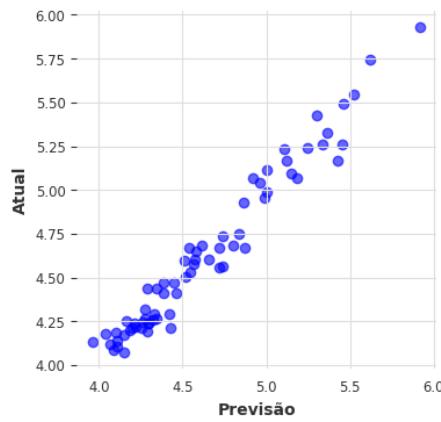
Figura 6.32: Resultados do modelo MLP Takens no conjunto de teste



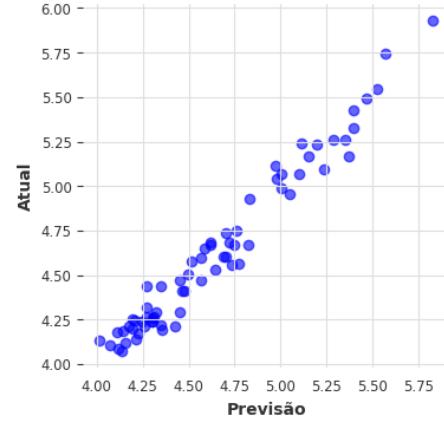
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

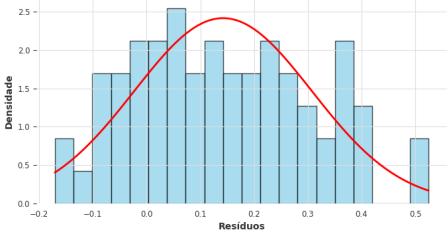


(c) Treinado com Biodiesel Nacional
sem tendência

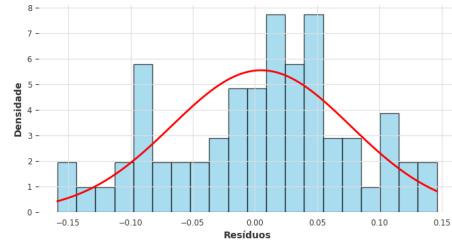


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

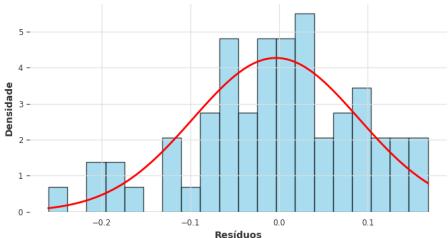
Figura 6.33: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo **MLP** Takens no conjunto de teste



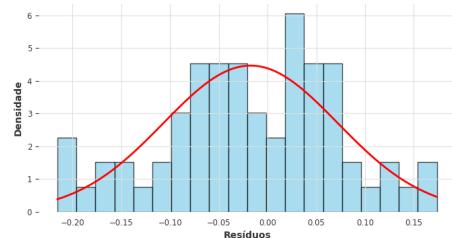
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.34: Histograma dos erros do modelo **MLP** Takens no conjunto de teste

visto na Tabela 6.3, com desvio padrão aproximadamente 50% do valor da medida.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura 6.39, observa-se que todos os modelos apresentam correlação positiva. No entanto, os modelos (a), (b) e (d) exibem maior dispersão, o que indica maior variabilidade nos valores previstos em comparação aos demais modelos.

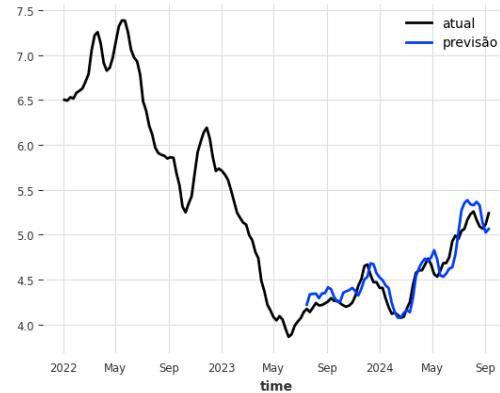
Por fim, ao examinar os histogramas apresentados na Figura 6.40, observa-se que os modelos (a) e (c) apresentam distribuições dos resíduos que se aproximam de uma distribuição normal, sugerindo maior consistência nos erros de previsão. Em contraste, os modelos (b) e (d) exibem distribuições com cauda longa, indicando a presença de valores extremos nos resíduos.

6.2.6 N-Linear

Os modelos N-Linear apresentaram bons resultados no conjunto de teste, conforme mostrado na Tabela 6.3 e ilustrado na Figura 6.41. Apesar disso, apenas os modelos



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

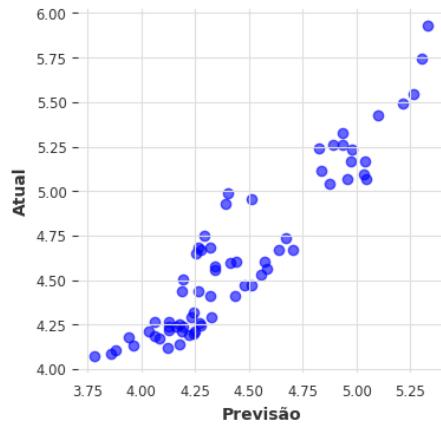


(c) Treinado com Biodiesel Nacional sem tendência

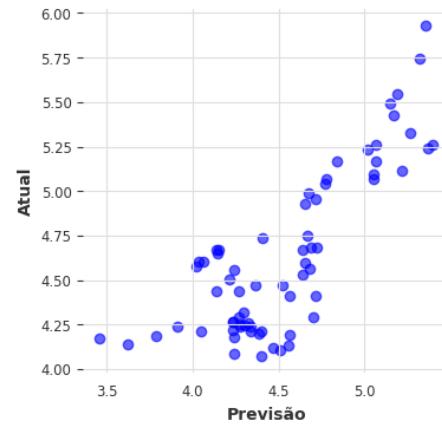


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

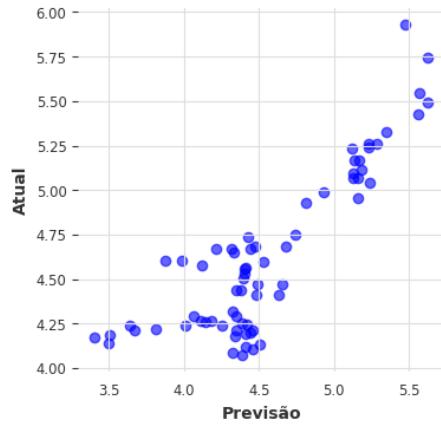
Figura 6.35: Resultados do modelo IMP Takens no conjunto de teste



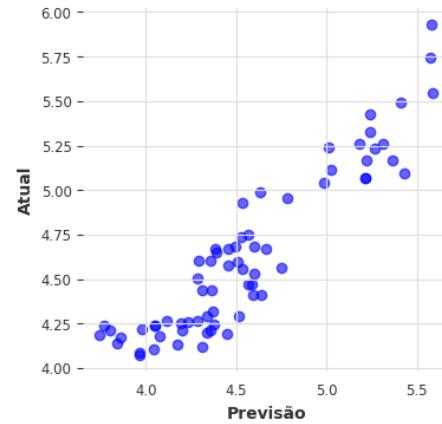
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

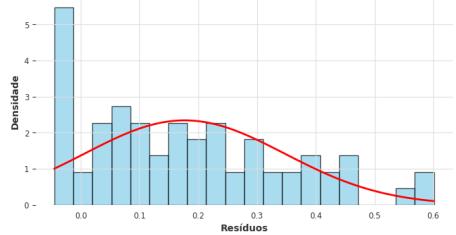


(c) Treinado com Biodiesel Nacional
sem tendência

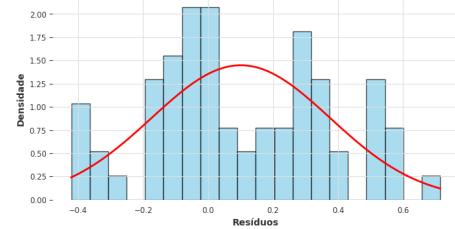


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

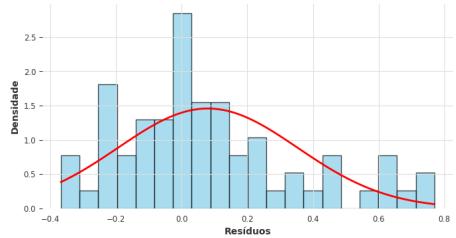
Figura 6.36: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo IMP Takens no conjunto de teste



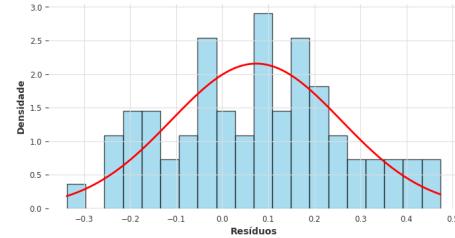
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.37: Histograma dos erros do modelo IMP Takens no conjunto de teste

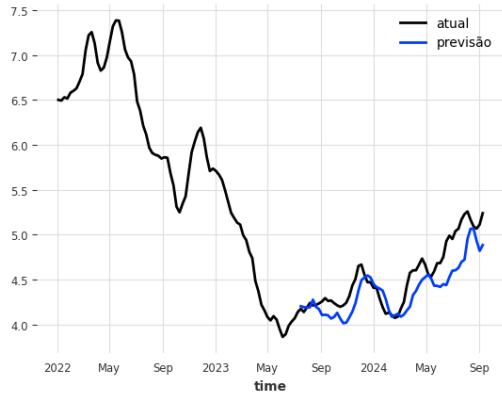
(b) e (c) apresentam resultados consistentes, mostrando que o modelo possui um número elevado de parâmetros e precisa de muitos dados para um treinamento efetivo.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura 6.42, observa-se que todos os modelos apresentam correlação com os valores reais. No entanto, o modelo (d) exibe maior dispersão para valores menores que 4,75, indicando maior variabilidade nos valores previstos nesse intervalo.

Por fim, ao examinar os histogramas apresentados na Figura 6.43, observa-se distribuições assimétricas e multimodais, sugerindo a presença de múltiplos picos nas distribuições dos resíduos, o que pode indicar complexidades nos erros de previsão.

6.2.7 NHITS

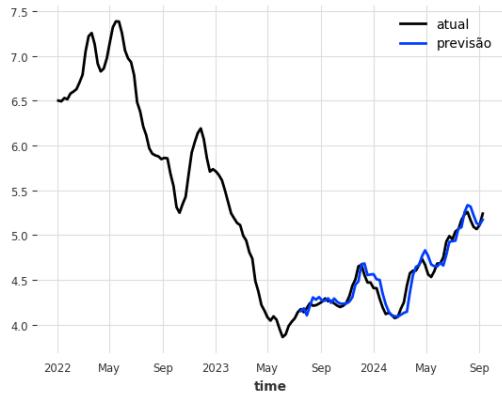
O modelo (b) obteve um bom resultado no conjunto de teste, como mostrado na Figura 6.44. Além disso, o modelo foi o de menor erro na validação cruzada



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

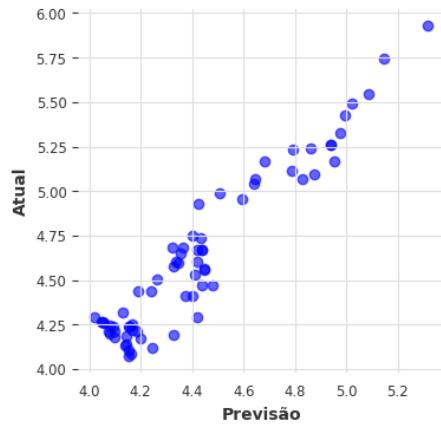


(c) Treinado com Biodiesel Nacional sem tendência

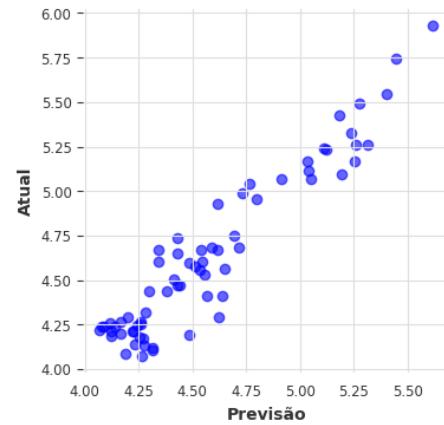


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

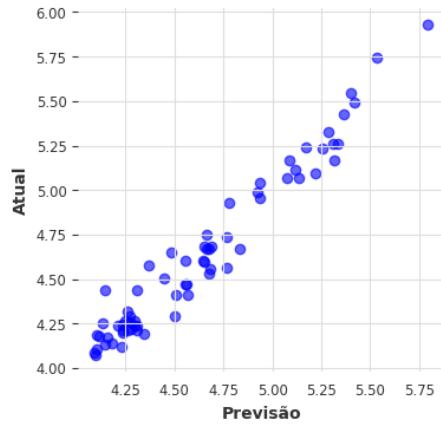
Figura 6.38: Resultados do modelo IDLN Takens no conjunto de teste



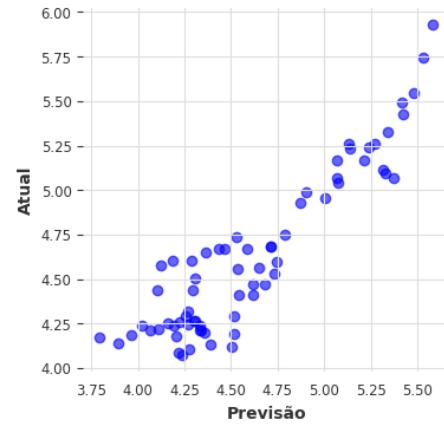
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

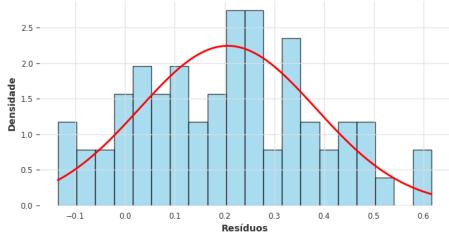


(c) Treinado com Biodiesel Nacional
sem tendência

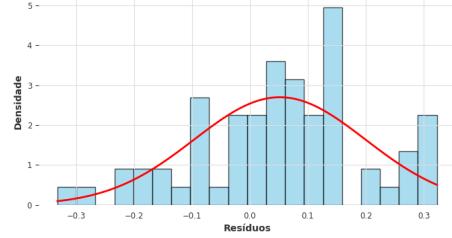


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

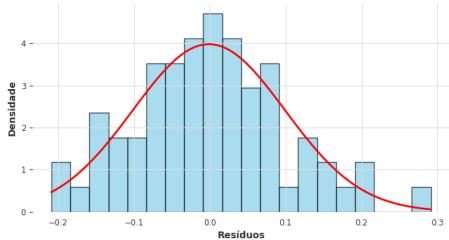
Figura 6.39: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo **IDLN** Takens no conjunto de teste



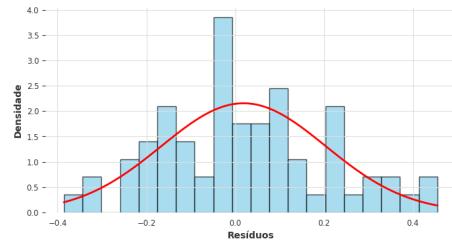
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja



(c) Treinado com Biodiesel Nacional sem tendência



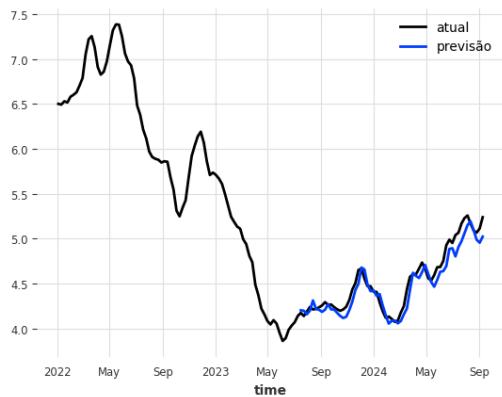
(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

Figura 6.40: Histograma dos erros do modelo IDLN Takens no conjunto de teste

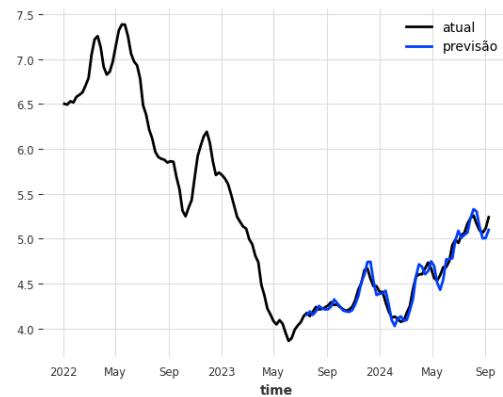
de forma consistente, como mostrado na Tabela 6.3, demonstrando a capacidade de generalização do modelo para conjuntos de tamanhos distintos.

Além disso, ao analisar os gráficos de dispersão apresentados na Figura 6.45, observa-se que todos os modelos apresentam correlação positiva com os valores reais. No entanto, os modelos (a), (c) e (d) exibem maior dispersão, indicando maior variabilidade nos valores previstos em comparação ao modelo (b).

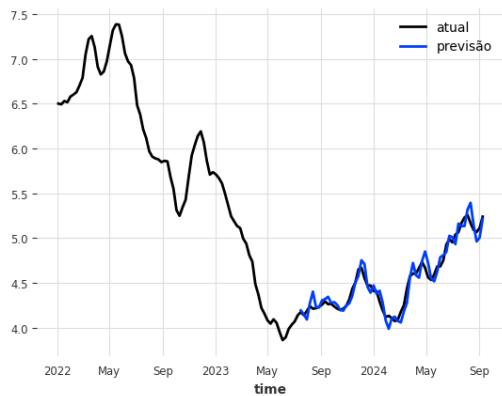
Por fim, ao examinar os histogramas apresentados na Figura 6.46, observa-se distribuições multimodais e com cauda longa, sugerindo a presença de valores extremos nos resíduos e uma estrutura complexa dos erros de previsão.



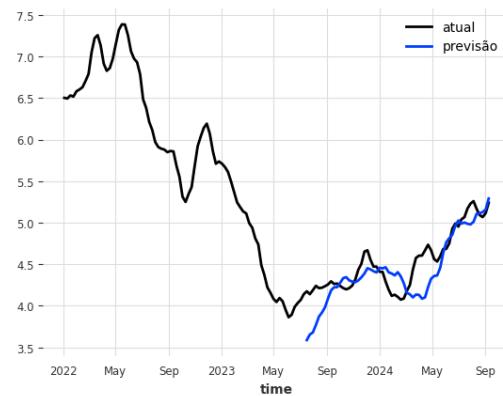
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

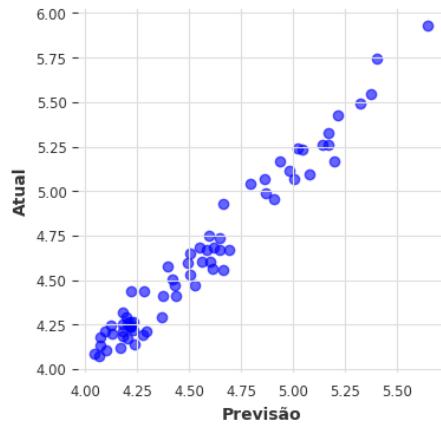


(c) Treinado com Biodiesel Nacional sem tendência

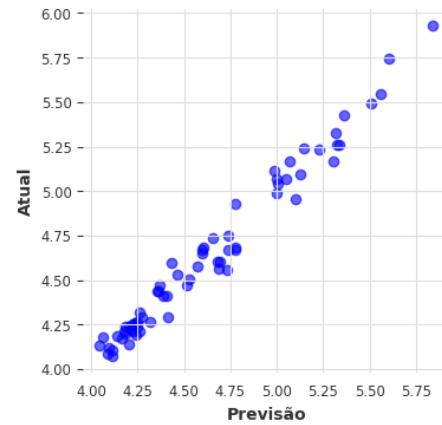


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

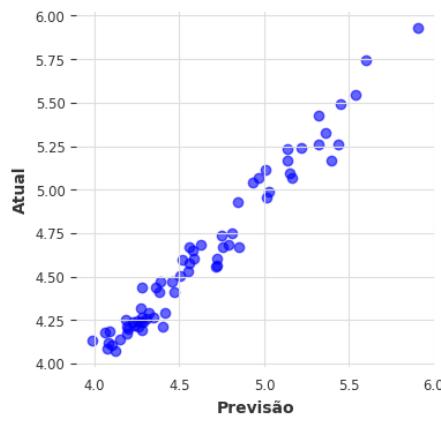
Figura 6.41: Resultados do modelo N-Linear Takens no conjunto de teste



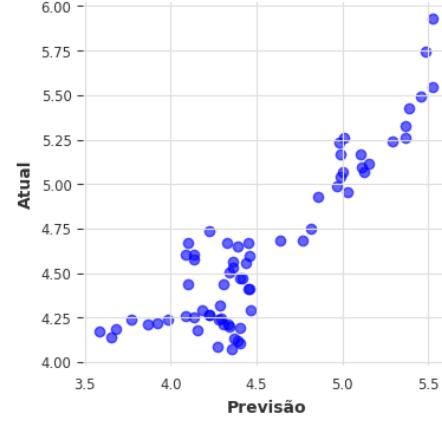
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja

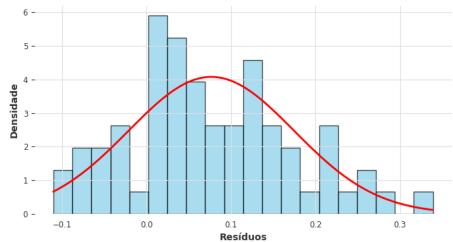


(c) Treinado com Biodiesel Nacional
sem tendência

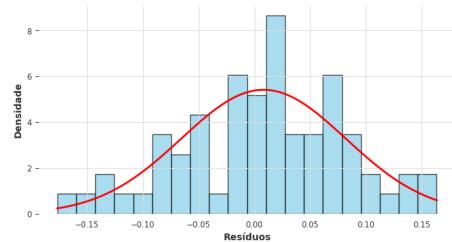


(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

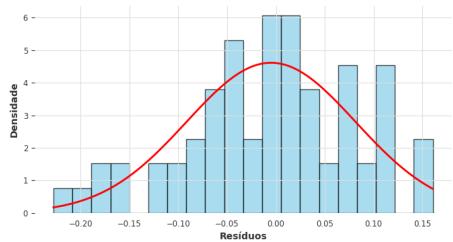
Figura 6.42: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo N-Linear Takens no conjunto de teste



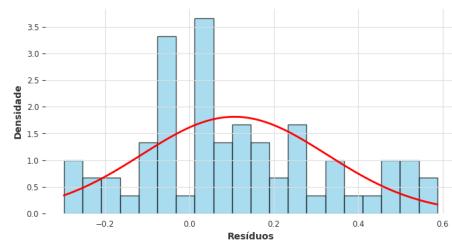
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja



(c) Treinado com Biodiesel Nacional
sem tendência



(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

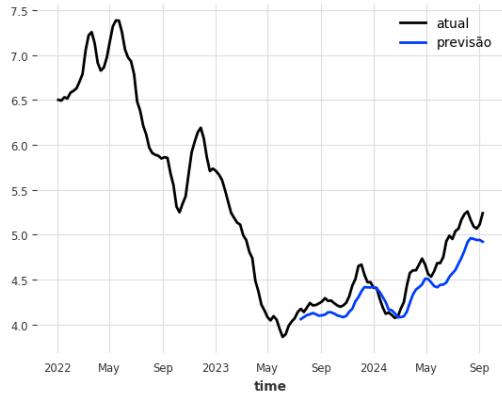
Figura 6.43: Histograma dos erros do modelo **N-Linear** Takens no conjunto de teste

Tabela 6.1: Resultados da Validação Cruzada usando Janela Deslizante

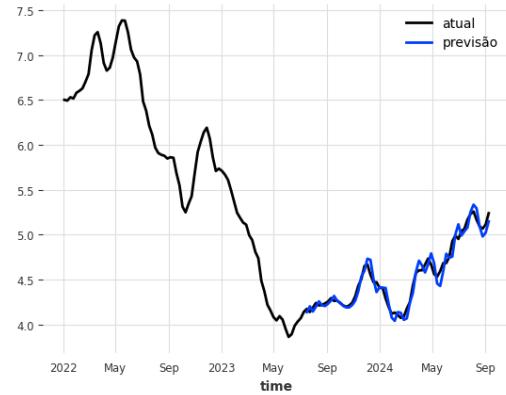
Modelo	Biodiesel Nacional						Biodiesel Nacional + Óleo de Soja					
	MAPE	SLE	MSE	RMSE	U_1	U_2	MAPE	SLE	MSE	RMSE	U_1	U_2
ARIMA	15,79	4,21	0,87	0,93	—	—	—	—	—	—	—	—
NARX	Original	5,72 ± 0,56	0,11 ± 0,04	0,15 ± 0,06	0,37 ± 0,08	0,97 ± 0,04	1,37 ± 0,49	1,55 ± 0,12	0,01 ± 0,00	0,10 ± 0,01	0,32 ± 0,18	1,21 ± 0,09
	Sem tendência	9,68 ± 5,26	0,49 ± 0,36	0,52 ± 0,43	0,59 ± 0,36	0,83 ± 0,37	2,84 ± 2,15	3,83 ± 0,28	0,06 ± 0,01	0,23 ± 0,02	0,45 ± 0,16	1,52 ± 0,22
DA-RNN	Original	5,30 ± 2,70	0,11 ± 0,09	0,16 ± 0,14	0,34 ± 0,20	0,70 ± 0,01	3,91 ± 0,97	1,93 ± 0,07	0,01 ± 0,00	0,12 ± 0,02	0,36 ± 0,17	1,80 ± 0,04
	Sem tendência	6,85 ± 4,14	0,25 ± 0,22	0,30 ± 0,28	0,45 ± 0,31	0,74 ± 0,37	2,73 ± 0,51	1,82 ± 0,33	0,01 ± 0,00	0,11 ± 0,03	0,18 ± 0,00	0,97 ± 0,14
MLP	Original	4,45 ± 2,23	0,10 ± 0,08	0,12 ± 0,11	0,30 ± 0,18	0,80 ± 0,10	1,74 ± 0,37	1,27 ± 0,02	0,01 ± 0,00	0,01 ± 0,00	0,08 ± 0,01	0,27 ± 0,13
	Sem tendência	5,92 ± 3,41	0,15 ± 0,13	0,19 ± 0,17	0,37 ± 0,23	0,60 ± 0,28	1,60 ± 0,68	1,55 ± 0,11	0,01 ± 0,00	0,01 ± 0,00	0,19 ± 0,01	0,48 ± 0,22
IMP	Original	9,42 ± 2,96	0,27 ± 0,14	0,38 ± 0,26	0,57 ± 0,23	0,87 ± 0,15	5,75 ± 1,24	4,15 ± 0,20	0,06 ± 0,01	0,07 ± 0,02	0,25 ± 0,05	0,67 ± 0,27
	Sem tendência	8,68 ± 1,78	0,39 ± 0,12	0,38 ± 0,22	0,59 ± 0,18	1,03 ± 0,14	1,84 ± 0,19	6,20 ± 0,02	0,14 ± 0,01	0,17 ± 0,04	0,41 ± 0,05	0,54 ± 0,05
IDLN	Original	4,99 ± 2,47	0,09 ± 0,07	0,13 ± 0,11	0,31 ± 0,17	0,73 ± 0,11	3,29 ± 2,15	5,29 ± 0,28	0,08 ± 0,00	0,10 ± 0,01	0,31 ± 0,02	0,77 ± 0,27
	Sem tendência	4,85 ± 2,31	0,11 ± 0,09	0,14 ± 0,12	0,32 ± 0,19	0,59 ± 0,20	2,06 ± 0,93	2,25 ± 0,11	0,02 ± 0,00	0,02 ± 0,00	0,14 ± 0,01	0,26 ± 0,03
N-Linear	Original	3,09 ± 1,56	0,05 ± 0,04	0,06 ± 0,05	0,21 ± 0,12	0,47 ± 0,03	2,28 ± 1,07	1,41 ± 0,13	0,01 ± 0,00	0,01 ± 0,00	0,09 ± 0,02	0,28 ± 0,14
	Sem tendência	3,74 ± 2,34	0,10 ± 0,09	0,12 ± 0,11	0,28 ± 0,20	0,36 ± 0,15	1,85 ± 1,38	1,33 ± 0,10	0,01 ± 0,00	0,01 ± 0,00	0,09 ± 0,02	0,18 ± 0,02
NHfS	Original	4,75 ± 0,91	0,09 ± 0,05	0,11 ± 0,07	0,31 ± 0,12	0,95 ± 0,05	2,05 ± 1,02	1,27 ± 0,03	0,01 ± 0,00	0,01 ± 0,00	0,09 ± 0,01	0,28 ± 0,15
	Sem tendência	6,85 ± 3,13	0,22 ± 0,16	0,25 ± 0,21	0,44 ± 0,24	0,82 ± 0,24	1,75 ± 0,70	1,38 ± 0,08	0,01 ± 0,00	0,01 ± 0,00	0,09 ± 0,01	0,18 ± 0,04

Tabela 6.2: Resultados do Conjunto de Teste usando Janela Deslizante

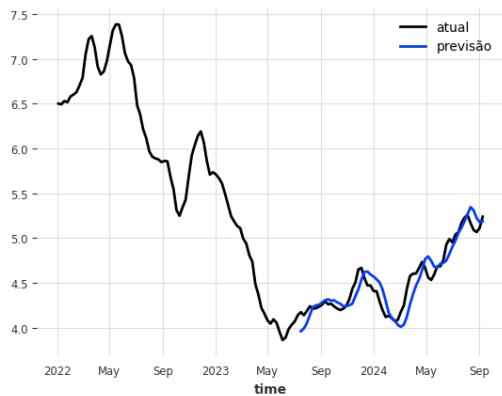
Modelo	Biodiesel Nacional						Biodiesel Nacional + Óleo de Soja						
	MAPE	SLE	MSE	RMSE	U_1	U_2	MAPE	SLE	MSE	RMSE	U_1	U_2	
NARX	Original	6,25	0,32	0,11	0,33	0,88	2,44	2,23	0,05	0,02	0,13	0,21	0,59
	Sem tendência	4,97	0,42	0,13	0,32	0,79	1,04	7,94	0,61	0,24	0,45	0,65	5,53
DA-RNN	Original	3,34	0,11	0,04	0,20	0,30	1,90	3,61	0,13	0,05	0,22	0,35	2,75
	Sem tendência	1,87	0,03	0,01	0,10	0,27	0,64	5,13	0,31	0,09	0,30	0,81	1,48
MLP	Original	3,12	0,09	0,03	0,18	0,32	2,00	6,91	0,40	0,14	0,37	0,69	2,10
	Sem tendência	1,58	0,02	0,01	0,09	0,24	0,21	1,62	0,02	0,01	0,09	0,23	0,52
RMP	Original	5,58	0,31	0,10	0,31	0,57	2,31	5,69	0,29	0,10	0,31	0,43	9,15
	Sem tendência	7,84	0,90	0,23	0,48	1,08	1,15	3,73	0,13	0,04	0,20	0,55	1,70
DLN	Original	3,42	0,10	0,04	0,19	0,32	1,08	7,70	0,85	0,21	0,46	0,58	3,23
	Sem tendência	1,66	0,03	0,01	0,10	0,27	0,17	3,42	0,11	0,04	0,19	0,52	0,45
N-Linear	Original	1,72	0,03	0,01	0,10	0,16	0,50	7,73	0,55	0,18	0,43	0,81	2,39
	Sem tendência	1,47	0,02	0,01	0,09	0,22	0,10	4,34	0,21	0,06	0,25	0,64	0,70
NHITS	Original	3,98	0,15	0,05	0,23	0,41	2,06	5,23	0,25	0,09	0,30	0,54	1,96
	Sem tendência	2,42	0,06	0,02	0,14	0,39	0,19	4,42	0,23	0,07	0,26	0,77	0,50



(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional e Óleo de Soja

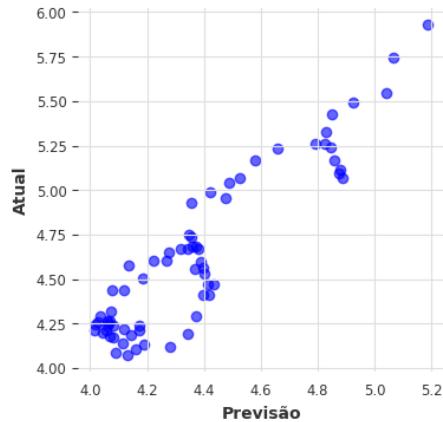


(c) Treinado com Biodiesel Nacional sem tendência

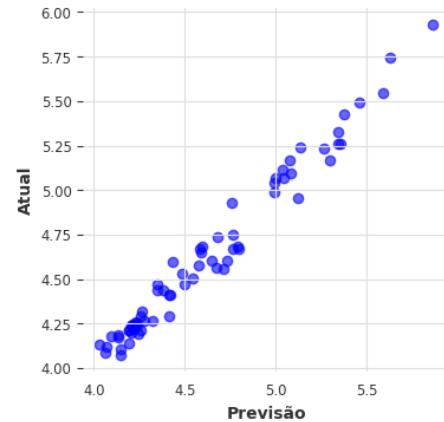


(d) Treinado com Biodiesel Nacional e Óleo de Soja sem tendência

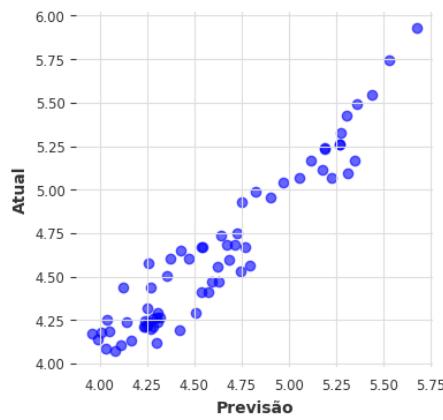
Figura 6.44: Resultados do modelo NHITS Takens no conjunto de teste



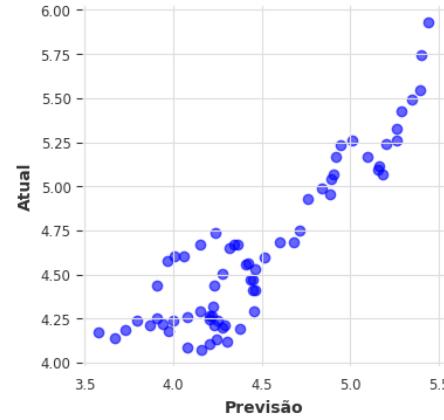
(a) Treinado com Biodiesel Nacional



(b) Treinado com Biodiesel Nacional
e Óleo de Soja



(c) Treinado com Biodiesel Nacional
sem tendência



(d) Treinado com Biodiesel Nacional
e Óleo de Soja sem tendência

Figura 6.45: Gráficos de dispersão dos valores atuais versus valores previstos pelo modelo NHITS Takens no conjunto de teste

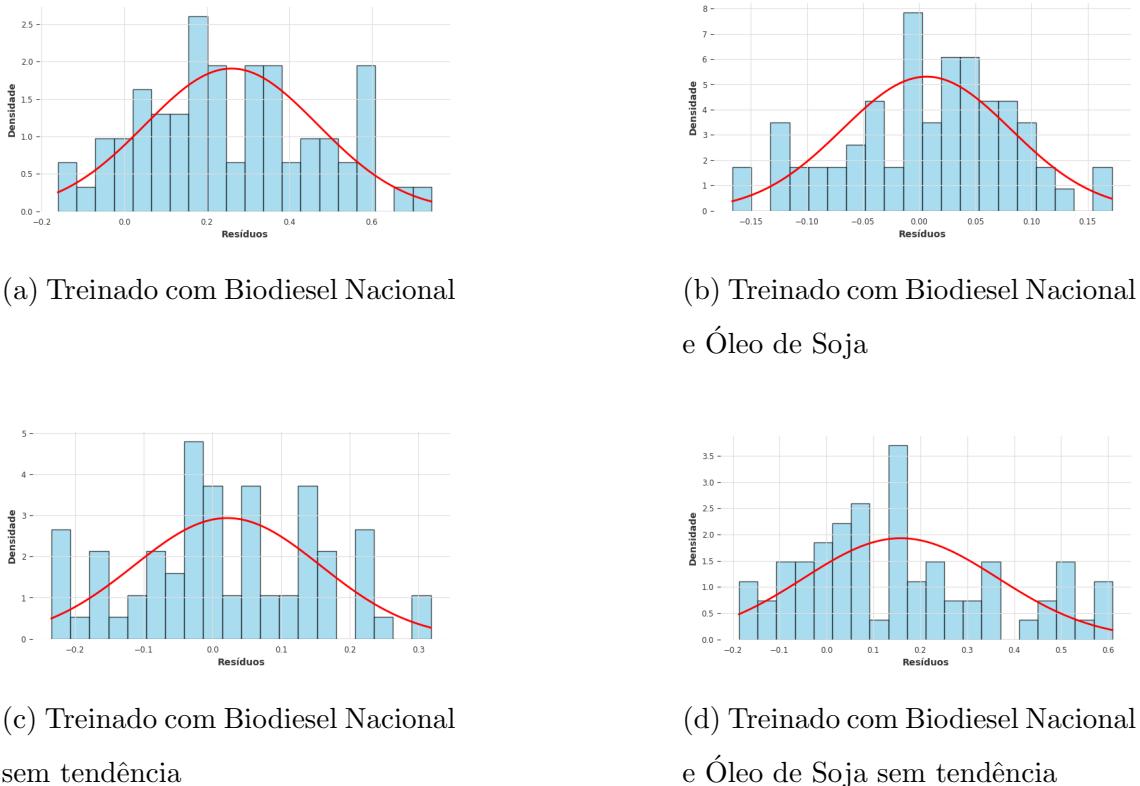
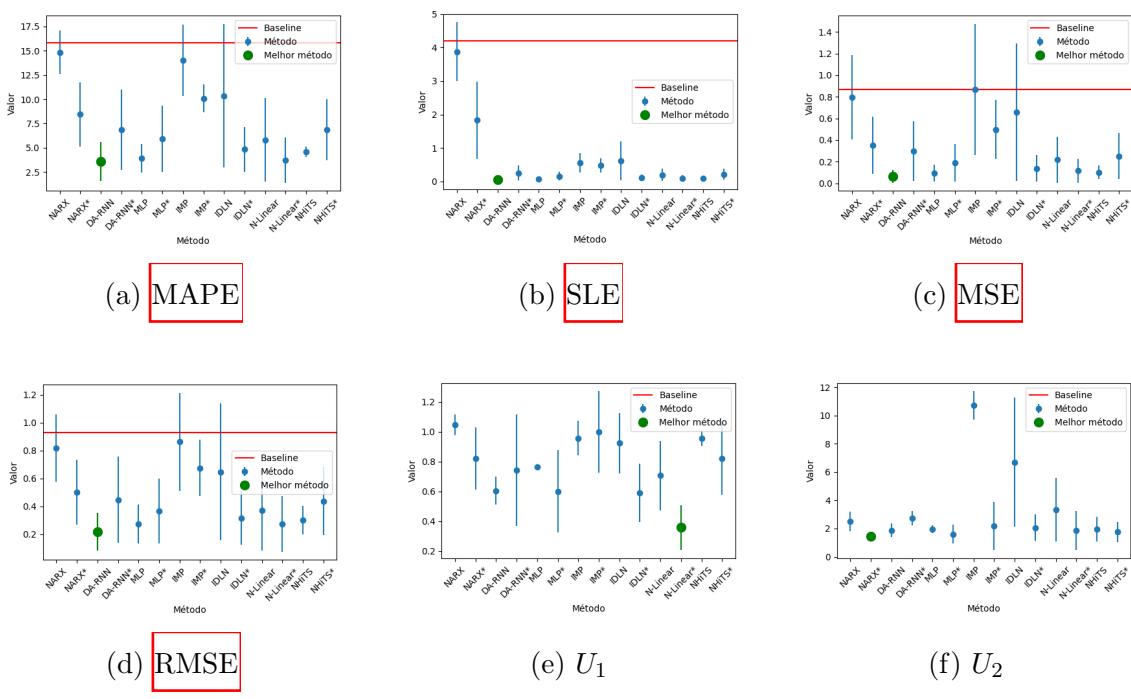


Figura 6.46: Histograma dos erros do modelo NHITS Takens no conjunto de teste



* sem tendência

Figura 6.47: Resultados da Validação Cruzada treinado com Biodiesel Nacional usando Janela de Takens

Tabela 6.3: Resultados da Validação Cruzada usando Janela de Takens

Modelo	Biodiesel Nacional						Biodiesel Nacional + Óleo de Soja						
	MAPE	SLE	MSE	RMSE	U_1	U_2	MAPE	SLE	MSE	RMSE	U_1	U_2	
ARIM A	15,79	4,21	0,87	0,93	—	—	—	—	—	—	—	—	
NARX	Original	14,85 ± 2,23	3,88 ± 0,89	0,80 ± 0,39	0,82 ± 0,24	1,05 ± 0,07	2,49 ± 0,70	1,51 ± 0,12	0,01 ± 0,00	0,10 ± 0,01	0,31 ± 0,16	0,79 ± 0,09	
	Sem tendência	8,45 ± 3,29	1,83 ± 1,15	0,35 ± 0,26	0,50 ± 0,23	0,82 ± 0,21	1,43 ± 0,28	6,65 ± 0,13	0,73 ± 0,04	0,19 ± 0,05	0,40 ± 0,05	0,57 ± 0,08	6,92 ± 5,53
DA-RNN	Original	3,58 ± 2,01	0,05 ± 0,04	0,07 ± 0,06	0,22 ± 0,14	0,60 ± 0,09	1,87 ± 0,48	2,05 ± 0,15	0,02 ± 0,00	0,13 ± 0,01	0,40 ± 0,17	1,95 ± 0,56	
	Sem tendência	6,85 ± 4,14	0,25 ± 0,22	0,30 ± 0,28	0,45 ± 0,31	0,74 ± 0,37	2,73 ± 0,51	1,82 ± 0,33	0,01 ± 0,00	0,01 ± 0,01	0,11 ± 0,03	0,18 ± 0,00	0,97 ± 0,14
MLP	Original	3,94 ± 1,46	0,08 ± 0,06	0,10 ± 0,08	0,27 ± 0,14	0,76 ± 0,01	1,93 ± 0,28	1,33 ± 0,00	0,01 ± 0,00	0,01 ± 0,00	0,09 ± 0,01	0,28 ± 0,14	1,07 ± 0,04
	Sem tendência	5,92 ± 3,41	0,15 ± 0,13	0,19 ± 0,17	0,37 ± 0,23	0,60 ± 0,28	1,60 ± 0,68	1,55 ± 0,11	0,01 ± 0,00	0,01 ± 0,00	0,10 ± 0,01	0,19 ± 0,05	0,48 ± 0,22
IMP	Original	14,03 ± 3,68	0,55 ± 0,29	0,87 ± 0,61	0,86 ± 0,35	0,96 ± 0,11	10,72 ± 1,01	3,21 ± 0,81	0,03 ± 0,01	0,04 ± 0,03	0,20 ± 0,07	0,55 ± 0,20	1,09 ± 0,19
	Sem tendência	10,10 ± 1,42	0,47 ± 0,21	0,50 ± 0,27	0,68 ± 0,20	1,00 ± 0,27	2,19 ± 1,70	6,20 ± 0,02	0,14 ± 0,01	0,17 ± 0,04	0,41 ± 0,05	0,54 ± 0,01	1,80 ± 0,77
IDLN	Original	10,37 ± 7,41	0,61 ± 0,58	0,66 ± 0,63	0,65 ± 0,49	0,92 ± 0,20	6,69 ± 4,57	4,41 ± 0,73	0,06 ± 0,02	0,06 ± 0,01	0,25 ± 0,01	0,65 ± 0,33	3,66 ± 1,23
	Sem tendência	4,85 ± 2,31	0,11 ± 0,09	0,14 ± 0,12	0,32 ± 0,19	0,59 ± 0,20	2,06 ± 0,93	2,25 ± 0,11	0,02 ± 0,00	0,02 ± 0,00	0,14 ± 0,01	0,26 ± 0,03	0,76 ± 0,33
N-Linear	Original	5,80 ± 4,31	0,19 ± 0,18	0,22 ± 0,21	0,37 ± 0,29	0,71 ± 0,23	3,32 ± 2,24	1,36 ± 0,11	0,01 ± 0,00	0,09 ± 0,02	0,28 ± 0,14	1,07 ± 0,02	
	Sem tendência	3,74 ± 2,34	0,10 ± 0,09	0,12 ± 0,11	0,28 ± 0,20	0,36 ± 0,15	1,85 ± 1,38	1,33 ± 0,10	0,01 ± 0,00	0,09 ± 0,02	0,18 ± 0,02	0,31 ± 0,14	
NHITS	Original	4,60 ± 0,56	0,08 ± 0,04	0,10 ± 0,06	0,30 ± 0,10	0,95 ± 0,05	1,97 ± 0,87	1,32 ± 0,01	0,01 ± 0,00	0,01 ± 0,00	0,09 ± 0,01	0,27 ± 0,13	1,10 ± 0,08
	Sem tendência	6,85 ± 3,13	0,22 ± 0,16	0,25 ± 0,21	0,44 ± 0,24	0,82 ± 0,24	1,75 ± 0,70	1,38 ± 0,08	0,01 ± 0,00	0,01 ± 0,01	0,09 ± 0,01	0,18 ± 0,04	0,50 ± 0,10

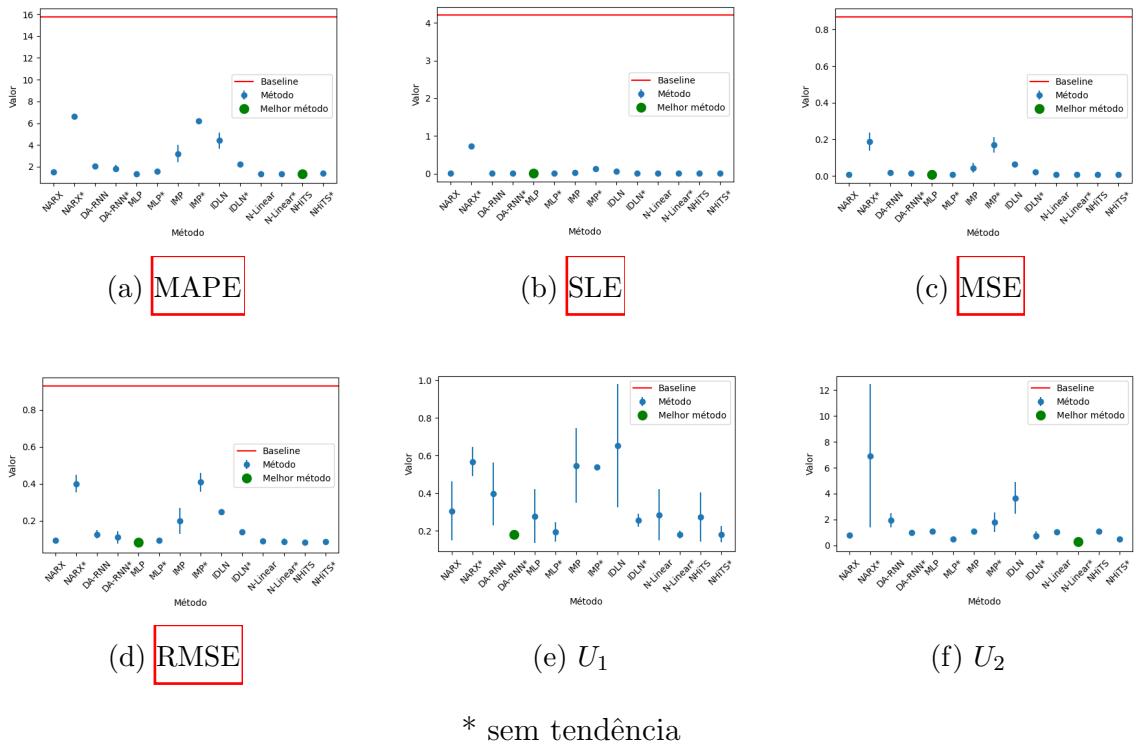


Figura 6.48: Resultados da Validação Cruzada treinado com Biodiesel Nacional e Óleo de Soja usando Janela de Takens

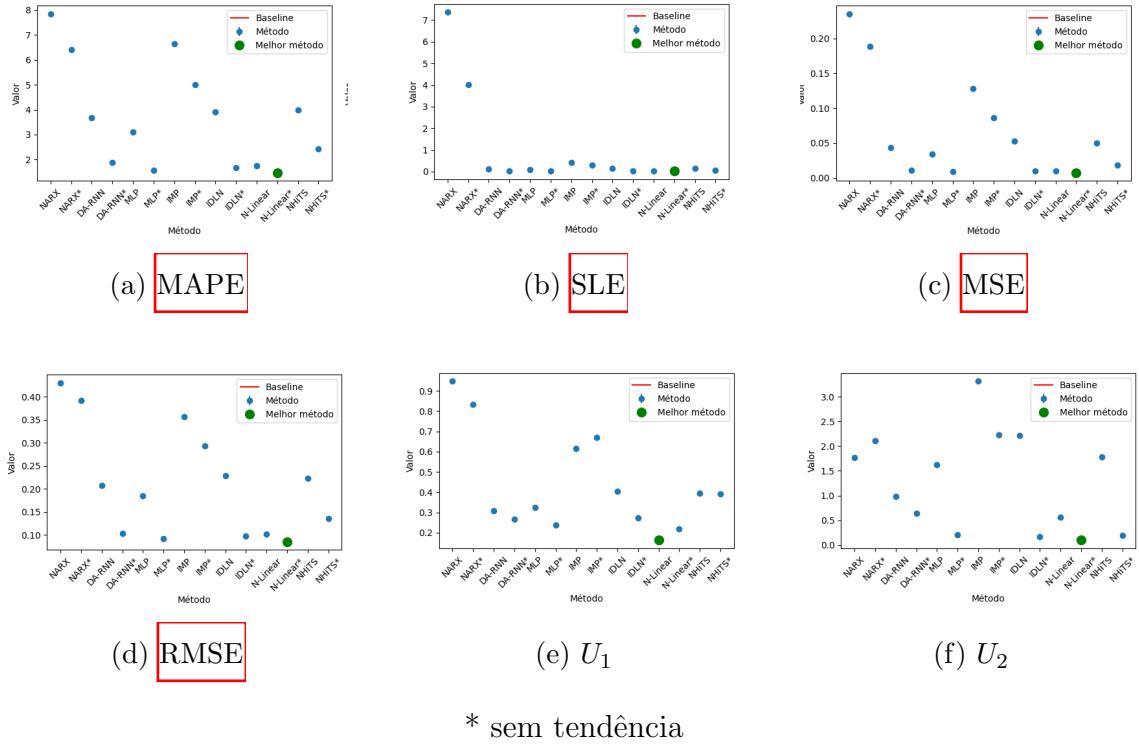
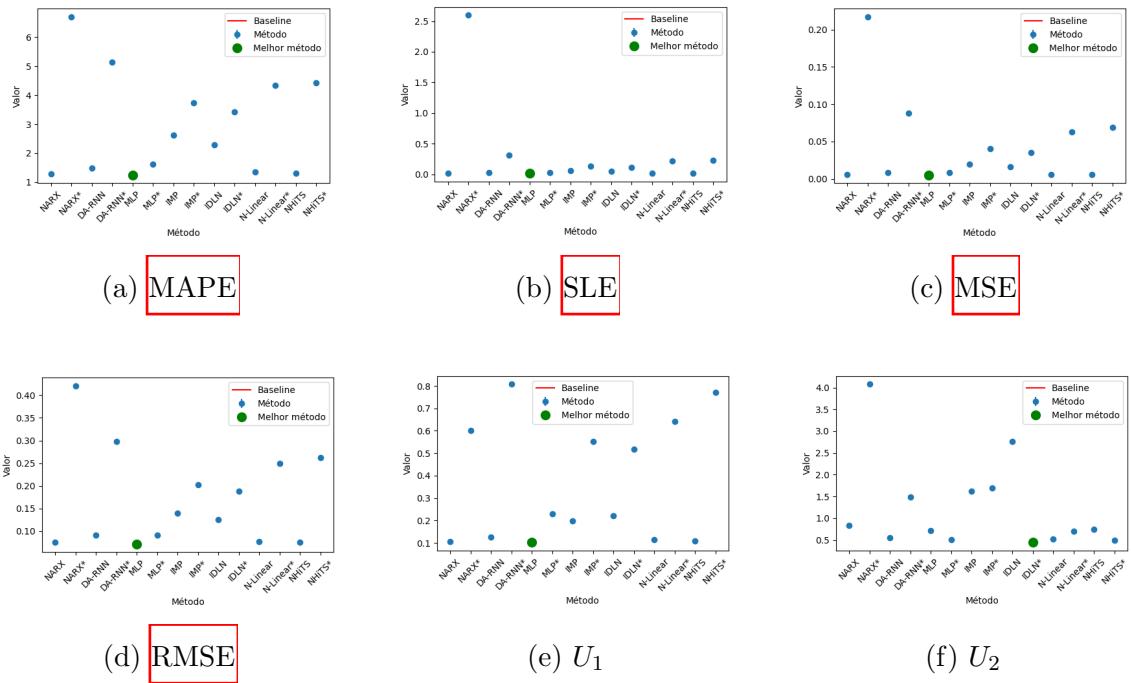


Figura 6.49: Resultados do Conjunto de Teste treinado com Biodiesel Nacional usando Janela de Takens

Tabela 6.4: Resultados do Conjunto de Teste usando Janela de Takens

Modelo	Biodiesel Nacional						Biodiesel Nacional + Óleo de Soja					
	MAPE	SLE	MSE	RMSE	U_1	U_2	MAPE	SLE	MSE	RMSE	U_1	U_2
NARX	Original	7,85	7,37	0,24	0,43	0,95	1,77	1,28	0,02	0,01	0,08	0,11
	Sem tendência	6,40	4,01	0,19	0,39	0,83	2,11	6,71	2,60	0,22	0,42	0,60
DA-RNN	Original	3,68	0,12	0,04	0,21	0,31	0,98	1,48	0,02	0,01	0,09	0,13
	Sem tendência	1,87	0,03	0,01	0,10	0,27	0,64	5,13	0,31	0,09	0,30	0,81
MLP	Original	3,10	0,10	0,03	0,19	0,32	1,63	1,25	0,01	0,01	0,07	0,10
	Sem tendência	1,58	0,02	0,01	0,09	0,24	0,21	1,62	0,02	0,01	0,09	0,23
RMP	Original	6,64	0,41	0,13	0,36	0,62	3,32	2,61	0,05	0,02	0,14	0,20
	Sem tendência	4,99	0,31	0,09	0,29	0,67	2,23	3,73	0,13	0,04	0,20	0,55
DLN	Original	3,92	0,15	0,05	0,23	0,41	2,22	2,28	0,05	0,02	0,13	0,22
	Sem tendência	1,66	0,03	0,01	0,10	0,27	0,17	3,42	0,11	0,04	0,19	0,52
N-Linear	Original	1,76	0,03	0,01	0,10	0,16	0,56	1,35	0,02	0,01	0,08	0,11
	Sem tendência	1,47	0,02	0,01	0,09	0,22	0,10	4,34	0,21	0,06	0,25	0,64
NHITS	Original	4,00	0,14	0,05	0,22	0,39	1,79	1,31	0,02	0,01	0,08	0,11
	Sem tendência	2,42	0,06	0,02	0,14	0,39	0,19	4,42	0,23	0,07	0,26	0,77



* sem tendência

Figura 6.50: Resultados do Conjunto de Teste treinado com Biodiesel Nacional e Óleo de Soja usando Janela de Takens

Capítulo 7

Conclusões e Trabalhos Futuros

Este estudo investiga a dinâmica subjacente à série temporal semanal dos preços do biodiesel, influenciada majoritariamente pelo custo das matérias-primas e pela regulamentação vigente, que impacta diretamente a demanda.

Com base na avaliação dos resultados, observa-se que, ao utilizar uma janela deslizante, a aplicação de técnicas de normalização e decomposição resultou em perda de consistência nos modelos neurais, exceto nos modelos morfológicos (**IMP** e **IDLN**), além de um aumento nos erros preditivos. Esses resultados sugerem que modelos baseados em decomposição linear não são adequados para essa série temporal. Ainda assim, o modelo **N-Linear** demonstrou desempenho superior em comparação aos demais, tanto no conjunto de teste quanto no de validação, quando o treinamento foi realizado exclusivamente com a série de preços do biodiesel.

Ao incluir a série de preços do óleo de soja como variável exógena, o modelo **NHiTS** apresentou o melhor desempenho, como esperado devido à sua arquitetura profunda, que requer um maior volume de dados para treinamento. Contudo, dada sua complexidade, o modelo não mostrou uma vantagem substancial em relação ao **MLP**. No conjunto de teste, o **MLP** demonstrou alta precisão, enquanto o **NHiTS** apresentou uma piora em relação ao desempenho observado na validação.

Na análise dos modelos com janela de Takens, verificou-se novamente a perda de consistência ao aplicar a normalização quando o treinamento foi realizado somente com a série do biodiesel. Além disso, os valores da medida U_2 excederam 1 em todos

os casos, indicando que os modelos analisados tiveram um desempenho inferior ao de um passeio aleatório. Apesar disso, os modelos **N-Linear** e **MLP** obtiveram bons resultados no conjunto de teste, demonstrando a capacidade de lidar com conjuntos de dados de tamanho reduzido.

Ao repetir a análise incluindo a série de preços do óleo de soja como covariável, observou-se uma melhora na consistência dos erros. Destacam-se, com baixos valores de erro, os modelos **IDLN** e **N-Linear**, ambos com valores reduzidos de U_2 . No conjunto de teste, ambos mantiveram o bom desempenho, mas o modelo **MLP** merece destaque por superar os demais em todas as medidas de desempenho.

A partir da análise realizada, observa-se que os modelos de redes neurais feed-forward e redes profundas demonstraram uma melhoria significativa com a inclusão de covariáveis, evidenciando sua capacidade de lidar com grandes volumes de dados. No entanto, nos testes, essa diferença foi observada em uma escala menor. As redes **NARX** e **DA-RNN** revelaram-se sensíveis ao conjunto de dados, exibindo variações consideráveis quando covariáveis foram adicionadas, em alguns casos apresentando melhores resultados com a série isolada. As redes morfológicas, por sua vez, mostraram eficiência com a janela deslizante, mas apresentaram desempenho inferior ao utilizar a janela de Takens.

Entre os modelos avaliados, a rede **N-Linear** destacou-se como a mais recomendada, pois apresentou melhor desempenho com a janela deslizante e manteve resultados consistentes ao utilizar a janela de Takens, evidenciando sua robustez em diferentes configurações de modelagem.

Como trabalhos futuros, propõe-se o aprimoramento dos modelos com a inclusão de variáveis exógenas adicionais, como séries temporais de produção e demanda de biodiesel, preços de outras matérias-primas e o percentual obrigatório de mistura de biodiesel no diesel, para avaliar o impacto nos preços finais. Além disso, sugere-se a incorporação de modelos baseados em Processamento de Linguagem Natural (do inglês *Natural Language Processing*, **NLP**) para análise do impacto de notícias no mercado. Recomenda-se também a aplicação de métodos de compactação de sinais,

como Análise de Componentes Principais (do inglês *Principal Component Analysis*, **PCA**) e Análise de Componentes Principais Não Lineares (do inglês *Nonlinear Principal Component Analysis*, **NLPCA**), para reduzir a dimensionalidade dos dados e analisar o comportamento dos modelos.

Adicionalmente, propõe-se a inclusão de Redes Neurais com Memória de Longo e Curto Prazo (do inglês *Long Short-Term Memory*, **LSTMs**) menores, com o objetivo de reduzir a complexidade dos modelos e melhorar sua generalização. Sugere-se também a aplicação de métodos alternativos para extração de tendências em séries temporais, bem como a introdução de índices setoriais, como o índice agropecuário e o dólar, para uma análise mais abrangente do mercado. Além disso, destaca-se a importância da análise e ajuste dos modelos com base em histogramas e gráficos de dispersão, o que pode fornecer análises adicionais e aprimorar a qualidade das previsões.

Referências Bibliográficas

- [1] BOX, G. E. P., JENKINS, G. M., REINSEL, G. C., *et al.*, *Time Series Analysis: Forecasting and Control*. 5 ed. Hoboken, NJ, Wiley, 2016.
- [2] HYNDMAN, R., ATHANASOPOULOS, G., *Forecasting: Principles and Practice*. 2 ed. Melbourne, Australia, OTexts, 2018. Acesso em: 20/06/2024.
- [3] MGHAZLI, Y.-S. B., *Redes Neurais Morfológicas Profundas para Previsão de Séries Temporais Financeiras*. M.Sc. dissertation, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2023.
- [4] SOUZA, R. M. P., “O que é série temporal?”, Março 2022, Disponível em: <https://estatjr.com.br/2022/03/25/o-que-e-serie-temporal/>, Acesso em: 20/06/2024.
- [5] ALVES, L., BARROS, G. S. D. C., “Desafios da precificação no mercado de biodiesel brasileiro”, Feb. 2022, Disponível em: <https://www.cepea.esalq.usp.br/br/opiniao-cepea/desafios-da-precificacao-no-mercado-de-biodiesel-brasileiro-1.aspx>.
- [6] SCIENCE, A., “Qual a importância da energia renovável”, 2024, Disponível em: <https://ambscience.com/energia-renovavel/>.
- [7] HAYKIN, S., *Neural networks and learning machines*, v. 3. Pearson Upper Saddle River, NJ, USA, 2009.
- [8] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Ministério de Minas e Energia, Governo Federal do Brasil, “Preços de produtores e importadores de derivados de petróleo e biodiesel”, 2024, Disponível em: <https://www.gov.br/anp/pt-br/assuntos/precos-e-defesa-da>

concorrencia/precos/precos-de-produtores-e-importadores-de-derivados-de-petroleo-e-biodiesel.

- [9] World Bank, “Historical Data: Monthly”, 2024, Disponível em: <https://thedocs.worldbank.org/en/doc/5d903e848db1d1b83e0ec8f744e55570-0350012021/related/CMO-Historical-Data-Monthly.xlsx>, Acesso em: 31 jul. 2024.
- [10] FAIER, J. M., *Análise de Componentes Independentes para Monitoração da Qualidade de Dados em Séries Temporais*. Ph.D. dissertation, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011. Acesso em 20 de maio de 2023.
- [11] DANTAS, A. C. H., *Sistema de Monitoração de Qualidade de Dados*. Ph.D. dissertation, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007. Acesso em 20 de maio de 2023.
- [12] HAYKIN, S., *Neural Networks: A Comprehensive Foundation*. New Jersey, Prentice Hall, 1998.
- [13] ZHENG, Z., ZHANG, Y., LIU, W., et al., “Forecasting with deep learning: A review”, *arXiv preprint arXiv:1704.02971*, , 2017.
- [14] ARAÚJO, R. A., “A Morphological Perceptron with Gradient-based Learning for Brazilian Stock Market Forecasting”, *Neural Networks*, , 2012.
- [15] ARAÚJO, R. D. A., *Mercado de Ações Brasileiro em Alta-Freqüência: Evidências de sua Previsibilidade com Modelagem Morfológica-Linear*. Ph.D. dissertation, Universidade Federal de Pernambuco, Recife, 2016. Acesso em 20 de maio de 2023.
- [16] ZENG, A., CHEN, M., ZHANG, L., et al., “Are Transformers Effective for Time Series Forecasting?”, *arXiv*, v. 3, 2022.
- [17] CHALLU, C., OLIVARES, K. G., ORESHKIN, B. N., et al., “N-HiTS: Neural Hierarchical Interpolation for Time Series Forecasting”, *arXiv*, v. 6, 2022.
- [18] LIAW, R., LIANG, E., NISHIHARA, R., et al., “Tune: A Research Platform for Distributed Model Selection and Training”, *arXiv preprint arXiv:1807.05118*, , 2018.

- [19] AKIBA, T., SANO, S., YANASE, T., *et al.*, “Optuna: A Next-generation Hyperparameter Optimization Framework”. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [20] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Ministério de Minas e Energia, Governo Federal do Brasil, “Biodiesel”, 2024, Disponível em: <https://www.gov.br/anp/pt-br/assuntos/producao-e-fornecimento-de-biocombustiveis/biodiesel>.
- [21] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Ministério de Minas e Energia, Governo Federal do Brasil, “Especificação do biodiesel”, 2024, Disponível em: <https://www.gov.br/anp/pt-br/assuntos/producao-e-fornecimento-de-biocombustiveis/biodiesel/biodiesel/especificacao-do-biodiesel>.
- [22] SILVA, F. R. D., SILVA, F. J. L. B. D., SANTOS, R. D. S., *et al.*, “O biodiesel no Brasil: Uma análise da produção, consumo e perspectivas na transição energética”, *Research, Society and Development*, v. 12, n. 11, pp. e43121143670, 2023. CC BY 4.0.
- [23] TotalEnergies, “O que é Biodiesel? Vantagens e Desvantagens”, 2024, Disponível em: <https://totalenergies.com.br/o-que-e-biodiesel-vantagens-e-desvantagens>, Acesso em: 20 ago. 2024.
- [24] InfoMoney, “O que é uma cesta básica?”, 2024, Disponível em: <https://www.infomoney.com.br/consumo/o-que-e-uma-cesta-basica/>, Acesso em: 20 ago. 2024.
- [25] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Ministério de Minas e Energia, Governo Federal do Brasil, “Boletim Mensal de Energia - Dezembro/2023”, 2024, Disponível em: <https://www.gov.br/mme/pt-br/assuntos/secretarias/sntep/publicacoes/boletins-mensais-de-energia/boletins/2023-1/portugues>.
- [26] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Ministério de Minas e Energia, Governo Federal do Brasil, “Boletim Mensal de Ener-

gia - Janeiro/2024”, 2024, Disponível em: <https://www.gov.br/mme/pt-br/assuntos/secretarias/sntep/publicacoes/boletins-mensais-de-energia/boletins/2024>.

- [27] EPBR, “Biodiesel: entenda o novo modelo de comercialização aprovado pela ANP”, 2021, Disponível em: <https://epbr.com.br/biodiesel-entenda-o-novo-modelo-de-comercializacao-aprovado-pela-anp/>, Acesso em: 31 jul. 2024.
- [28] Reuters, “Com oferta limitada, Brasil reduz mistura de biodiesel para novembro e dezembro”, *G1 Globo*, , 10 2020.
- [29] Presidência da República, “Lei nº 14.993, de 8 de outubro de 2024”, 2024, Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2024/lei/L14993.htm.
- [30] Agência Senado, “Lei que incentiva ”combustíveis do futuro”é sancionada com vetos”, 2024, Disponível em: <https://www12.senado.leg.br/noticias/materias/2024/10/10/lei-que-incentiva-combustiveis-do-futuro-e-sancionada-com-vetos>, Acesso em: 8 nov. 2024.
- [31] Canal Rural, “Biodiesel: setor diz que está pronto para atender mistura de 20%”, 2024, Disponível em: <https://www.canalrural.com.br/economia/biodiesel-setor-diz-que-esta-pronto-para-atender-mistura-de-20/>, Acesso em: 20 ago. 2024.
- [32] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Ministério de Minas e Energia, Governo Federal do Brasil, “Resolução ANP nº 910, de 18 de novembro de 2022”, 2022, Disponível em: <https://atosoficiais.com.br/anp/resolucao-n-910-2022-dispoe-sobre-a-previa-anuencia-da-anp-para-o-uso-experimental-ou-especifico-de-biodiesel-ou-de-sua-mistura-com-oleo-diesel-a-oleo-diesel-bx-em-quantidade-superior-ao-percentual-de-adicao-de-biodiesel-obrigatoria>.
- [33] ROSENBLATT, F., “The perceptron: A probabilistic model for information storage and organization in the brain”, *Psychological Review*, v. 65, n. 6, pp. 386–408, 1958.

- [34] MINSKY, M., PAPERT, S., *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA, MIT Press, 1969.
- [35] RUMELHART, D. E., HINTON, G. E., WILLIAMS, R. J., “Learning representations by back-propagating errors”. In: *Nature*, Springer, pp. 533–536, 1986.
- [36] VASCONCELOS, M., “Distribuição normal e classe Standard Scaler da biblioteca Scikit-learn em Python. Por que padronizar/normalizar os dados?”, 2024, Disponível em: <https://medium.com/@mhvasconcelos/distribui%C3%A7%C3%A3o-normal-e-a-biblioteca-standard-scaler-em-python-f21c52070c6b>, Acesso em: 20 ago. 2024.
- [37] GEEKSFORGEEKS, “Data Pre-processing with Sklearn Using Standard and MinMax Scaler”, 2023, Disponível em: <https://www.geeksforgeeks.org/data-pre-processing-wit-sklearn-using-standard-and-minmax-scaler/>, Acessado em: 19-agosto-2024.
- [38] ORACLE, “MAPE: Mean Absolute Percentage Error”, Disponível em: https://docs.oracle.com/cloud/help/pt_BR/pbcs_common/PFUSU/insights_metrics.MAPE.html GUID-C33B0F01-83E9-468B-B96C-413A12882334, Acessado em 19 de agosto de 2024.
- [39] FILHO, M., “MAPE: Erro Absoluto Percentual Médio em Machine Learning”, 2023, Disponível em: <https://mariofilho.com/mape-erro-absoluto-percentual-medio-em-machine-learning/#como-interpretar-o-mape>, Acessado em 19 de agosto de 2024.
- [40] HACKERS, D., “Prevendo Números: Entendendo Métricas de Regressão”, 2023, Disponível em: <https://medium.com/data-hackers/prevendo-n>
- [41] THEIL, B., “Understanding the Construction and Interpretation of Forecast Evaluation Statistics Using Computer-Based Tutorial Exercises”, *Journal of Forecasting*, v. 11, n. 2, 1992.
- [42] HURST, H. E., “Long-term storage capacity of reservoirs”, *Transactions of the American Society of Civil Engineers*, v. 116, pp. 770–799, 1951.

- [43] Presidência da República, “Lei nº 12.527, de 18 de novembro de 2011”, 2011, Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm.
- [44] SAHA, D., “Machine Learning: Training Set, Validation Set, and Test Set”, 2024, Disponível em: <https://medium.com/@dipan.saha/machine-learning-training-set-validation-set-and-test-set-48d10ca389db>, Acesso em: 2024-06-20.

Apêndice A

Arquitetura dos Modelos para Previsão de Séries Temporais

Este apêndice contém a descrição de algumas arquiteturas de rede neural usadas neste trabalho, além de descrever os hiperparâmetros de cada rede.

Alguns hiperparâmetros são comuns a todos os modelos, são eles:

- **Comprimento da Janela de Entrada** : Número de instantes de tempo passados usados como entrada para o modelo.
- **Comprimento da Janela de Saída** : Número de passos à frente que o modelo prevê.
- **Comprimento do Lote** : especifica o número de amostras processadas simultaneamente durante o treinamento, o que afeta tanto a eficiência computacional quanto a qualidade da convergência.
- **Número de Épocas** : Número de vezes que o modelo passa por todo o conjunto de treinamento.
- **Taxa de Aprendizado**: Define a rapidez com que o modelo ajusta os pesos das conexões durante o treinamento.
- **Otimizador** : Algoritmo que ajusta os pesos das conexões durante o treinamento.

- **Função de Perda** : Função que calcula a diferença entre a previsão e o valor real.

A.1 Rede Neural Autoregressiva Não-linear com Entradas Exógenas (do inglês *Nonlinear Autoregressive Exogenous Model, NARX*)

O modelo **NARX** é uma classe de modelos não lineares, onde a saída $y(n+1)$ é uma função dos valores de entrada e dos valores atrasados da saída, conforme mostra a Equação A.1 [12].

$$y(n+1) = f^l(x(n), x(n-1), \dots, x(n-p), y(n-q), \dots, y(n)) \quad (\text{A.1})$$

Sendo f^l a função não-linear de grau l do modelo, y os valores de saída, x os valores de entrada do modelo, n o instante de tempo atual, p o número de atrasos apresentados ao modelo na entrada, q o número de atrasos recuperados da saída para a entrada e o o horizonte de previsão. O diagrama de blocos do modelo **NARX** é mostrado na Figura A.1.

Os principais hiperparâmetros que ajustam o treinamento do modelo são:

- **Tamanho da Camada Escondida** : Número de neurônios nas camadas escondidas, influenciando a capacidade de abstração do modelo.
- **Número de Atrasos na Entrada** : Número de atrasos apresentados ao modelo na entrada, influenciando a capacidade de capturar dependências temporais.

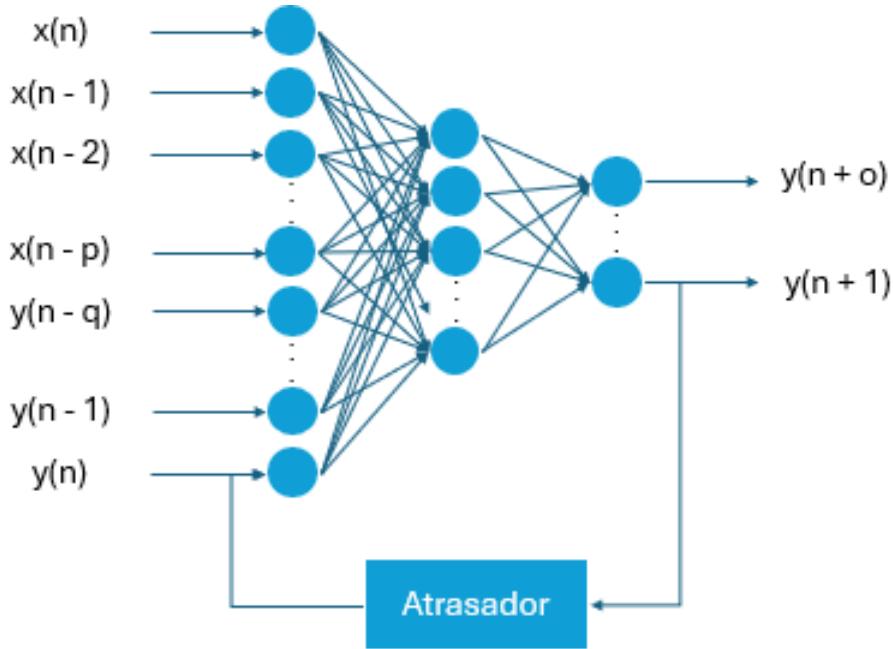


Figura A.1: Diagrama da NARX

Fonte: Elaborado pelo autor, baseado em [12]

A.2 Rede Neural Recorrente Baseada em Atenção de Dois Estágios (do inglês *Dual-Stage Attention-Based Recurrent Neural Network*, DA-RNN)

O modelo DA-RNN representa uma abordagem alternativa, baseada em RNN para a previsão de séries temporais. Este modelo é projetado para lidar com dados sequenciais, utilizando um mecanismo de atenção em duas etapas, o que melhora a precisão das previsões ao capturar dependências de longo prazo [13]. O diagrama é mostrado na Figura A.2.

Sendo x o vetor de entradas, n o instante atual, y o vetor de saídas e i o tamanho da janela de entrada.

Os principais hiperparâmetros que ajustam o treinamento do modelo são:

- **Tamanho da Camada Escondida** : Número de neurônios nas camadas escondidas, influenciando a capacidade de abstração do modelo.

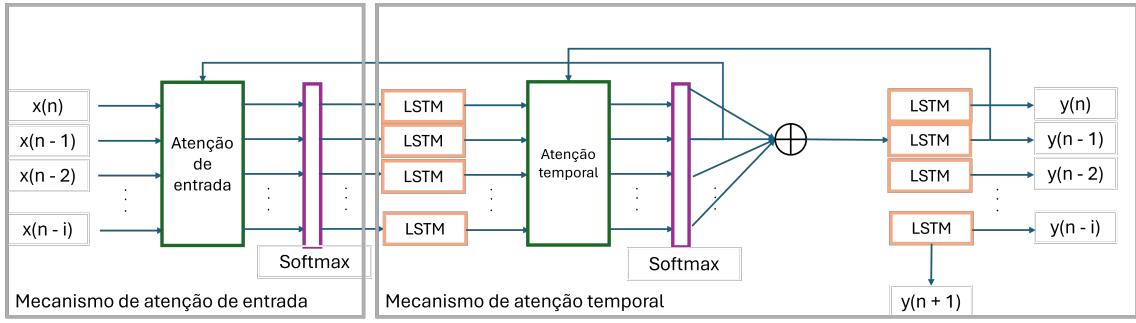


Figura A.2: Diagrama da DA-RNN

Fonte: Elaborado pelo autor, baseado em [13]

- **Número de Cabeças de Atenção**: Número de cabeças no mecanismo de atenção, permitindo ao modelo focar em diferentes partes da sequência de entrada.
- **Taxa de Abandono (do inglês *dropout*)**: Fração de neurônios que são desativados durante o treinamento, evitando o sobreajuste.

A.3 Perceptron Multicamadas (do inglês *Multi-layer Perceptron, MLP*)

A MLP é uma rede neural feedforward composta por múltiplas camadas de neurônios. Cada neurônio aplica uma função de ativação a uma combinação linear das entradas, permitindo modelar relações não lineares, como mostrado na Equação A.2. A estrutura básica envolve uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. O aprendizado é realizado por meio do algoritmo de retropropagação, que ajusta os pesos das conexões para minimizar a diferença entre a previsão e a saída desejada [12]. O diagrama é mostrado na Figura A.3.

$$y = W_2\sigma(W_1x + b_1) + b_2 \quad (\text{A.2})$$

Sendo x o vetor de entrada, W_1 e W_2 as matrizes de pesos da camada de entrada para a camada oculta e da camada oculta para a camada de saída, respectivamente, b_1 e b_2 os vetores de bias e σ uma função de ativação não linear.

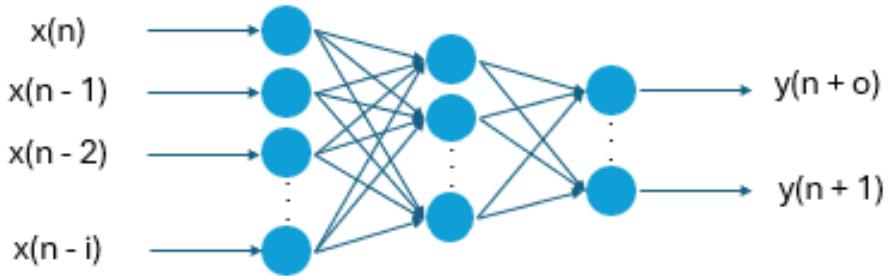


Figura A.3: Diagrama da MLP

Fonte: Elaborado pelo autor, baseado em [12]

Os principais hiperparâmetros que ajustam o treinamento do modelo são:

- **Tamanho da Camada Escondida** : Número de neurônios nas camadas escondidas, influenciando a capacidade de abstração do modelo.

A.4 Perceptron Morfológico Crescente (do inglês *Increasing Morphological Perceptron, IMP*)

O modelo IMP consiste em uma combinação convexa entre duas decomposições morfológicas, por erosões e dilatações. Como mostrado em Araújo, 2012 [14] na Figura A.4.

Sendo n o tamanho da janela de entrada, α a decomposição de operadores morfológicos crescentes por dilatações, β a decomposição de operadores morfológicos crescentes por erosões, v e w são os vetores que definem seus respectivos operadores morfológicos.

Os principais hiperparâmetros que ajustam o treinamento do modelo são os comuns a todos os modelos.

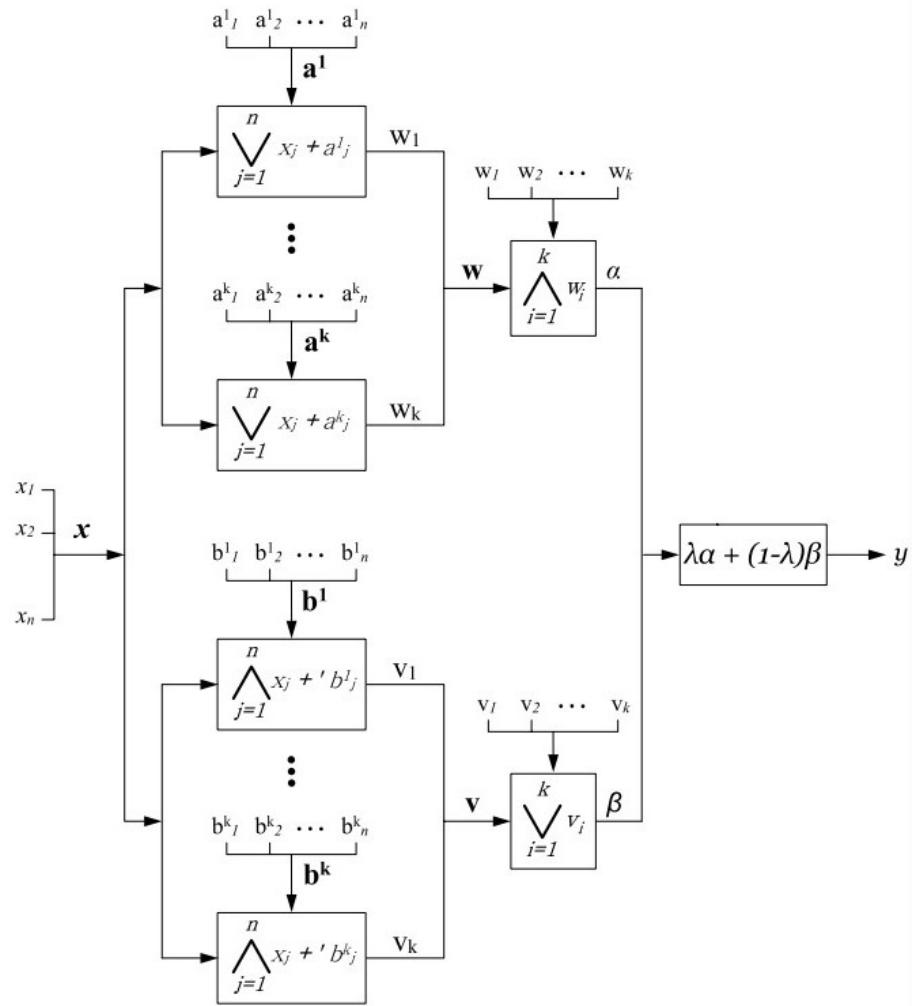


Figura A.4: Diagrama da IMP

Fonte: Araújo, 2012 [14]

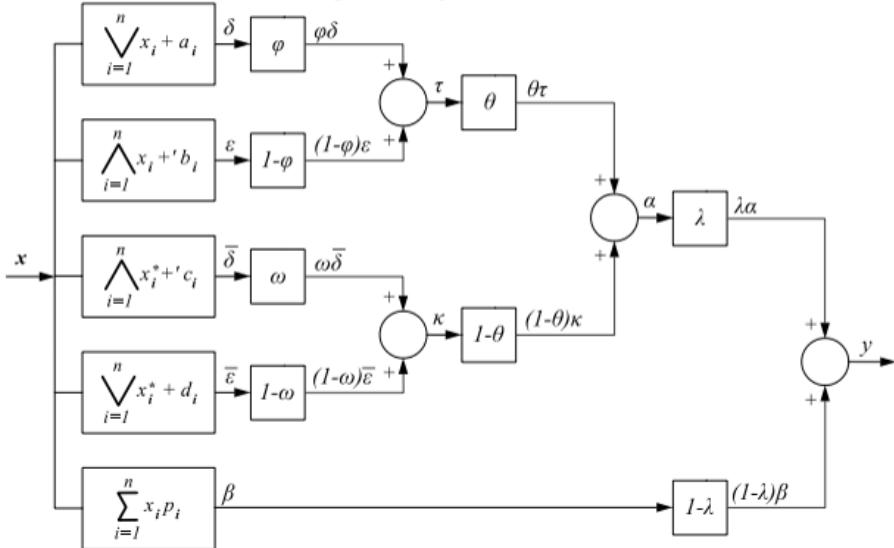


Figura A.5: Diagrama da **IDLN**

Fonte: Araújo, 2016 [15]

A.5 Neurônio Crescente Decrescente Linear (do inglês *Increasing-Decreasing Linear Neuron*, **IDLN**)

O modelo **IDLN** consiste em uma combinação de operadores morfológicos de dilatação, erosão, anti-dilatação e anti-erosão. Conforme mostrado em Araújo, 2016 [15] na Figura A.5.

Os principais hiperparâmetros que ajustam o treinamento do modelo são os comuns a todos os modelos.

A.6 Linear Normalizado (do inglês *Normalized Linear*, **N-Linear**)

O modelo **N-Linear** possui uma arquitetura simples, semelhante ao **MLP**, com a camada de entrada sendo substituída por uma etapa de normalização. Onde cada entrada é subtraída do último valor da sequencia, de modo a evitar outliers [16]. O diagrama é mostrado na Figura A.6.

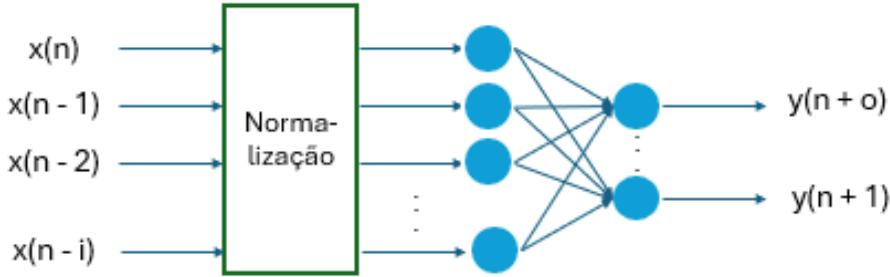


Figura A.6: Diagrama da N-Linear

Fonte: Elaborado pelo autor, baseado em [16]

Os principais hiperparâmetros que ajustam o treinamento do modelo são os comuns a todos os modelos.

A.7 Interpolação Neural Hierárquica para Previsão de Séries Temporais (do inglês *Neural Hierarchical Interpolation for Time Series Forecasting, NHITS*)

O modelo NHITS compõe a classe dos modelos de aprendizado profundo. Sua arquitetura é em pilhas e blocos, e o modelo é composto por uma série de P pilhas, cada pilha é composta por uma série de B blocos residuais que tem por objetivo zerar os resíduos e cada bloco residual é composto por alguns modelos MLP com não linearidades aplicadas pela ReLU [17].

O funcionamento do modelo consiste em dividir o sinal multi-taxa em bandas distintas, para que cada pilha seja especializada em uma banda. Cada pilha possui blocos residuais, onde o sinal é filtrado por um max pool e passa por 3 fluxos de rede MLP, seguido de transformações não lineares e não lineares, para fornecer valores de previsão e retroprojeção. No final a previsão resultante é a soma da previsão de cada pilha com o valor residual da última pilha. O diagrama é mostrado na Figura A.7.

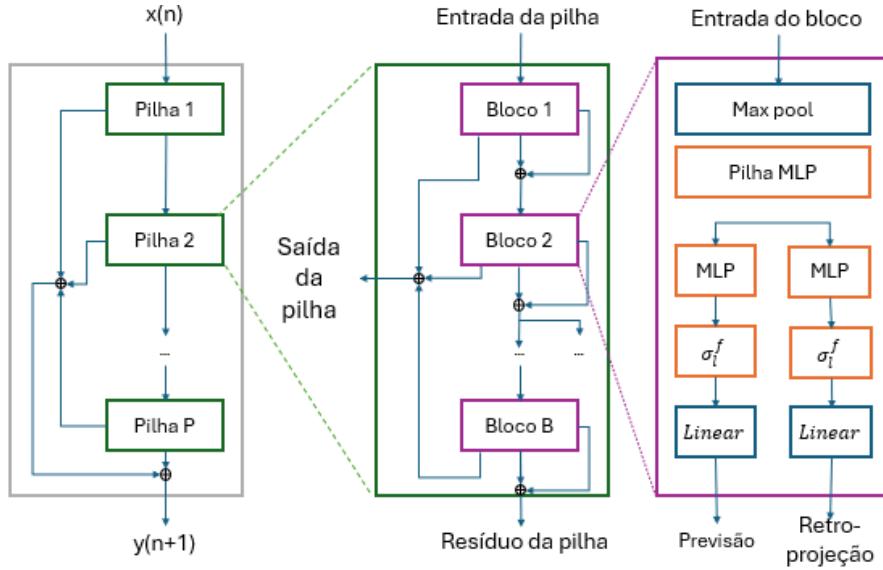


Figura A.7: Diagrama da NHITS

Fonte: Elaborado pelo autor, baseado em [17]

Os principais hiperparâmetros que ajustam o treinamento do modelo são:

- **Número de Pilhas** : Número de pilhas independentes de blocos (*NHiTS-Block*) que compõem a arquitetura. Cada pilha pode aprender diferentes padrões temporais.
- **Número de Blocos** : Número de blocos em cada pilha. Mais blocos permitem que o modelo capture padrões mais complexos ao passar pela rede várias vezes.
- **Número de Camadas** : Número de camadas dentro de cada bloco. Um maior número de camadas adiciona mais profundidade ao modelo, potencialmente melhorando sua capacidade de abstração.
- **Tamanho da Camada** : Número de neurônios por camada dentro de cada bloco. Camadas maiores aumentam a capacidade de aprendizado do modelo.