



INSTITUTO TECNOLÓGICO VALE



**Programa de Pós-Graduação em Instrumentação, Controle e
Automação de Processos de Mineração (PROFICAM)
Escola de Minas, Universidade Federal de Ouro Preto (UFOP)
Associação Instituto Tecnológico Vale (ITV)**

Dissertação

**PREVISÃO DE PREÇOS DE MINÉRIO DE FERRO UTILIZANDO
MODELOS DE INTELIGÊNCIA COMPUTACIONAL**

Hamilton Tonidandel Júnior

**Ouro Preto
Minas Gerais, Brasil
2022**

Hamilton Tonidandel Júnior

**PREVISÃO DE PREÇOS DE MINÉRIO DE FERRO UTILIZANDO
MODELOS DE INTELIGÊNCIA COMPUTACIONAL**

Dissertação apresentada ao Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração da Universidade Federal de Ouro Preto e do Instituto Tecnológico Vale, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Controle e Automação.

Orientador: Prof. Frederico Gadelha Guimarães, D.Sc.

Ouro Preto

2022

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

T665p Tonidandel Junior, Hamilton.
Previsão de preços de minério de ferro utilizando modelos de
inteligência computacional. [manuscrito] / Hamilton Tonidandel Junior. -
2022.

62 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Frederico Gadelha Guimarães.

Dissertação (Mestrado Profissional). Universidade Federal de Ouro
Preto. Programa de Mestrado Profissional em Instrumentação, Controle e
Automação de Processos de Mineração. Programa de Pós-Graduação em
Instrumentação, Controle e Automação de Processos de Mineração.

Área de Concentração: Engenharia de Controle e Automação de
Processos Minerais.

1. Análise de séries temporais - Séries temporais nebulosas. 2.
Minérios de ferro - Economia do mercado - Previsão. 3. Minérios de ferro.
I. Gadelha Guimarães, Frederico. II. Universidade Federal de Ouro Preto.
III. Título.

CDU 681.5:622.2

Bibliotecário(a) Responsável: Maristela Sanches Lima Mesquita - CRB-1716



FOLHA DE APROVAÇÃO

Hamilton Tonidandel Júnior

Previsão de preços de minério de ferro utilizando modelos de inteligência computacional

Dissertação apresentada ao Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração (PROFICAM), Convênio Universidade Federal de Ouro Preto/Associação Instituto Tecnológico Vale - UFOP/ITV, como requisito parcial para obtenção do título de Mestre em Engenharia de Controle e Automação na área de concentração em Instrumentação, Controle e Automação de Processos de Mineração.

Aprovada em 13 de junho de 2022

Membros da banca

Doutor - Frederico Gadelha Guimarães - Orientador - Universidade Federal de Minas Gerais
Doutor - Agnaldo José da Rocha Reis - Universidade Federal de Ouro Preto
Doutor - Petrônio Cândido de Lima e Silva - Instituto Federal do Norte de Minas
Doutora - Rosângela Ballini - Universidade Estadual de Campinas

Frederico Gadelha Guimarães, orientador do trabalho, aprovou a versão final e autorizou seu depósito no Repositório Institucional da UFOP em 05/09/2022



Documento assinado eletronicamente por **Bruno Nazário Coelho, COORDENADOR(A) DE CURSO DE PÓS-GRADUAÇÃO EM INST. CONTROLE AUTOMAÇÃO DE PROCESSOS DE MINERAÇÃO**, em 08/09/2022, às 12:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0394675** e o código CRC **FAB88582**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.012469/2022-01

SEI nº
0394675

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000
Telefone: (31)3552-7352 - www.ufop.br

Agradecimentos

Agradeço a Deus, em primeiro lugar, pelo dom da inteligência e pela resiliência em situações adversas.

Em especial, dedico este trabalho aos meus pais, Hamilton (*in memoriam*) e Fátima, exemplos de amor e esmero, e que nunca mediram esforços para que eu me tornasse um homem de bem.

Agradeço aos meus irmãos, Karina e Wagner, por me inspirarem e proporcionarem confiança em minhas decisões.

Aos meus padrinhos, Graça e José Reis, pelo modelo de correção, serviço e caridade; e aos demais familiares que sempre me apoiaram.

À minha noiva, Natália, pelo seu amor, carinho e compreensão. Minha companheira de todas as horas.

Ao meu orientador, Prof. Dr. Frederico, pela paciência e excelência na transmissão de seu vasto conhecimento.

Aos amigos e colegas do CEFET, da UFOP e do ITV; e a todos os demais colegas da PROGEP.

À UFOP, Instituição que tenho orgulho de servir, pelo suporte em viabilizar uma educação superior de qualidade; e ao Instituto Tecnológico Vale, que tornou possível a realização deste Mestrado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Código de Financiamento 001; do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG); e da Vale SA.

“A chave de todas as ciências é negavelmente o ponto de interrogação.”
(Honoré de Balzac)

Resumo

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

PREVISÃO DE PREÇOS DE MINÉRIO DE FERRO UTILIZANDO MODELOS DE INTELIGÊNCIA COMPUTACIONAL

Hamilton Tonidandel Júnior

Junho, 2022

Orientador: Prof. Frederico Gadelha Guimarães, D.Sc.

A estimativa do comportamento futuro da série temporal de preços de minério de ferro a curto prazo é uma ferramenta importante na elaboração de projetos de mineração relacionados ao planejamento operacional. O presente trabalho propõe, neste contexto, um estudo comparativo entre modelos preditivos construídos por meio técnicas de inteligência computacional, e avalia a acuracidade dos modelos nebulosos PWFTS (*Probabilistic Weighted Fuzzy Time Series*) e FDT-FTS (*Fuzzy Decision Trees*) em relação aos modelos preditivos ARIMA, *Multilayer Perceptron* (MLP) e *Xgboost*. Para garantir a variabilidade dos padrões de entrada, os dados são distribuídos em subconjuntos de janelas deslizantes. Em um contexto multivariado, variáveis preditoras são selecionadas por meio de análise correlacional e causal com a série temporal principal, com destaque para a inclusão do excedente produtivo de minério de ferro da Vale S.A., que apresentou considerável grau de contribuição no poder de generalização dos modelos. Os resultados indicam superioridade dos modelos nebulosos nas métricas RMSE e MAPE, e na estatística U de *Theil*.

Palavras-chave: Séries temporais nebulosas; Previsão; Minério de ferro.

Macrotema: Logística; **Linha de Pesquisa:** Tecnologias da Informação, Comunicação e Automação Industrial; **Tema:** Redução de Incerteza no Planejamento da Mina.

Abstract

Abstract of the Dissertation presented to the Graduate Program on Instrumentation, Control and Automation of Mining Process as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

IRON ORE PRICE PREDICTION USING COMPUTATIONAL INTELLIGENCE MODELS

Hamilton Tonidandel Júnior

June, 2022

Advisor: Prof. Frederico Gadelha Guimarães, D.Sc.

The estimation of the future behavior of a short-term iron ore price time series is an important tool in the elaboration of mining projects related to operational planning. The present work proposes, in this context, a comparative study between different predictive models, built through computational intelligence techniques, and evaluates the accuracy of the fuzzy models PWFTS (Probabilistic Weighted Fuzzy Time Series) and FDT-FTS (Fuzzy Decision Trees) in relation to the models ARIMA, Multilayer Perceptron (MLP) and Xgboost. To ensure the variability of the input patterns, the data is distributed in subsets of sliding windows. In a multivariate context, predictor variables are selected through correlational and causal analysis with the main time series, with emphasis on the inclusion of the productive surplus of iron ore from Vale S.A., which presented a considerable degree of contribution to the generalization power of the models. The results indicate superiority of the fuzzy models in the RMSE and MAPE metrics, and in Theil's U statistic.

Keywords: Fuzzy time series; Forecasting; Iron ore.

Macrotheme: Logistics; **Research Line:** Information Technologies, Communication and Industrial Automation; **Theme:** Uncertainty Reduction in Mine Planning.

Lista de Figuras

Figura 2.1: Arquitetura básica de um <i>perceptron</i>	21
Figura 2.2: Arquitetura de uma rede MLP com duas camadas ocultas.	21
Figura 2.3: Árvore de regressão construída para previsão de preços de minério de ferro.....	22
Figura 2.4: Processo de treinamento de um modelo nebuloso simplificado.	25
Figura 2.5: Exemplos de conjuntos de dados com diferentes níveis de correlação.....	27
Figura 3.1: Fluxograma do procedimento metodológico.	30
Figura 3.2: Série de preços de minério de ferro, em US\$/t.	31
Figura 3.3: Divisão dos dados em treino e teste.	35
Figura 4.1: Correlação das variáveis exógenas ao longo das janelas de dados.	42
Figura 4.2: Histograma dos dados da série de preços de minério de ferro.....	44
Figura 4.3: Diagramas de autocorrelação ACF e PACF para a série principal.	44
Figura 4.4: Comparação entre o valor real e o valor previsto para a série de teste univariada.....	48
Figura 4.5: <i>Boxplots</i> para os subconjuntos de RMSE e MAPE das séries univariadas.	49
Figura 4.6: <i>Boxplots</i> para os subconjuntos de RMSE e MAPE das séries multivariadas.....	52
Figura 4.7: Comparação entre o valor real e o valor previsto para a série de teste multivariada.	53
Figura 4.8: Grau de importância de cada variável preditora nos modelos FDT-FTS.	55

Lista de Tabelas

Tabela 2.1: Classificação das previsões de acordo com o horizonte de previsibilidade.....	17
Tabela 3.1: Conjuntos de dados avaliados.....	32
Tabela 3.2: Modelos considerados neste trabalho.	37
Tabela 4.1: Correlação das variáveis exógenas em relação à série de preços de minério de ferro.....	41
Tabela 4.2: Teste de causalidade de Granger.	43
Tabela 4.3: Estatísticas para o RMSE dos modelos univariados.....	46
Tabela 4.4: Estatísticas para o MAPE dos modelos univariados.	47
Tabela 4.5: Estatísticas para o RMSE dos modelos multivariados.	50
Tabela 4.6: Estatísticas para o MAPE dos modelos multivariados.	51
Tabela 4.7: <i>U-Theil</i> para os modelos testados.....	54
Tabela 4.8: Estatísticas para o RMSE dos modelos sem o excedente produtivo da Vale S.A.	55
Tabela 4.9: Estatísticas para o MAPE dos modelos sem o excedente produtivo da Vale S.A.	56

Sumário

1.	Introdução	12
1.1.	Motivação	13
1.2.	Objetivos	14
1.3.	Organização do texto	14
2.	Referencial Teórico e Fundamentação Científica.....	16
2.1.	Análise e previsão de séries temporais	16
2.2.	Modelos preditivos baseados em estatística	18
2.2.1.	ARIMA	18
2.3.	Modelos preditivos baseados em inteligência artificial	20
2.3.1.	Redes neurais artificiais	20
2.3.2.	Árvores de decisão.....	22
2.3.3.	Séries temporais nebulosas	23
2.4.	Correlação e causalidade das variáveis predictoras	26
2.5.	Trabalhos relacionados	28
3.	Materiais e Métodos	30
3.1.	Descrição dos dados e análise das séries temporais	31
3.1.1.	Avaliação da estacionariedade.....	33
3.1.2.	Diagramas de autocorrelação.....	33
3.2.	Correlação e causalidade das variáveis exógenas	34
3.3.	Construção dos modelos preditivos	35
3.4.	Métricas de desempenho e avaliação de acurácia.....	39
4.	Resultados e Discussão	41
4.1.	Correlação e causalidade entre a série principal e as variáveis exógenas	41

4.2. Análise exploratória da série de preços de minério de ferro	43
4.3. Modelos preditivos univariados	45
4.4. Modelos preditivos multivariados	49
5. Conclusões	57
6. Trabalhos futuros	58
Referências Bibliográficas	59

1. Introdução

Em uma perspectiva econômica, as *commodities* são entendidas como matérias-primas ou produtos com baixo grau de transformação que podem ser comprados ou vendidos, sendo usadas como insumos para a produção de itens de maior valor agregado. De acordo com Bain (2013), uma vez que as *commodities* possuem características uniformes de produção e com pouca variação qualitativa entre os diferentes locais nos quais são encontradas, havendo, portanto, uma cotação mundial, o impacto na oscilação do preço de uma *commodity* é sentido em todo o mercado financeiro global.

Dentre as *commodities* minerais, o minério de ferro se destaca por ser um recurso vital para a economia e para a sociedade, sendo utilizado como matéria-prima na produção do ferro-gusa, posteriormente transformado em aço bruto, que é então empregado na produção de automóveis, estruturas metálicas para construção civil, entre outros fins. Neste aspecto, o Brasil representa um importante papel, sendo um dos principais exportadores desta *commodity*. De fato, o mercado de abastecimento de minério de ferro atual se concentra majoritariamente nos chamados “*big three*”, isto é, Vale – Brasil, Rio Tinto – Austrália, e BHP Biliton – Austrália (MA; ZHEN, 2020). Segundo dados da plataforma ComexStat (2020), o Brasil exportou aproximadamente 326 milhões de toneladas de minério de ferro em 2020, o que coloca a *commodity* na segunda posição dos itens de maior valor exportados pelo Brasil naquele ano. Do ponto de vista da importação de minério de ferro, a China concentra a maior parte das transações, movimentando mais de 1 bilhão de toneladas em 2020 (TRADEMAP, 2020).

De acordo com Aminrostamkolaei *et al.* (2017), a expectativa futura de comportamento do preço do minério de ferro tem se tornado uma questão central para a elaboração de projetos de mineração, sendo uma variável importante no planejamento operacional, especialmente para o melhor aproveitamento do ciclo de vida de uma mina em atividade, contribuindo nas tomadas de decisão da empresa mineradora. No caso de uma projeção de preços em queda, por exemplo, as opções de atraso momentâneo na produção, desaceleração ou mesmo abandono da mina podem ser consideradas. Por outro lado, dada uma projeção de alta nos preços, é possível aplicar estratégias de aceleração, ou mesmo expansão das operações na mina.

Atestada a importância da previsão dos preços de minério de ferro no plano de tomada de decisões da indústria mineradora, esta pesquisa estuda a aplicação de modelos preditivos univariados e multivariados para este fim, utilizando técnicas de inteligência computacional, e

tendo como conjunto principal de dados a série temporal de preços médios de compra e venda de minério de ferro refinado com teor de 62%, em dólares por tonelada.

De acordo com Ewees *et al.* (2020), um dos maiores desafios de se elaborar um modelo preditivo multivariado com boa acurácia para uma série financeira é a escolha das variáveis preditoras externas, isto é, variáveis que possuam correlação com a série principal e ajam como auxiliares na previsão. Para tanto, esta pesquisa contempla a análise da correlação de tais variáveis com a série de preços de minério de ferro, evidenciando o impacto na performance dos modelos multivariados em comparação aos univariados, em relação à acuracidade na estimativa, diante do horizonte de previsão predefinido.

Segundo Egrioglu *et al.* (2020), os métodos de modelagem por séries temporais nebulosas vêm se destacando na literatura, e se apresentam como uma alternativa viável para previsão de variáveis diversas relacionadas ao mercado financeiro. Neste sentido, uma vez que a base dos modelos de séries nebulosas são regras *fuzzy*, a interpretabilidade se torna uma característica importante desses modelos, em contraponto aos ditos “caixa-preta”, como os baseados em redes neurais artificiais, por exemplo. Em vista disto, esta pesquisa também aborda a aplicação de modelos nebulosos no contexto da previsão de preços de minério de ferro.

1.1. Motivação

Considerando o caráter estratégico que o minério de ferro representa na economia brasileira e mundial, o preço desta *commodity* se configura como um importante parâmetro econômico que impacta fortemente os países que a negociam e, em primeira instância, as empresas mineradoras associadas.

Levando-se em conta a posição atual da Vale S.A. como uma das maiores empresas de mineração do mundo e também a maior exportadora de minério de ferro, esta pesquisa busca a criação de modelos preditivos que possam ser utilizados posteriormente como contribuição na tomada de decisões na fase de planejamento operacional. Motivado pela “*theory of storage*”, desenvolvida por Working (1949), estuda-se ainda a correlação do excedente produtivo da Vale S.A. com o preço do minério de ferro, analisando o impacto no poder de previsão dos modelos multivariados. A teoria em questão sugere que preços de *commodities* podem ser afetados pelas quantidades produzidas e estocadas, em uma relação inversamente proporcional. Inicialmente

proposta para o trigo, a teoria foi confirmada para o mercado de *commodities* minerais pelo trabalho de Geman e Smith (2009).

Segundo Li *et al.* (2020), diferentes técnicas estatísticas e baseadas em inteligência artificial têm sido sugeridas para a previsão de séries financeiras aplicadas na previsão de *commodities*, com destaque para o trabalho de Ewees *et al.* (2020), que estudou o impacto da aplicação de um algoritmo de otimização em conjunto com redes neurais artificiais para previsão de preços de minério de ferro.

Ademais, destaca-se que, no presente momento, não são visualizados na literatura trabalhos que apliquem a técnica de modelagem por séries temporais nebulosas no contexto da previsão de preços de minério de ferro.

1.2. Objetivos

O objetivo principal do trabalho é elaborar um estudo comparativo entre modelos preditivos capazes de estimar o comportamento futuro da série temporal de preços de minério de ferro refinado com teor de 62%, empregando, para isso, técnicas de inteligência computacional. Para atingir este objetivo, são definidos os seguintes objetivos específicos:

- Analisar variáveis externas que possam estar correlacionadas com a série de preços de minério de ferro;
- Elaborar modelos baseados em redes neurais, árvores de regressão e lógica nebulosa, para a série univariável e multivariável;
- Analisar o impacto do excedente produtivo da Vale S.A. na acurácia dos modelos;
- Aplicar métricas de acuracidade para definir os melhores modelos preditivos.

1.3. Organização do texto

A estrutura deste texto está organizada como se segue:

No capítulo 1, é feita uma introdução sobre o conceito de séries temporais e a importância da previsão dos preços de minério de ferro para a indústria mineradora, destacando a motivação e os objetivos do trabalho. O capítulo 2 trata da revisão bibliográfica, abordando o conceito de horizonte de previsão e análise exploratória, bem como é dada uma introdução aos modelos preditivos e variáveis correlacionadas, finalizando com os trabalhos correlatos. O

capítulo 3 contém os dados utilizados e o procedimento metodológico para a condução da pesquisa. Por fim, no capítulo 4, são discutidos os resultados deste estudo, finalizando com as conclusões no capítulo 5 e as sugestões para futuros trabalhos no capítulo 6.

2. Referencial Teórico e Fundamentação Científica

2.1. Análise e previsão de séries temporais

Uma série temporal pode ser definida como uma sequência ordenada de valores distribuídos em intervalos de tempo igualmente espaçados. Em outras palavras, para estruturar uma série temporal, os dados devem ser coletados em intervalos regulares de tempo. Neste contexto, prever uma série temporal significa, essencialmente, estimar quais serão os próximos valores da variável de interesse, ou *target*, no intervalo de tempo predefinido (NIST/SEMATECH, 2012).

De acordo com Hyndman e Athanasopoulos (2020), certas variáveis possuem maior previsibilidade do que outras. Neste ponto, alguns fatores contribuem para a capacidade de previsão de uma série temporal, a citar:

- O grau de entendimento dos fatores correlacionados à variável ou ao evento estudado;
- A quantidade de dados disponíveis;
- Se as ações tomadas em função das previsões podem afetar os valores futuros.

A título de exemplificação, previsões de demanda de eletricidade podem ser altamente precisas, uma vez que todas as três condições são normalmente satisfeitas, isto é, se há um histórico suficientemente aceitável de dados sobre a demanda de eletricidade e demais variáveis correlacionadas, é possível desenvolver um modelo de previsão que garanta alta acuracidade. Por outro lado, previsões de taxas de câmbio não compartilham das mesmas premissas, apenas de que há alta quantidade de dados à disposição, não sendo possível compreender claramente, contudo, os fatores que afetam tal flutuação cambial. Outro ponto é que, uma vez divulgada publicamente uma previsão para o valor de uma taxa de câmbio, o mercado tenderá a ajustar os preços rapidamente de acordo com as previsões. Como consequência, a previsão de uma flutuação cambial com boa acurácia é uma tarefa consideravelmente mais complexa (HYNDMAN; ATHANASOPOULOS, 2020).

A construção dos modelos preditivos está associada, inicialmente, à determinação do horizonte de previsão. Neste aspecto, é possível classificar uma previsão como *short-term*, *mid-term* ou *long-term*, isto é, previsões de curto, médio e longo prazo. Conforme evidencia Beasley

(2006), apesar de tais nomenclaturas serem usuais, o significado real de cada termo varia de acordo com a variável estudada. Numa previsão de demanda de energia para implementação de uma usina hidrelétrica, por exemplo, um intervalo de cinco anos seria considerado curto prazo, já cinquenta anos seria longo prazo. Entretanto, na maior parte das aplicações comerciais e industriais, uma previsão de até seis meses é considerada de curto prazo, entre seis meses e dois anos médio prazo, e acima de dois anos longo prazo, conforme ilustra a Tabela 2.1.

Tabela 2.1: Classificação das previsões de acordo com o horizonte de previsibilidade.

Classificação da previsão	Escala temporal	Tipo de decisão	Exemplos
<i>Short-term</i>	Até 3 a 6 meses	Operacional	Controle de inventário; Planejamento da produção, distribuição
<i>Mid-term</i>	De 3 a 6 meses, a 2 anos	Tática	Locação de instalações; contratação de funcionários.
<i>Long-term</i>	Acima de 2 anos	Estratégica	Pesquisa e desenvolvimento; aquisições e fusões; mudanças de produto

Fonte: BEASLEY (2006).

A razão da necessidade de se definir um horizonte de previsão conforme a classificação evidenciada na Tabela 2.1 é que diferentes modelos preditivos se aplicam a diferentes situações. Um modelo de previsão bem construído e com boa acurácia para prever o próximo mês, por exemplo (uma previsão de curto prazo, portanto), possivelmente seria um método inadequado para uma previsão de cinco anos (longo prazo). Em linhas gerais, métodos qualitativos são mais adequados para previsões de longo prazo, ao passo que a análise preditiva de séries temporais, especialmente associadas a variáveis externas, como ocorre em modelos econômicos multivariados, é mais eficaz em um horizonte de previsão de curto prazo (BEASLEY, 2006).

A realização de uma análise exploratória prévia da série temporal é uma etapa importante para a elaboração dos modelos. Nesta fase, busca-se identificar os principais componentes de uma série, como a tendência, a sazonalidade e o erro. Visto que o princípio que orienta uma análise preditiva de um conjunto de dados ordenados é a busca por padrões não aleatórios que se repetem e que, possivelmente, se repetirão no futuro, a identificação de tais componentes é fundamental para verificar se os fenômenos da estacionariedade, da sazonalidade e da autocorrelação estão presentes. Neste sentido, uma série é dita estacionária quando as propriedades estatísticas, como a média e a variância, são constantes ao longo do

tempo. A sazonalidade, por outro lado, se refere a flutuações periódicas. Já a autocorrelação define a influência de valores passados de uma série temporal (*lags*) no valor presente da mesma série (NIELSEN, 2019).

A etapa seguinte à análise exploratória dos dados é a construção dos modelos preditivos. Neste ponto, é válido destacar que a acuracidade de uma previsão depende fortemente do método escolhido (BAUER *et al.*, 2020). Segundo Wolpert e Macready (1997), não há um único método capaz de produzir modelos que apresentem boa performance em qualquer série temporal. Isto posto, é válido submeter a série a vários modelos, construídos com diferentes técnicas, e comparar o resultado final por meio de métricas de desempenho e testes de previsibilidade.

De acordo com Wang *et al.* (2012), é usual distribuir os modelos preditivos em dois grupos, abordados na sequência: os modelos baseados em estatística e os modelos baseados em inteligência artificial.

2.2. Modelos preditivos baseados em estatística

Dentre os modelos matemáticos tradicionais mais utilizados na previsão de séries temporais, pode-se citar o ETS, um modelo de espaço de estados resultante da suavização exponencial sazonal de *Holt-Winters*, e o ARIMA (*Autoregressive Integrated Moving Average*), um modelo formado pela combinação de três estratégias: autorregressão, médias móveis e diferenciação. Esses modelos têm sido tradicionalmente a força motriz da previsão de séries temporais e continuam a ser aplicados em uma ampla gama de situações, desde a pesquisa acadêmica até a modelagem industrial (HYNDMAN; ATHANASOPOULOS, 2020). Neste contexto, considera-se para este trabalho a aplicação de modelos ARIMA como ferramenta de comparação aos métodos de inteligência artificial.

2.2.1. ARIMA

O modelo ARIMA é a combinação de três modelos mais simples. O primeiro deles se trata de uma autorregressão ($AR(p)$), isto é, a variável observada é representada como uma combinação linear de seus valores anteriores, em que a ordem p determina o número de valores passados (*lags*) a serem considerados (NIELSEN, 2019). Um modelo $AR(p)$, portanto, pode ser formulado pela equação (1), em que y_{t-i} são os valores passados, φ_i são os parâmetros da

combinação, ϵ_t é o ruído branco (isto é, uma sequência de variáveis aleatórias não autocorrelacionadas, com média nula e variância finita) e c é a constante.

$$y_t = c + \epsilon_t + \sum_{i=1}^p \varphi_i \cdot y_{t-i} . \quad (1)$$

O segundo componente do ARIMA é o modelo construído por médias móveis (MA(q)). Nesta abordagem, os erros de previsão anteriores são utilizados para representar o valor observado como uma combinação linear, em que a ordem q define o número de erros de previsão passados considerados. Um modelo ARMA(p, q) é representado em (2), em que ϵ_{t-j} representa os erros de previsão anteriores (ruído branco) e θ_j são os parâmetros da combinação.

$$y_t = c + \epsilon_t + \sum_{i=1}^p \varphi_i \cdot y_{t-i} + \sum_{j=1}^q \theta_j \cdot \epsilon_{t-j} . \quad (2)$$

Neste ponto, considera-se o modelo ARMA apropriado para uma série estacionária. Contudo, uma vez que a estacionariedade não for observada, é possível transformar os dados por meio de uma ou mais diferenciações. O acréscimo deste processo no método ARMA resulta no ARIMA(p, d, q), em que d é o grau da diferenciação aplicado. Em outras palavras, o ARIMA é a aplicação do processo de diferenciação em uma série não-estacionária, seguida da estimação pelo ARMA. A formulação final do modelo ARIMA é exibida em (3), em que B é o operador de defasagem, tal que $B^i y_t = y_{t-i}$ (BAUER et al., 2020).

$$\left(1 - \sum_{i=1}^p \varphi_i \cdot B^i\right) (1 - B)^d y_t = c + \left(1 + \sum_{j=1}^q \theta_j \cdot B^j\right) \epsilon_t . \quad (3)$$

Uma variante importante do ARIMA é o *Seasonal* ARIMA, capaz de modelar padrões sazonais. Neste caso, cada componente não sazonal do ARIMA é combinado com sua contraparte sazonal, que é especificada pelos parâmetros P , Q , D e o número de observações por ciclo (HYNDMAN; ATHANASOPOULOS, 2020).

2.3. Modelos preditivos baseados em inteligência artificial

Devido à crescente disponibilidade de dados históricos, os modelos de aprendizado de máquina se estabeleceram como fortes concorrentes aos modelos estatísticos clássicos. Em linhas gerais, boa parte dos métodos de inteligência artificial aplicados à previsão de séries temporais utilizam a técnica de aprendizado supervisionado para o treinamento dos modelos, isto é, dado um conjunto de dados de entrada e de saída previamente conhecidos, um algoritmo é utilizado para aprender a função de mapeamento entre ambos. Previsões são feitas iterativamente sobre os dados de treinamento e o modelo é corrigido até que atinja um nível de performance aceitável (BONTEMPI; TAIEB; LE BORGNE, 2012). No contexto dos modelos preditivos baseados em inteligência artificial, apresenta-se, a seguir, uma introdução às redes neurais artificiais, árvores de decisão e séries temporais nebulosas.

2.3.1. Redes neurais artificiais

As redes neurais artificiais são modelos computacionais simplificados que simulam o processamento feito pelo cérebro humano, permitindo relacionamentos entre as variáveis de saída e seus preditores (variáveis de entrada).

O *perceptron* (Figura 2.1) é a forma mais simples de uma rede neural artificial, usado para classificar padrões linearmente separáveis. Neste modelo, há um único neurônio que comporta todas as camadas. Para cada entrada, é associado um peso que reflete a importância da respectiva entrada para a rede. A combinação linear das entradas com os pesos, somada a um *bias*, produz um valor *net*, aplicado posteriormente a uma função de ativação, que define a saída do neurônio (NIELSEN, 2019).

As redes *perceptron* de múltiplas camadas, ou *multilayer perceptron* (MLP), são redes do tipo *feedforward* que representam uma generalização do *perceptron* de camada única, e são caracterizadas pela presença de uma ou mais camadas intermediárias de neurônios (*hidden layers*). As saídas dos nós em uma camada são as entradas para a próxima camada. Segundo Haykin (2007), as redes MLP se tornaram viáveis por meio do treinamento supervisionado pelo algoritmo de repropagação de erro (*error backpropagation*), e sua precisão está associada a três aspectos: a determinação do número de camadas ocultas, a definição do número de neurônios em cada camada oculta e a especificação dos pesos correspondentes às entradas.

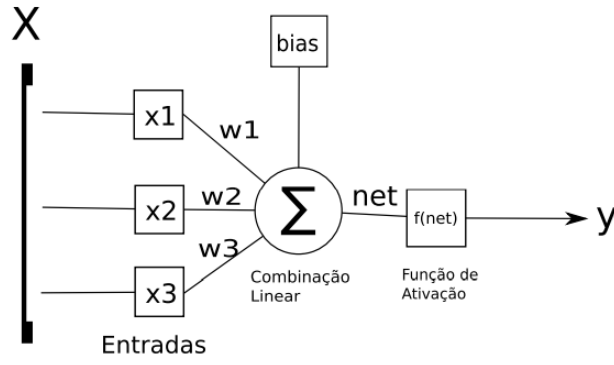


Figura 2.1: Arquitetura básica de um *perceptron*.

Fonte: Adaptado de HAYKIN, 2007.

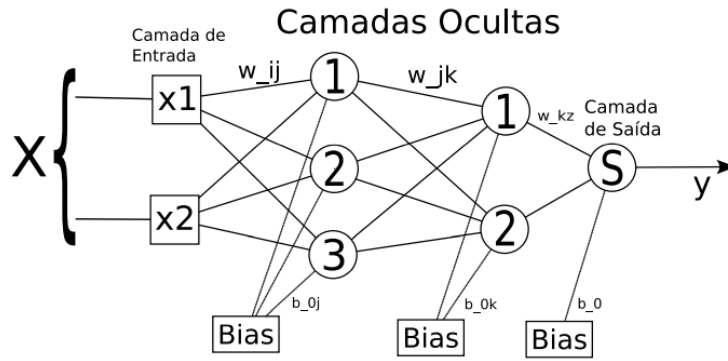


Figura 2.2: Arquitetura de uma rede MLP com duas camadas ocultas.

Fonte: Adaptado de HAYKIN, 2007.

No contexto de previsão de séries temporais, uma rede neural simples pode ser formulada utilizando valores correspondentes aos *lags* da série para treinamento. A equação (4) representa esta formulação, considerando uma única camada oculta. Neste modelo, f representa a rede neural, sendo ϵ_t o erro e $\Phi_l(y_{t-1})$ o vetor que contém l lags, isto é, $\Phi_l(y_{t-1}) = (y_{t-1}, \dots, y_{t-l})$ (BAUER et al., 2020).

$$y_t = f(\Phi_l(y_{t-1})) + \epsilon_t. \quad (4)$$

É válido mencionar que existem outros tipos de redes neurais aplicadas à previsão de séries temporais, como as redes LSTMs (*long short-term memory*) e CNN (*convolutional*

neural networks), por exemplo, que mostraram bons resultados na previsão de diversas séries temporais, especialmente em séries financeiras (LIVIERIS; PINTELAS; PINTELAS, 2020).

2.3.2. Árvores de decisão

A árvore de decisão é um modelo de aprendizado de máquina supervisionado, utilizado para estimar um *target* por meio do aprendizado de regras de decisão a partir de atributos. A construção de uma árvore é feita por particionamento recursivo, isto é, começando do nó raiz, cada nó pode ser dividido em nós filhos. Estes nós são então novamente divididos e se tornam os pais de seus filhos resultantes. Os últimos nós são as folhas, e as linhas que interligam os pais aos filhos são os ramos (ROKACH; MAIMON, 2005).

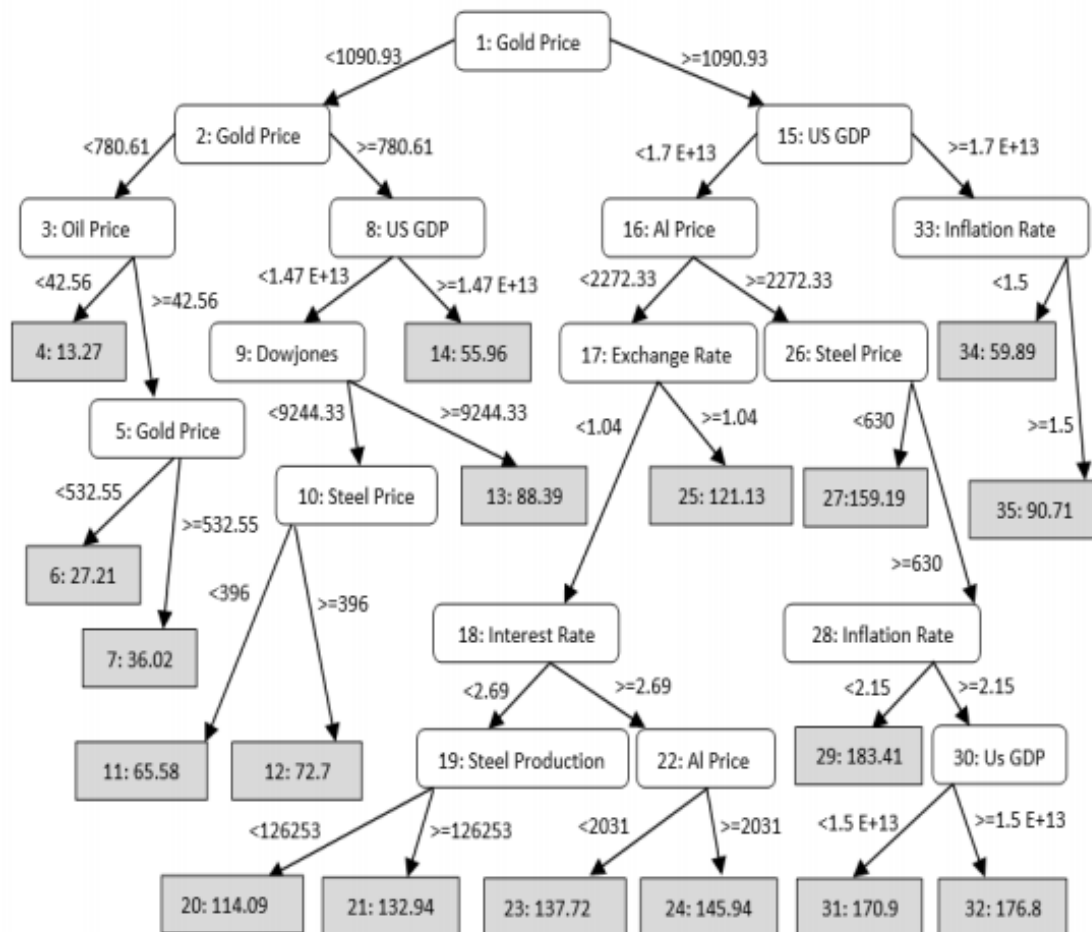


Figura 2.3: Árvore de regressão construída para previsão de preços de minério de ferro.

Fonte: LI *et al.*, 2020.

Em problemas de regressão e, consequentemente, predição, a construção de uma árvore é baseada no particionamento recursivo binário. Inicialmente, todas as amostras de treino são usadas para determinar a estrutura da árvore. O algoritmo então divide os dados usando todas as divisões binárias possíveis e seleciona a divisão que minimiza o desvio padrão nas partes separadas. O processo de divisão é aplicado posteriormente a cada um dos novos ramos e continua até que cada nó alcance um tamanho mínimo especificado, tornando-se um nó folha (XU et al., 2005). A Figura 2.3 ilustra uma árvore de regressão construída para previsão de preços de minério de ferro, considerando algumas possíveis variáveis preditoras externas (LI et al., 2020).

Dentre os métodos baseados em árvores de decisão, destaca-se o algoritmo *XGboost*, ou *extreme gradient boosting*, que é integrado em diversos modelos de *ensembles* de árvores de classificação e regressão. É considerado uma variante do algoritmo GBDT (*gradient boosting decision tree*). A ideia do algoritmo GBDT é adicionar novas árvores a cada iteração, criadas a partir de transformações dos atributos, com o objetivo de ajustar o erro residual da última previsão realizada. O *XGboost* é uma melhoria do GBDT, e tem o intuito de minimizar o treinamento necessário, com um algoritmo mais rápido e com mais acurácia (PENG; HUANG; HAN, 2019).

No contexto da previsão de séries temporais, os modelos baseados em *XGboost* tem mostrado resultados satisfatórios na previsão de preços, com destaque para os trabalhos de Gumus e Kiran (2017) e Peng et al. (2019), que utilizaram o *XGboost* na previsão de preços de petróleo e imóveis residenciais, respectivamente.

2.3.3. Séries temporais nebulosas

As séries temporais nebulosas, ou *fuzzy time series* (FTS), são métodos de previsão inicialmente propostos por Song e Chisson (1993), tomando como base a teoria dos conjuntos nebulosos (*fuzzy sets*). Desde então, os métodos FTS têm sido utilizados na previsão de diversas variáveis em inúmeras áreas do conhecimento, com destaque para as previsões relacionadas ao mercado financeiro (CHENG; CHEN, 2018). De modo geral, observa-se que os estudos das séries temporais nebulosas são relativamente recentes, e ainda há diversas áreas passíveis de aplicação desta técnica (COSTA, 2018). Os primeiros trabalhos conduzidos com lógica nebulosa aplicada em séries temporais elaboraram o conceito de modelagem de primeira ordem, ou *first-order FTS*. Neste método, a estimativa pontual de uma série temporal é determinada

por meio de um único preditor. A ordem indica, neste caso, o número de atrasos de tempo, ou *lags*, usados no modelo (CHEN, 1996).

Sendo U o universo de discurso em que $U = \{u_1, u_2, \dots, u_b\}$, A_i é um conjunto nebuloso de U , definido como:

$$A_i = \{f_{A_i}(u_1)/u_1, f_{A_i}(u_2)/u_2, \dots, f_{A_i}(u_b)/u_b\}, \quad (5)$$

em que f_{A_i} é a função de pertinência do conjunto nebuloso A_i , $f_{A_i} : U \rightarrow [0, 1]$; u_a é um elemento genérico do conjunto A_i ; $f_{A_i}(u_a)$ é o grau de pertinência de u_a a A_i , e $1 \leq a \leq b$.

Neste contexto, o algoritmo proposto por Chen (1996) para a elaboração de um modelo de primeira ordem consiste nas etapas a seguir, sendo a primeira e a terceira etapas referentes ao processo de treinamento, e a quarta etapa referente ao processo de previsão.

Etapa 1: Definir o universo de discurso (U), particionando U em intervalos de mesmo tamanho, tal que $U = \{u_1, u_2, \dots, u_n\}$;

Etapa 2: Transformar os valores numéricos da série em conjuntos nebulosos A_i , em U (*fuzzificação*);

Etapa 3a: Estabelecer relações lógicas nebulosas, ou FLR (*fuzzy logical relationships*), de primeira ordem. Por exemplo, $F(t-1) \rightarrow F(t)$, em que $F(t)$ representa a série fuzzificada no instante t , e $t-1$ representa o instante anterior. São relações do tipo $A_i \rightarrow A_j$, em que A_i é o conjunto nebuloso precedente, no momento $t-1$, e A_j é o conjunto nebuloso consequente, no momento t .

Etapa 3b: Estabelecer grupos usando as FLR. Os grupos FLR são construídos pelo agrupamento das FLR que possuem o mesmo lado esquerdo da relação lógica.

Etapa 4: *Defuzzificar* e realizar a previsão, transformando o conjunto de pertinências de $F(t)$ em um valor pontual na série, tal como se segue:

- Se, no instante t , o lado direito da FLR contiver apenas um conjunto nebuloso na sequência, isto é, $A_{i_1} \rightarrow A_j$, então a previsão em $t+1$ é igual a M_j , em que M_j é o valor defuzzificado de A_j . Em outras palavras, M_j é o centro do intervalo representado por A_j ;

- Se, no instante t , o lado direito da FLR contiver mais de um conjunto nebuloso, isto é, $A_{i_1} \rightarrow A_{j_1}, A_{j_2}, \dots, A_{j_m}$, então a previsão em $t + 1$ é igual a $(M_{j_1} + M_{j_2} + \dots + M_{j_m})/m$;
- Se, no instante t , o lado direito da FLR não contiver um conjunto nebuloso na sequência, isto é, $A_{i_1} \rightarrow \#$, então a previsão em $t + 1$ é igual a $(M_{i_1} + M_{i_2} + \dots + M_{i_p})/p$.

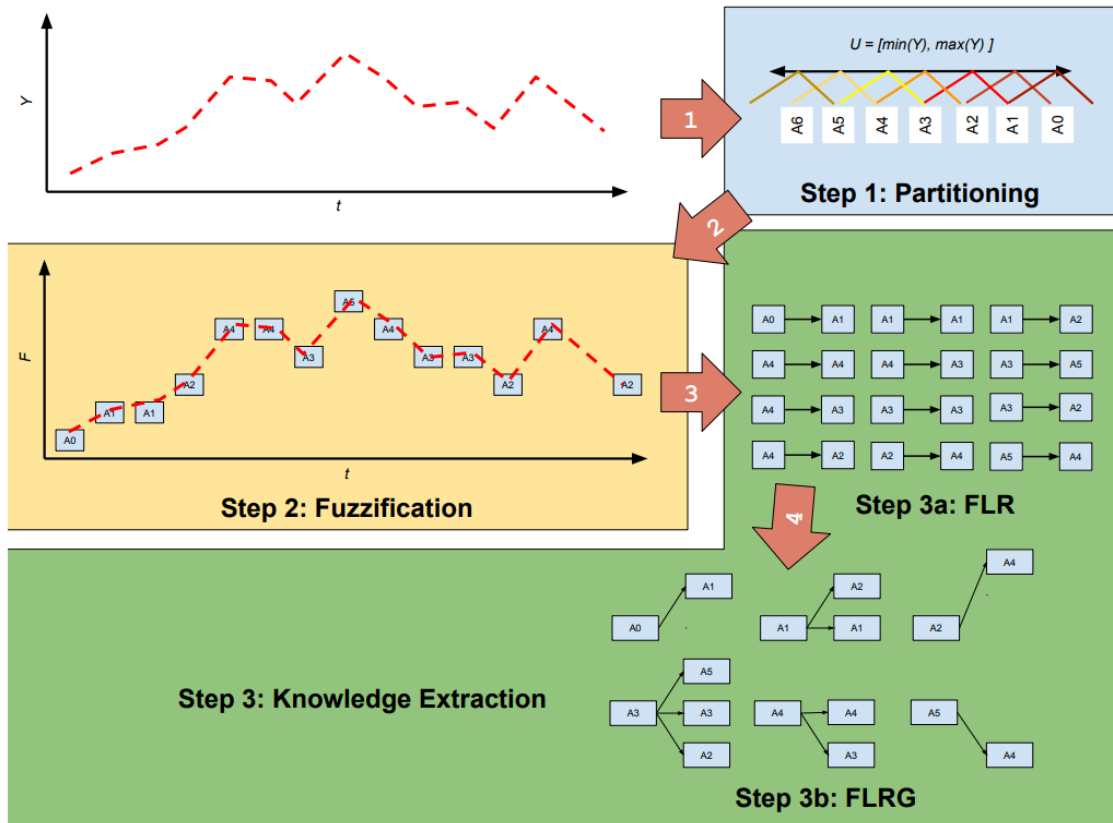


Figura 2.4: Processo de treinamento de um modelo nebuloso simplificado.

Fonte: (SILVA, 2019).

O processo de treinamento de um modelo *fuzzy* simplificado de primeira ordem, como descrito acima, pode ser visualizado pela Figura 2.4. Neste ponto, Silva (2019) destaca uma importante característica dos modelos de séries temporais nebulosas, frente aos modelos caixa-preta: a interpretabilidade. Ao final da fase de treinamento, a base de conhecimento do modelo é formada, essencialmente, pelos grupos FLR, que constituem padrões interpretáveis (caixa-

branca), usados na fase de previsão (Etapa 4), em que o valor predito *fuzzy* é identificado entre as regras, e então *defuzzyficado* para obtenção do valor numérico.

A partir do modelo de primeira ordem descrito acima, Hwang *et al.* (1998) implementou uma técnica de modelagem de alta ordem, isto é, utilizando mais de um *lag* como preditor, que foi aprimorada com a introdução de pesos no consequente das regras. No contexto dos modelos de múltipla ordem, destaca-se o trabalho de Silva (2019), que desenvolveu o método *Probabilistic Weighted Fuzzy Time Series*, ou PWFTS. Este método atribui pesos no consequente e no precedente das regras, aumentando o potencial de acuracidade dos modelos, ao ponderar os conjuntos de regras mais influentes

Já entre os métodos nebulosos voltados para séries multivariadas, destacam-se os modelos híbridos desenvolvidos por Silva (2021). O método FDT-FTS (*Fuzzy Decision Trees*), combina lógica nebulosa com árvores de decisão construídas pelos algoritmos ID3 (*Iterative Dichotomiser 3*), CART (*Classification and Regression Tree*) e RF (*Random Forest*).

De modo geral, observa-se que os estudos das séries temporais nebulosas são relativamente recentes, e ainda há diversas áreas passíveis de aplicação desta técnica (COSTA, 2018).

2.4. Correlação e causalidade das variáveis preditoras

Um modelo preditivo de uma série temporal pode ser construído de forma univariável, isto é, quando os dados passados da própria série de interesse são utilizados para a previsão; ou multivariável, quando um ou mais conjuntos de dados de variáveis externas auxiliam o modelo, melhorando sua acuracidade. Neste aspecto, define-se aqui o conceito de correlação, que é a relação estatística existente entre duas variáveis (HYNDMAN; ATHANASOPOULOS, 2020).

Uma correlação pode ser positiva, quando ambas as variáveis se movem na mesma direção; negativa, quando se movem em direções opostas; ou neutra, o que significa que as variáveis não estão correlacionadas. Segundo Brownlee (2018), em um modelo preditivo para séries temporais, pode ser válida a aplicação de uma ou mais variáveis correlacionadas à série principal, que podem melhorar o poder preditivo do modelo.

A título de exemplificação, a Figura 2.5 ilustra exemplos de conjuntos de dados com diferentes níveis de correlação. É fácil notar que, na primeira coluna, o nível de correlação entre

os dados é alto (positivo ou negativo). Se o conjunto de dados analisado neste caso estivesse representando, por exemplo, as séries temporais da variável de interesse e de uma variável externa possivelmente relacionada, seria uma boa prática, portanto, considerar esta variável externa como preditora na construção do modelo.

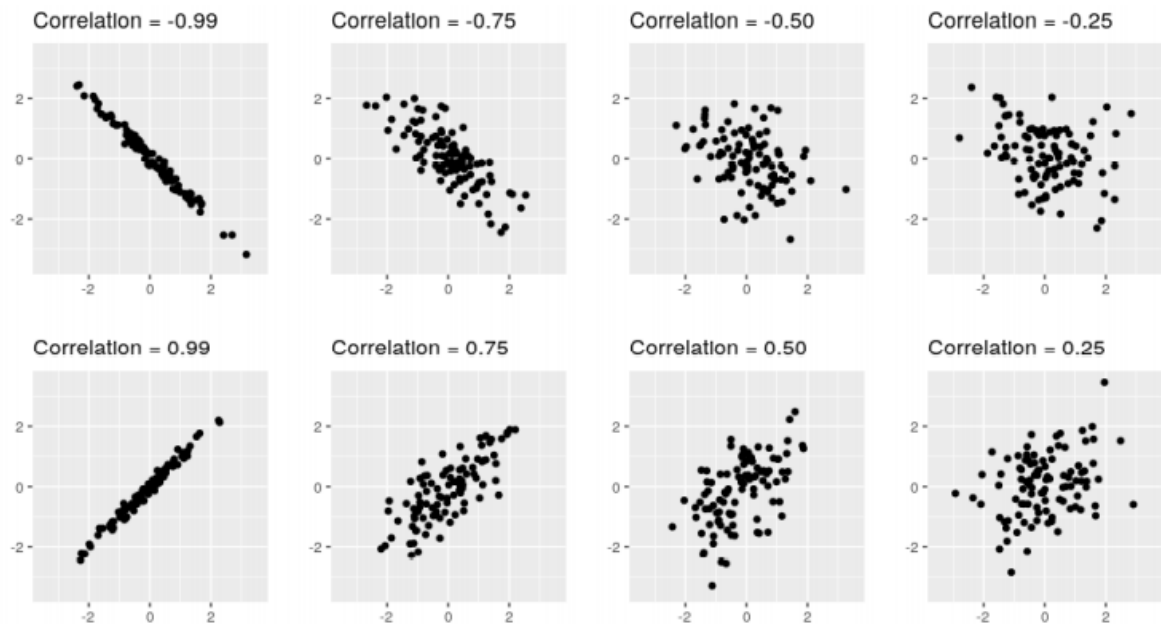


Figura 2.5: Exemplos de conjuntos de dados com diferentes níveis de correlação.

Fonte: (HYNDMAN; ATHANASOPOULOS, 2020).

Atestada a importância da análise correlacional entre a variável *target* e suas possíveis preditoras, é preciso estabelecer uma diferenciação entre correlação e causalidade. De acordo com Patricio (2016), a análise correlacional é insuficiente para estabelecer nexos causais entre as variáveis testadas. Conforme evidencia Calude e Longo (2017), é preciso ter cautela ao estabelecer relações de causa e efeito, especialmente em séries temporais do mercado financeiro.

Usualmente, aplica-se testes estatísticos para evidenciar a existência de causalidade entre as variáveis testadas, como o teste de causalidade de Granger, que se baseia no poder preditivo dos sinais temporais. Segundo Granger, se um sinal temporal X “G-cause” outro sinal temporal Y, então há grandes chances de que os *lags* de X contenham informações que possam ajudar a prever Y, além dos próprios *lags* de Y. Neste contexto, é usual utilizar o termo “G-cause”, que é um acrônimo para “Granger-cause”, que indica uma causalidade relativa, obtida

no teste de Granger, o que pode não significar uma verdade de causa absoluta. É importante salientar ainda que a formulação matemática deste teste é baseada na modelagem por regressão linear (SETH, 2007).

2.5. Trabalhos relacionados

Considerando a proposta desta dissertação, é válido ressaltar alguns trabalhos recentes disponíveis na literatura, em especial os que compartilham do mesmo objetivo principal, ou específicos.

Em Li *et al.* (2020), são consideradas diversas variáveis preditoras externas auxiliares na previsão do minério de ferro, a citar: o preço do aço bruto e sua produção em toneladas, o preço do ouro e do petróleo, a taxa de juros e inflação, a produção de ferro e o preço do alumínio. Neste trabalho, desenvolveu-se um modelo a partir do método *group method of data handling* (GMDH), uma técnica baseada em computação evolutiva. Os testes foram realizados em comparação com modelos de redes neurais artificiais, ARIMA, regressão por vetores de suporte (SVR) e árvores de classificação e regressão. O resultado do trabalho indica superioridade do modelo GMDH na base de dados testada, de periodicidade mensal, considerando as métricas de desempenho RMSE e MAPE.

Em Ewees *et al.* (2020), o estudo considera o preço do cobre, do ouro, da prata, do petróleo e as taxas de inflação e juros da China e dos Estados Unidos, para o desenvolvimento de um modelo de previsão de preços de minério de ferro que integra um algoritmo de otimização meta-heurístico baseado em GOA (*grasshopper optimization algorithm*), chamado de CGOA (*chaotic grasshopper optimization algorithm*) em uma rede MLP. Os resultados empíricos indicam superioridade deste método híbrido em relação aos demais modelos MLP, *k-nearest neighbors* (KNN) e regressão linear. As métricas consideradas foram a RMSE e a MAE.

Em relação às séries temporais nebulosas, considera-se o trabalho de Cheng e Chen (2018), que utiliza a modelagem por séries nebulosas com relações lógicas de alta ordem, e aplica o modelo na previsão dos preços de ouro e taxas de câmbio. O modelo apresenta um algoritmo de regras de associações ponderadas, que calcula o comprimento dos intervalos para que os limites superior e inferior do universo do discurso sejam determinados automaticamente. Os resultados mostram uma melhor acurácia deste modelo, em relação aos métodos ARIMA,

FTS de primeira ordem e GRNN (*general regression neural network*) nas métricas MAPE e RMSE.

3. Materiais e Métodos

Este capítulo descreve os passos a serem realizados para a execução desta pesquisa. Inicialmente, os conjuntos de dados são apresentados e, em seguida, trata-se do procedimento para avaliação da correlação e causalidade entre o preço do minério de ferro e as variáveis exógenas selecionadas. Na sequência, apresenta-se a metodologia para a execução da análise das séries temporais e a construção dos modelos preditivos. Por fim, é tratado das métricas de desempenho utilizadas na avaliação e comparação dos modelos. O procedimento metodológico está sumarizado no fluxograma exibido na Figura 3.1.

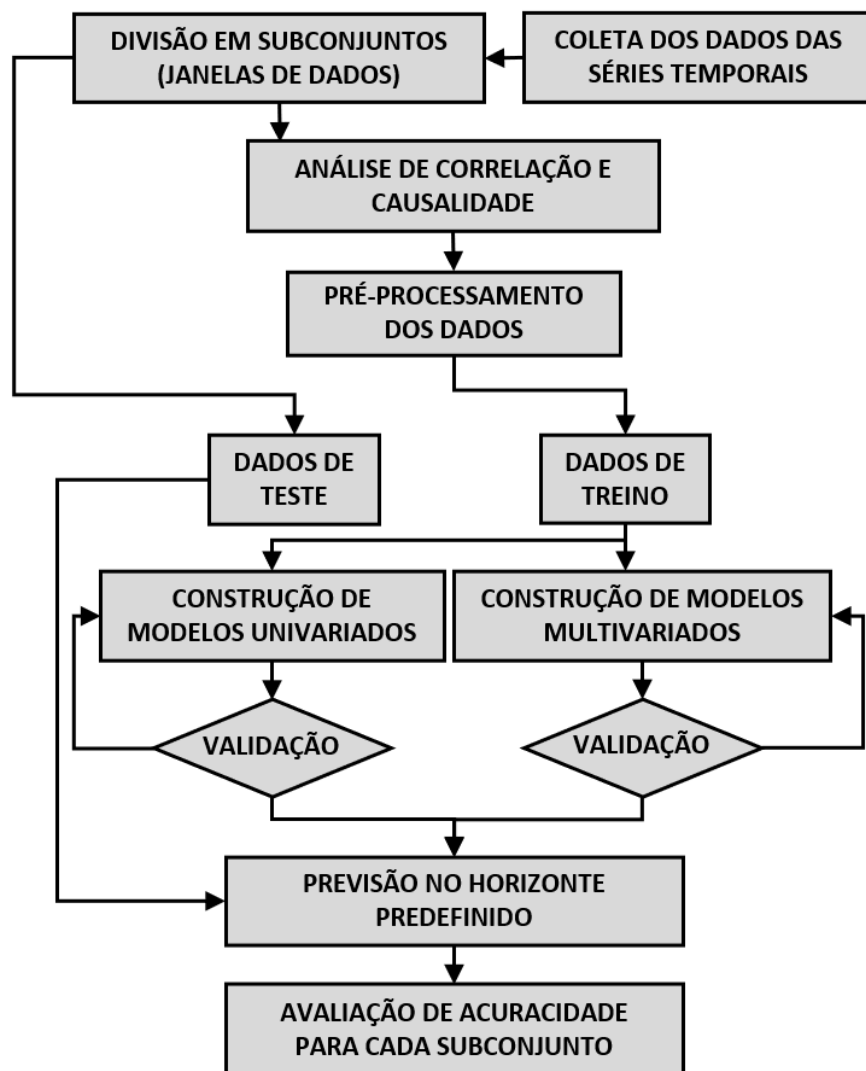


Figura 3.1: Fluxograma do procedimento metodológico.

Fonte: Autor (2020).

3.1. Descrição dos dados e análise das séries temporais

O conjunto de dados para a formação da série de preços de minério de ferro em dólares americanos por tonelada pode ser obtido de diversas fontes, gratuitamente. Para esta pesquisa, considera-se a série temporal mensal de preços médios de contratos futuros de minério de ferro refinado com teor de 62%, em relação ao período compreendido entre julho de 2015 e janeiro de 2022. A Figura 3.2 representa a série temporal em destaque.

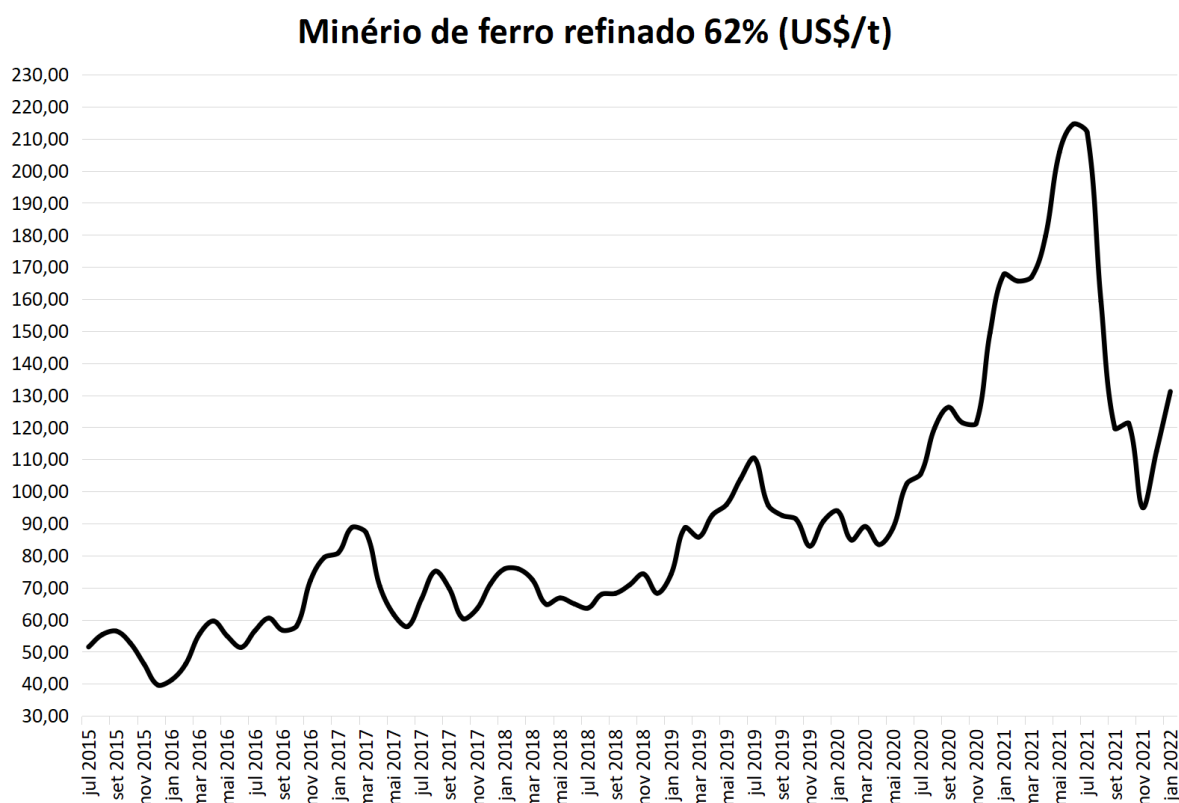


Figura 3.2: Série de preços de minério de ferro, em US\$/t.

Fonte: INVESTING.COM, 2021.

Além do conjunto de dados da série principal, considera-se ainda os dados referentes às variáveis exógenas, que são aplicadas aos modelos preditivos multivariados. Tomando como base os principais *datasets* sugeridos nos trabalhos correlatos apresentados na seção 2.5, esta pesquisa adota como possíveis predictoras as variáveis explicitadas e referenciadas na Tabela 3.1. Todos os conjuntos de dados referentes às variáveis exógenas são coletados também de fontes gratuitas, disponibilizadas *online*.

Tabela 3.1: Conjuntos de dados avaliados.

Variáveis	Unidade	Fonte de dados
Preço do minério de ferro refinado 62% Fe	US\$/t	<i>Investing.com</i> https://br.investing.com/commodities/iron-ore-62-cfr-futures
Preço do petróleo	US\$/barril	<i>U.S. Energy Information Administration</i> https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm
Preço do ouro	US\$/onça troy	<i>Investing.com</i> https://br.investing.com/commodities/gold
Preço da prata	US\$/onça troy	<i>Investing.com</i> https://br.investing.com/commodities/silver
Preço do cobre	US\$/onça troy	<i>Investing.com</i> https://br.investing.com/commodities/copper
Preço da sucata de ferro	US\$/onça troy	<i>Organization for Economic Co-operation and Development</i> https://fred.stlouisfed.org/series/PCU4299304299301
Produção de aço bruto na China	1000t	<i>Worldsteel Association</i> https://www.worldsteel.org/steel-by-topic/statistics/
Quant. de minério de ferro importado pela China	1000t	<i>Trade statistics for international business development</i> https://www.intracen.org/itc/market-info-tools/statistics-import-product-country-monthly/
Taxa de inflação na China (IPC)	%	<i>Organization for Economic Co-operation and Development</i> https://fred.stlouisfed.org/series/CPALTT01CNM659N#0
Excedente produtivo de minério de ferro da Vale	Mt	<i>Vale.com</i> www.vale.com/EN/investors/information-market/quarterly-results/

Fonte: Autor (2020).

Como visto, três destas variáveis referem-se diretamente ao mercado chinês, uma vez que a China é o país que concentra a maior parte das importações de minério de ferro do mundo, sendo também a maior produtora de aço bruto. Neste contexto, é natural supor que a quantidade de minério de ferro importada pela China, a produção de aço bruto e a taxa de inflação chinesa (Índice de Preços ao Consumidor, ou IPC) apresentem alguma correlação ao preço do minério de ferro.

Neste ponto, vale destacar que uma nova variável não incluída nos trabalhos da literatura é analisada nesta pesquisa: o excedente produtivo de minério de ferro da Vale S.A. Os dados desta série, de periodicidade mensal, podem ser coletados por meio dos relatórios de produção da Vale, disponibilizados trimestralmente. A motivação da inclusão desta variável se dá,

conforme abordado no capítulo 1, pela “*theory of storage*”, assumindo a hipótese de que o excedente produtivo de minério de ferro (quantidade de minério produzida menos a quantidade vendida) possa ter alguma correlação com os preços praticados, e que essa correlação, caso exista, seja negativa (inversamente proporcional aos valores da série principal). Outro ponto é que o excedente produtivo é uma variável conhecida previamente. Logo, a confirmação desta correlação pode contribuir na melhora da capacidade de previsão do modelo a ser utilizado posteriormente no planejamento operacional da Vale S.A.

3.1.1. Avaliação da estacionariedade

Uma série é dita estacionária quando sua média e variância são constantes ao longo do tempo. Quando uma série possui forte elemento de tendência, é possível pressupor que não será estacionária. Ademais, é possível aplicar o teste estatístico de raiz unitária de *Dickey-Fuller* aumentado (Teste ADF), no qual um valor- $p \leq 0,05$ indica que a série é estacionária, e um valor- $p > 0,05$ sugere a não-estacionariedade. No R, isto pode ser feito pela função “*adf.test()*”, da biblioteca “*tseries*”.

O fenômeno da estacionariedade é particularmente necessário para aplicação correta dos modelos preditivos utilizados neste estudo. Sendo uma série não-estacionária, é indicada a transformação dos dados por meio da aplicação de uma ou mais diferenciações no pré-processamento, para a formulação de modelos com maior acurácia. O número de diferenciações necessárias para tornar a série estacionária pode ser verificado por meio da função “*ndiffs()*”, do pacote “*forecast*”, no R.

3.1.2. Diagramas de autocorrelação

Os diagramas de autocorrelação indicam se há correlação entre as observações de uma série temporal, sendo utilizados na construção dos modelos preditivos para determinar os intervalos (*lags*) mais significativos. O diagrama ACF (*autocorrelation function*) verifica a correlação entre cada intervalo t com seus vizinhos passados, e o diagrama PACF (*partial autocorrelation function*) retorna a correlação entre observações que não é contabilizada pelos intervalos mais curtos entre essas observações.

No R, é possível visualizar os diagramas ACF e PACF por meio da função “*tsdisplay()*”, contida no pacote “*forecast*”. Os gráficos retornados indicam uma linha tracejada que

representa o intervalo de confiança de 95%, de modo que, dentro deste intervalo, estão contidos os valores que são estatisticamente não-significativos.

3.2. Correlação e causalidade das variáveis exógenas

Seguindo a proposta de investigar a relação entre os preços de minério de ferro e as variáveis exógenas mencionadas, o presente estudo busca estabelecer uma correlação que se preocupa em determinar, essencialmente, a relação linear entre cada grupo de duas variáveis. Para tal, utiliza-se o coeficiente ρ de correlação de Pearson, que assume valores entre -1 e 1, tal que:

- Se $\rho = 1$, existe uma correlação perfeita positiva entre as duas variáveis;
- Se $\rho = -1$, existe uma correlação perfeita negativa entre as duas variáveis;
- Se $\rho = 0$, as duas variáveis não dependem linearmente uma da outra.

Para o cálculo de ρ , utiliza-se a fórmula:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (6)$$

em que \bar{x} e \bar{y} são as médias aritméticas de ambas as variáveis.

É válido destacar que a elaboração dos algoritmos pertinentes à análise correlacional e às demais etapas do projeto é feita com o auxílio de pacotes e *frameworks* dedicados à inteligência computacional e à previsão de séries temporais, disponíveis para as linguagens *Python* e *R*.

Os dados coletados referentes à variável de interesse e às variáveis exógenas candidatas a preditoras dos modelos são extraídos de suas plataformas correspondentes e organizados em planilhas. Para avaliar a correlação, os dados de todas as variáveis são lidos das planilhas e distribuídos em uma matriz, dividindo-a em janelas de dados (janelas deslizantes, seção 3.3), com as variáveis exógenas defasadas de 3 meses em relação à série principal. Na sequência, utiliza-se o comando “*cor()*”, pertencente à biblioteca “*stats*” do *R*, que retorna os coeficientes de Pearson para cada par de variáveis. A matriz de correlação final pode ser exportada pela função “*corrplot()*”, disponível na biblioteca de mesmo nome, e também pela função “*chart.Correlation()*”, do pacote “*PerformanceAnalytics*”.

Em seguida, verifica-se se há, estatisticamente, uma relação de causa e efeito entre as variáveis exógenas correlacionadas e o preço do minério de ferro. Para tal, utiliza-se o teste de causalidade de Granger: Dadas duas séries temporais X e Y estacionárias, a hipótese nula é de que o valor de entrada X “não G-causa” Y, considerando uma dada defasagem temporal. Desta forma, o teste é feito em pares, testando a causalidade entre a série de preços de minério de ferro e cada uma das variáveis exógenas. No R, o teste de Granger é feito pela função “*grangertest()*”, do pacote “*lmtest*”, de forma bidirecional, isto é, testa-se X “não G-causa” Y e também Y “não G-causa” X.

3.3. Construção dos modelos preditivos

Feita a análise de estacionariedade, correlação e causalidade das séries temporais e identificados os elementos para a construção dos modelos, como a presença de dados estacionários e a determinação dos *lags* significativos nos diagramas ACF e PACF, parte-se para a definição das técnicas utilizadas.

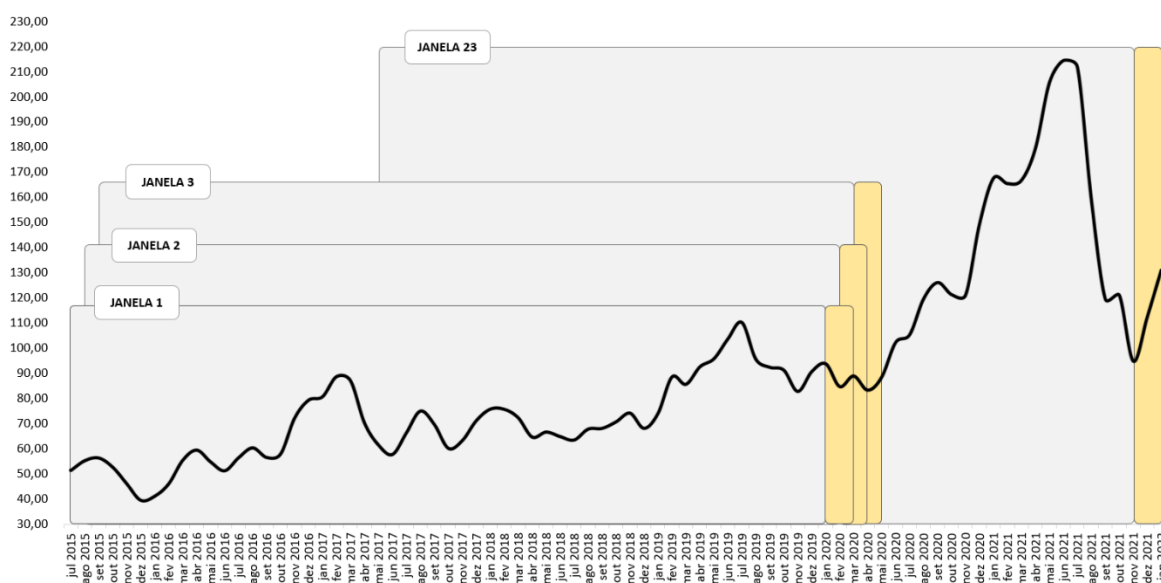


Figura 3.3: Divisão dos dados em treino e teste.

Fonte: Autor (2022).

Inicialmente, a série de preços de minério de ferro é tratada de forma univariável, isto é, são utilizados os próprios *lags* da série para construção dos modelos. Em um segundo momento, é feita a modelagem multivariável, considerando a inclusão das variáveis exógenas

que possuam correlação significativa com a série principal, conforme o procedimento descrito na seção 3.1. Em ambos os casos, os dados são divididos em 23 conjuntos de janelas deslizantes (Figura 3.3), que avançam mês a mês, e cada janela se divide em dois grupos: treino (90%) e teste (10%), sendo os dados de teste correspondentes a uma previsão trimestral, isto é, do tipo *short-term*. É importante salientar que os dados das variáveis exógenas são defasados em, no mínimo, três meses, para evitar vazamento de dados no teste.

O procedimento geral para cada modelo testado segue, portanto, conforme o pseudocódigo:

INÍCIO

Dados \leftarrow Planilha com os dados coletados;

Parametros \leftarrow Parâmetros para cada modelo;

n \leftarrow Tamanho da janela de treino;

m \leftarrow Tamanho da janela de teste;

p \leftarrow Número de janelas deslizantes;

k $\leftarrow 0$;

Enquanto (*k* < *p*) **Faça**

TreinoInicio $\leftarrow k$;

TreinoFinal $\leftarrow n + k$;

TesteInicio \leftarrow *TreinoFinal* + 1;

TesteFinal \leftarrow *TreinoFinal* + *m*;

DadosTreino[*k*] \leftarrow *Preprocessamento*(*Dados*[*TreinoInicio* : *TreinoFinal*]);

DadosTeste[*k*] \leftarrow *Dados*[*TesteInicio* : *TesteFinal*];

Modelo \leftarrow *Treino*(*DadosTreino*[*k*], *Parametros*)

DadosPrevistos[*k*] \leftarrow *Teste*(*Modelo*, *DadosTeste*[*k*]);

AvaliaAcuracidade[*k*];

k $\leftarrow k + 1$;

Fim Enquanto

FIM

Neste estudo, considera-se a técnica estatística ARIMA (*autoregressive integrated moving average*), em comparação às técnicas de inteligência artificial: baseadas em redes neurais artificiais, pela arquitetura multicamadas (*multilayer perceptron*), ou MLP; *ensembles*

de árvores de regressão, pelo algoritmo *XGboost* (*extreme gradient boosting*); e lógica nebulosa, por meio das *fuzzy time series*, ou FTS.

A Tabela 3.2 explicita, sumariamente, os modelos construídos para a previsão da série de preços de minério de ferro, indicando as bibliotecas específicas utilizadas no *Python* e no *R*, bem como os parâmetros configurados manualmente em cada caso.

Tabela 3.2: Modelos considerados neste trabalho.

Método	Bibliotecas	Parâmetros configurados	Sintaxe dos modelos
ARIMA	<i>forecast</i> – <i>R</i>	$p = \text{n}^\circ \text{ de lags significativos no PACF}$	ARIMA(p, d, q)
		$d = \text{n}^\circ \text{ de diferenciações}$	
		$q = \text{n}^\circ \text{ de lags significativos no ACF}$	
MLP	<i>nnfor</i> , <i>neuralnet</i> – <i>R</i>	$hd = \text{n}^\circ \text{ de neurônios em cada camada oculta}$	MLP hd
		$lags = \text{defasagens significativas no ACF}$	
		$eta = \text{taxa de aprendizado}$	
Xgboost	<i>xgboost</i> , <i>caret</i> – <i>R</i>	$max_depth = \text{profundidade máxima de uma árvore}$	XGBOOST
		$eta = \text{taxa de aprendizado}$	
FTS	<i>pyFTS</i> – <i>Python</i>	$model = \text{tipo do modelo utilizado}$	PWFTS $npart_order$ FDTFTS_ID3 FDTFTS_CART FDTFTS_RF
		$npart = \text{n}^\circ \text{ de partições do universo de discurso}$	
		$order = \text{n}^\circ \text{ de lags significativos considerados (ordem do modelo)}$	

Fonte: Autor (2022).

Para os modelos ARIMA, os componentes p , d e q são determinados pela função “*auto.arima()*”, no *R*, presente no pacote “*forecast*”, que usa uma variação do algoritmo desenvolvido por Hyndman e Khandakar. No caso dos modelos MLP, a sintaxe está de acordo com o número de neurônios na camada oculta. Um modelo MLP10, por exemplo, se refere a uma rede que possui 10 neurônios na camada oculta.

Os modelos FTS tomam como base o procedimento metodológico descrito na seção 2.3.3. Neste trabalho, a biblioteca *pyFTS* é utilizada para aplicar o conceito de séries temporais

nebulosas à previsão de preços de minério de ferro. Esta biblioteca foi desenvolvida pelo laboratório *MINDS Lab*, da Universidade Federal de Minas Gerais, e é disponibilizada gratuitamente pelo repositório *pyFTS*.

O treinamento de um modelo FTS na biblioteca *pyFTS* segue os seguintes passos:

- Definição e particionamento do universo de discurso em intervalos sobrepostos;
- Definição dos conjuntos nebulosos a partir dos dados da série temporal (fuzzificação);
- Criação dos padrões temporais (*fuzzy logical relationships*);
- Agrupamento dos padrões e formação das regras, que constituem o modelo FTS propriamente dito.

Partindo do modelo criado, segue-se à etapa de inferência, em que se realiza a previsão para o horizonte definido, seguindo os passos:

- Fuzzificação do valor de entrada, que se deseja prever;
- Busca por regras compatíveis no conjunto de padrões temporais organizados na etapa de treinamento;
- Defuzzificação dos dados, que transforma o resultado previsto em valor numérico.

Em relação aos métodos FTS, considera-se nesta pesquisa os modelos de múltipla ordem, isto é, em que dois ou mais *lags* são usados para a construção das regras. Os métodos avaliados são variações do tipo PWFTS (com pesos probabilísticos no precedente e no consequente das regras), para a modelagem univariável; e do tipo FDT-FTS (*Fuzzy Decision Trees*), método que combina lógica nebulosa às árvores de decisão, construídas pelos algoritmos ID3 (*Iterative Dichotomiser 3*), CART (*Classification and Regression Tree*) e RF (*Random Forest*). O método FDT-FTS é aplicado neste trabalho para a estratégia de modelagem multivariável.

A sintaxe da Tabela 3.2 para os modelos FTS indica o número de partições do universo de discurso e a ordem, que são os principais parâmetros que afetam diretamente a acuracidade dos modelos univariados. Um modelo do tipo PWFTS15_2, por exemplo, se refere à variação pelo método PWFTS, de segunda ordem, e o universo de discurso particionado em 15 partições.

3.4. Métricas de desempenho e avaliação de acurácia

Após a etapa de previsão, realiza-se um estudo comparativo dos modelos para identificar os de maior acuracidade entre os avaliados. Na literatura, diversos indicadores são encontrados. Para esta pesquisa, contudo, considera-se dois destes, o *root mean square error* (RMSE) e o *mean absolute percentage error* (MAPE), sendo o RMSE um parâmetro dependente de escala, e o MAPE um parâmetro percentual, portanto independente de escala. Em ambos os casos, quanto menor, melhor.

O RMSE e o MAPE são calculados como se segue:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}; \quad (8)$$

$$\text{MAPE} = \frac{1}{n} \sum_{j=1}^n \frac{|y_j - \hat{y}_j|}{y_j} \cdot 100\%, \quad (9)$$

sendo y_j o j -ésimo valor real da série e \hat{y}_j o j -ésimo valor previsto, considerando um vetor de tamanho n . Finalmente, os modelos construídos podem ser avaliados de acordo com sua acuracidade, em relação a cada janela de previsão. Para tanto, são analisadas as estatísticas descritivas do conjunto de RMSE e MAPE, em que são verificadas a posição central (mediana), a dispersão do erro (intervalo interquartílico, ou IQR, que é a diferença entre o terceiro e o primeiro quartil), e a presença de valores discrepantes (*outliers*).

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{y}_{t+1} - y_{t-1}}{y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{y_{t+1} - y_{t-1}}{y_t} \right)^2}}, \quad (10)$$

Ademais, o índice de desempenho U de *Theil* também é usado nesta pesquisa para identificar o desempenho das previsões em relação ao modelo ingênuo (*Naive*). A previsão ingênua considera que os valores previstos das janelas de três meses no conjunto de teste são iguais aos três últimos do conjunto de treino. O coeficiente U de *Theil* é obtido pela equação

(10), sendo que um coeficiente U maior que 1 indica superioridade do modelo ingênuo, ao passo que valores de U menores que 1 mostram superioridade do método testado em relação ao *Naive*.

4. Resultados e Discussão

Este capítulo trata dos resultados e discussões, considerando os procedimentos metodológicos da seção anterior. Inicialmente, verifica-se os resultados da análise da correlação e da causalidade entre a série principal e as variáveis exógenas, em que são selecionadas as variáveis modeladas como preditoras nos modelos multivariados. Em seguida, discute-se acerca da análise da série de preços de minério de ferro para, finalmente, tratar dos modelos preditivos univariados e multivariados em um estudo comparativo, estudando ainda a contribuição do excedente produtivo da Vale S.A. nos modelos.

4.1. Correlação e causalidade entre a série principal e as variáveis exógenas

A análise de correlação entre as variáveis exógenas indicadas na Tabela 3.1 foi realizada conforme indicado na seção 3.2. A Tabela 4.1 contém a média dos coeficientes de correlação de Pearson (ρ) em cada uma das 23 janelas de dados, e também o coeficiente global, para toda a série. Em negrito, estão representados os coeficientes de maior relevância.

Tabela 4.1: Correlação das variáveis exógenas em relação à série de preços de minério de ferro.

Variável	ρ global	Média de ρ em cada janela de dados
Preço do petróleo	0,09	-0,10
Preço do ouro	0,87	0,70
Preço da prata	0,77	0,50
Preço do cobre	0,51	0,25
Preço da sucata de ferro	0,34	0,14
Produção de aço bruto na China	0,76	0,72
Quant. de minério de ferro importado pela China	0,75	0,68
Taxa de inflação na China (IPC)	0,52	0,65
Excedente produtivo da Vale	-0,51	-0,51

Fonte: Autor (2020).

De modo geral, nota-se que, das variáveis analisadas, destacam-se as seguintes:

- Preço do ouro, $\rho = 0,70$;
- Preço da prata, $\rho = 0,50$;
- Quantidade de minério de ferro importado pela China, $\rho = 0,68$;
- Produção de aço bruto na China, $\rho = 0,72$;
- Excedente produtivo de minério de ferro da Vale, $\rho = -0,51$;

A Figura 4.1 ilustra o comportamento da correlação ao longo das janelas deslizantes. Neste ponto percebe-se que, nem sempre, o grau de correlação se manteve constante em todo o conjunto de janela de dados. O preço do petróleo e da sucata de ferro, por exemplo, apresentaram grandes variações, percorrendo o campo das correlações positivas e negativas, e resultando em uma média geral baixa. Vale salientar, todavia, que as variáveis referentes ao mercado chinês se mantiveram fortemente correlacionadas com o preço do minério de ferro ao em todo o período analisado.

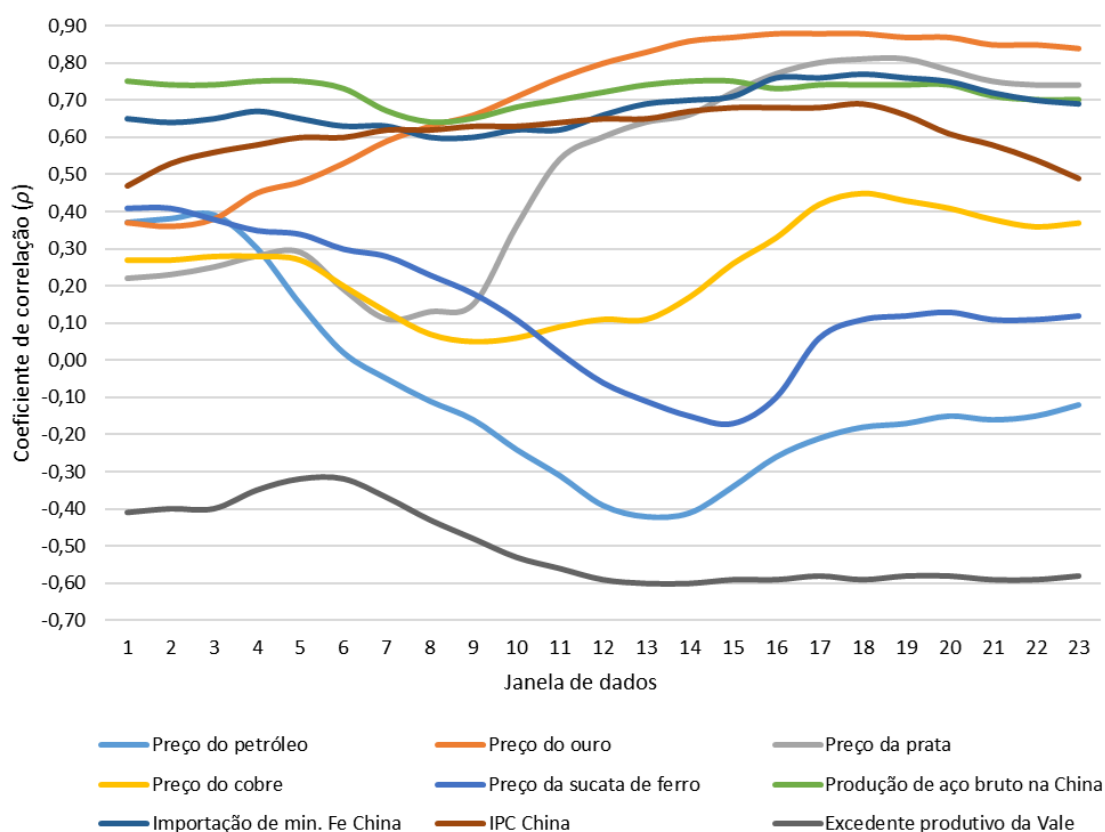


Figura 4.1: Correlação das variáveis exógenas ao longo das janelas de dados.

Fonte: Autor (2022).

Analisando o excedente produtivo de minério de ferro da Vale, nota-se que, a princípio, o coeficiente de correlação com a série temporal principal exibe uma tendência de crescimento modular a partir da sexta janela, e se mantém estável até o fim do conjunto de dados. O valor de correlação $\rho = -0,51$ confirma a hipótese inicial de que o excedente produtivo da Vale possui, de fato, correlação com os preços praticados da *commodity*. Isto indica a possibilidade de que esta variável, assim como as demais destacadas em negrito na Tabela 4.1, se comportem como preditoras para os modelos multivariados, vistos na seção 4.4.

A Tabela 4.2 exibe o resultado dos p-valores para o teste de causalidade de Granger das séries estacionárias, descartando-se do teste o preço do petróleo, do cobre e da sucata de ferro, por não apresentarem resultados satisfatórios na análise de correlação. Neste teste, verifica-se a G-causalidade bidirecional, sendo que um valor-p menor que 0,05 é necessário para se rejeitar a hipótese nula de não-causalidade.

Tabela 4.2: Teste de causalidade de Granger.

Hipótese nula	p-Valores			
	Lag 3	Lag 6	Lag 9	Lag 12
Exc. da Vale não causa Preço do min. Fe	0,36	0,04	0,05	0,05
Preço do min. Fe não causa Exc. da Vale	0,66	0,15	0,12	0,36
Prod. aço China não causa Preço do min. Fe	0,02	0,01	0,04	0,42
Preço do min. Fe não causa Prod. aço China	0,25	0,19	0,41	0,52
Imp. min. Fe China não causa Preço do min. Fe	0,07	0,10	0,00	0,00
Preço do min. Fe não causa Imp. min. Fe China	0,00	0,01	0,00	0,00
Preço da prata não causa Preço do min. Fe	0,87	0,91	0,66	0,01
Preço do min. Fe não causa Preço da prata	0,03	0,00	0,00	0,03
Preço do ouro não causa Preço do min. Fe	0,05	0,01	0,00	0,00
Preço do min. Fe não causa Preço do ouro	0,04	0,12	0,00	0,00
IPC China não causa Preço do min. Fe	0,07	0,03	0,06	0,00
Preço do min. Fe não causa IPC China	0,54	0,02	0,08	0,40

Fonte: Autor (2022).

Para o conjunto final de dados de treino e teste nos modelos multivariados, considerou-se as variáveis: preço do ouro, com defasagem de 6 meses; preço da prata, com defasagem de 12 meses; quantidade de minério de ferro importado pela China, com defasagem de 6 meses; produção de aço bruto na China, com defasagem de 3 meses; e excedente produtivo da Vale, com defasagem de 6 meses.

4.2. Análise exploratória da série de preços de minério de ferro

Na construção dos modelos univariados, foi considerada apenas a série de preços de minério de ferro, em que os próprios *lags* são utilizados como preditores para estimar o comportamento futuro. Pelo histograma da Figura 4.2, nota-se que os preços se concentram na maior parte do conjunto de dados, entre 63,48 e 102,11 dólares por tonelada (primeiro e terceiro quartil), com pico de 214,55 em junho de 2021. Observando a Figura 4.3, percebe-se forte padrão de tendência de crescimento dos preços até junho de 2021, quando apresentaram uma

queda considerável em um espaço de seis meses, retomando o crescimento em dezembro do mesmo ano.

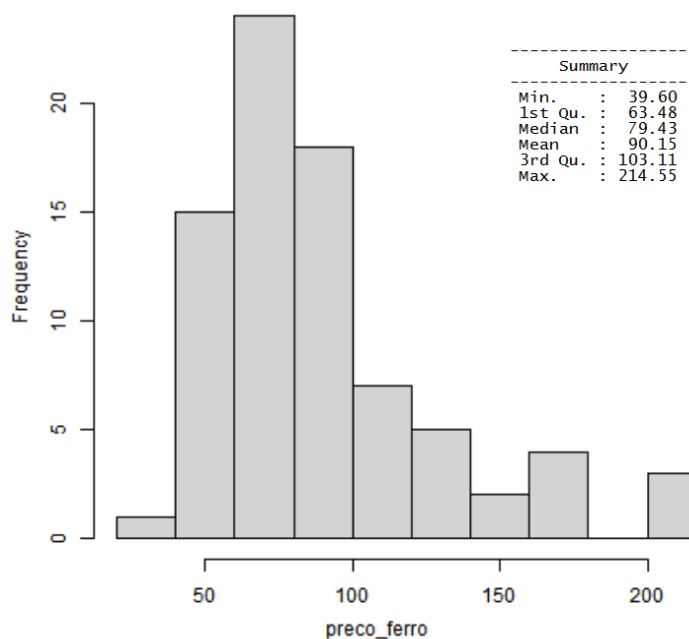


Figura 4.2: Histograma dos dados da série de preços de minério de ferro.

Fonte: Autor (2022).

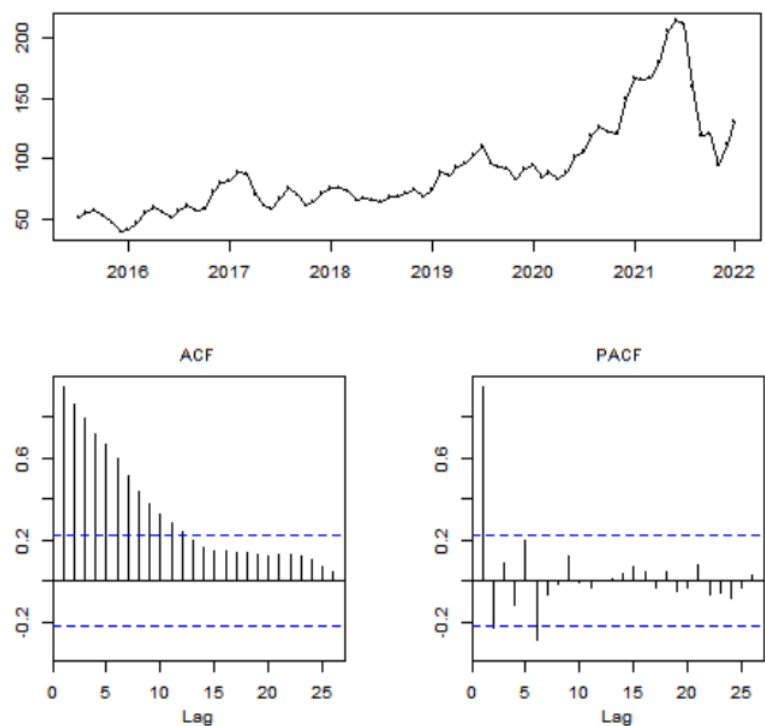


Figura 4.3: Diagramas de autocorrelação ACF e PACF para a série principal.

Fonte: Autor (2022).

Para confirmar a não-estacionariedade evidenciada pelo padrão de tendência, realizou-se o teste de raiz unitária de *Dickey-Fuller* aumentado (teste ADF), que retornou *p-valor* igual a 0,93, portanto maior que 0,05, sugerindo de fato uma série não-estacionária. Neste caso, identificou-se que uma diferenciação de primeira ordem seria suficiente para tornar estacionária a série temporal do conjunto de dados de treino.

Os diagramas de autocorrelação podem ser vistos também na Figura 4.3. Os *lags* ditos significativos, em cada diagrama, devem estar fora do intervalo de confiança compreendido pela linha tracejada azul. Tomando a série com periodicidade mensal, nota-se forte autocorrelação com os *lags* vizinhos (até o décimo segundo mês de atraso), apresentando baixo decaimento dos valores de autocorrelação ao longo do tempo, o que indica que os *lags* vizinhos são altamente significativos para explicar o comportamento da série em um dado instante. Devido à alta dependência dos *lags* vizinhos, espera-se bom desempenho do modelo ingênuo (*Naive*) em um horizonte de previsão trimestral.

4.3. Modelos preditivos univariados

Inicialmente, todos os modelos passaram pela fase de pré-processamento, com uma etapa de diferenciação e normalização na fase de treino.

Em relação ao modelo ARIMA, os coeficientes foram escolhidos por otimização interna da função “*auto.arima()*”. Os resultados da tabela são, portanto, referentes aos modelos ARIMA do tipo (0,1,1).

Nos modelos MLP, considerou-se a variação do número de neurônios da camada oculta (*hd*) em 5, 10 e 20, tomando a taxa de aprendizado (*eta*) como 0,4. Foi cogitado, a princípio, o uso de mais de uma camada oculta, contudo os resultados permaneceram similares, não justificando o maior custo computacional. Os *lags* 1, 2, 3 e 4 foram utilizados no treinamento, e cada modelo MLP passou por 10 rodadas de treino e teste para cada janela de dados, sendo o resultado final formado pela média aritmética das 10 rodadas.

Acerca do modelo *Xgboost*, os parâmetros configurados manualmente foram “*max_depth*” = 10, e “*eta*” = 0,4. O algoritmo possui uma ferramenta de otimização interna para a escolha dos demais parâmetros. Similar ao modelo MLP, a série teve 10 rodadas de treino e teste para cada janela de dados.

Os modelos FTS foram construídos a partir de três variações do método PWFTS com 10, 15 e 20 partições. Todos os modelos testados são de terceira ordem. Os dados foram fuzzificados considerando uma função de pertinência do tipo triangular, com conjuntos nebulosos de tamanhos iguais.

A Tabela 4.3 e a Tabela 4.4 exibem as estatísticas descritivas das métricas RMSE e MAPE, respectivamente, para os 23 subconjuntos de dados de teste, que também podem ser visualizadas nos *boxplots* da Figura 4.5. Nota-se baixo poder de generalização do modelo *Xgboost*, que apresentou as maiores medianas e grande dispersão tanto no RMSE quanto no MAPE. Os modelos ARIMA e MLP obtiveram melhores resultados do que o *Xgboost*, porém evidencia-se a presença de fortes outliers em ambos os casos. Os modelos FTS, por outro lado, se destacam por apresentarem as menores medianas, com melhores mínimos e máximos, e baixa variabilidade, pelo IQR reduzido. A Figura 4.4 ilustra a comparação entre os valores reais da série e os valores preditos por cada modelo univariados.

Tabela 4.3: Estatísticas para o RMSE dos modelos univariados.

Janela de testes	ARIMA	MLP5	MLP10	MLP20	XGBOOST	PWFTS 10_3	PWFTS 15_3	PWFTS 20_3
1	5,60	7,32	10,11	13,96	11,76	4,34	3,89	3,71
2	5,50	12,97	12,63	15,58	5,92	6,57	12,42	13,34
3	3,89	16,29	16,09	16,64	7,20	7,64	3,97	2,99
4	8,71	13,34	15,13	19,22	6,94	17,71	17,28	17,28
5	19,15	17,32	18,36	15,60	3,69	13,78	14,92	10,21
6	18,55	23,84	21,70	22,56	14,17	20,78	16,89	12,98
7	21,08	30,04	29,86	24,78	24,09	16,98	16,71	16,83
8	19,44	11,05	12,70	11,10	29,57	10,03	8,29	3,03
9	3,63	4,69	9,32	9,72	35,16	7,68	4,87	3,50
10	14,42	26,37	25,13	25,35	44,00	31,05	23,00	23,00
11	31,58	33,17	28,91	27,45	59,25	28,86	28,86	28,86
12	40,21	35,19	28,96	29,41	71,68	9,39	9,39	9,39
13	13,89	9,95	9,52	10,41	79,00	4,22	2,41	2,72
14	8,36	9,13	9,12	8,01	85,23	15,29	16,92	10,93
15	23,45	20,72	21,75	19,30	98,44	27,01	27,01	27,01
16	32,89	21,35	16,18	18,83	108,87	20,85	21,06	20,79

17	24,53	33,00	43,14	32,53	111,93	2,69	18,18	24,24
18	45,69	44,75	38,41	31,94	92,30	42,08	44,39	44,39
19	71,82	58,37	52,16	61,83	65,69	59,52	58,21	58,21
20	88,79	107,43	101,80	85,90	23,31	21,57	22,06	22,07
21	21,63	32,45	46,03	71,23	15,24	24,66	20,81	20,81
22	27,63	21,17	14,95	11,34	26,17	11,31	12,59	13,48
23	21,82	22,97	26,89	23,91	33,56	23,84	24,34	29,89
Mínimo	3,63	4,69	9,12	8,01	3,69	2,69	2,41	2,72
Mediana	21,08	21,35	21,70	19,30	33,56	16,98	16,92	16,83
Máximo	88,79	107,43	101,80	85,90	111,93	59,52	58,21	58,21
IQR	18,30	19,57	15,58	13,66	60,64	15,72	11,62	13,82

Fonte: Autor (2022).

Tabela 4.4: Estatísticas para o MAPE dos modelos univariados.

Janela de testes	ARIMA	MLP5	MLP10	MLP20	XGBOOST	PWFTS 10_3	PWFTS 15_3	PWFTS 20_3
1	5,68	7,07	10,79	13,69	4,27	4,48	2,88	3,38
2	5,89	15,03	14,21	17,56	7,45	5,46	13,11	12,33
3	4,07	17,59	16,61	17,86	3,71	7,79	3,97	2,24
4	8,44	14,21	16,57	21,07	6,03	18,88	18,72	18,72
5	18,18	14,98	14,59	12,78	7,36	13,91	15,16	8,71
6	14,44	18,87	17,23	16,86	8,06	19,05	15,19	10,17
7	15,35	23,92	23,96	20,50	16,28	14,37	14,00	14,15
8	15,57	8,44	1,90	0,87	23,88	6,52	5,67	2,11
9	3,22	3,31	6,01	6,36	21,58	5,35	6,37	2,59
10	7,50	12,81	14,23	14,36	20,06	23,59	15,48	15,48
11	15,73	15,80	13,82	13,24	34,75	16,87	16,87	16,87
12	24,25	20,47	16,33	17,04	45,63	4,03	4,02	4,02
13	8,24	5,42	4,96	5,47	33,35	2,10	1,29	1,37
14	3,11	5,26	5,27	4,39	27,56	8,03	9,61	5,36
15	8,80	9,55	10,87	10,42	36,33	12,60	12,60	12,60
16	14,66	10,23	8,07	8,87	47,12	8,32	8,80	7,91
17	11,43	15,19	19,87	15,19	48,26	7,90	7,87	10,69
18	19,99	20,43	14,45	14,84	24,58	18,28	18,65	18,65

19	43,69	35,48	32,27	37,12	21,58	26,15	28,18	25,16
20	67,73	81,94	77,71	65,58	19,29	10,91	9,34	9,39
21	17,33	26,90	34,77	54,84	10,86	17,59	14,10	14,10
22	21,78	15,34	10,47	10,26	13,16	8,63	11,40	12,32
23	16,31	20,14	22,95	19,13	23,96	19,99	21,01	24,66
Mínimo	3,11	3,31	1,90	0,87	3,71	2,10	1,29	1,37
Mediana	14,66	15,19	14,45	14,84	21,58	10,91	12,60	10,69
Máximo	67,73	81,94	77,71	65,58	48,26	26,15	28,18	25,16
IQR	9,89	10,40	7,92	8,15	20,99	10,78	8,22	10,12

Fonte: Autor (2022).

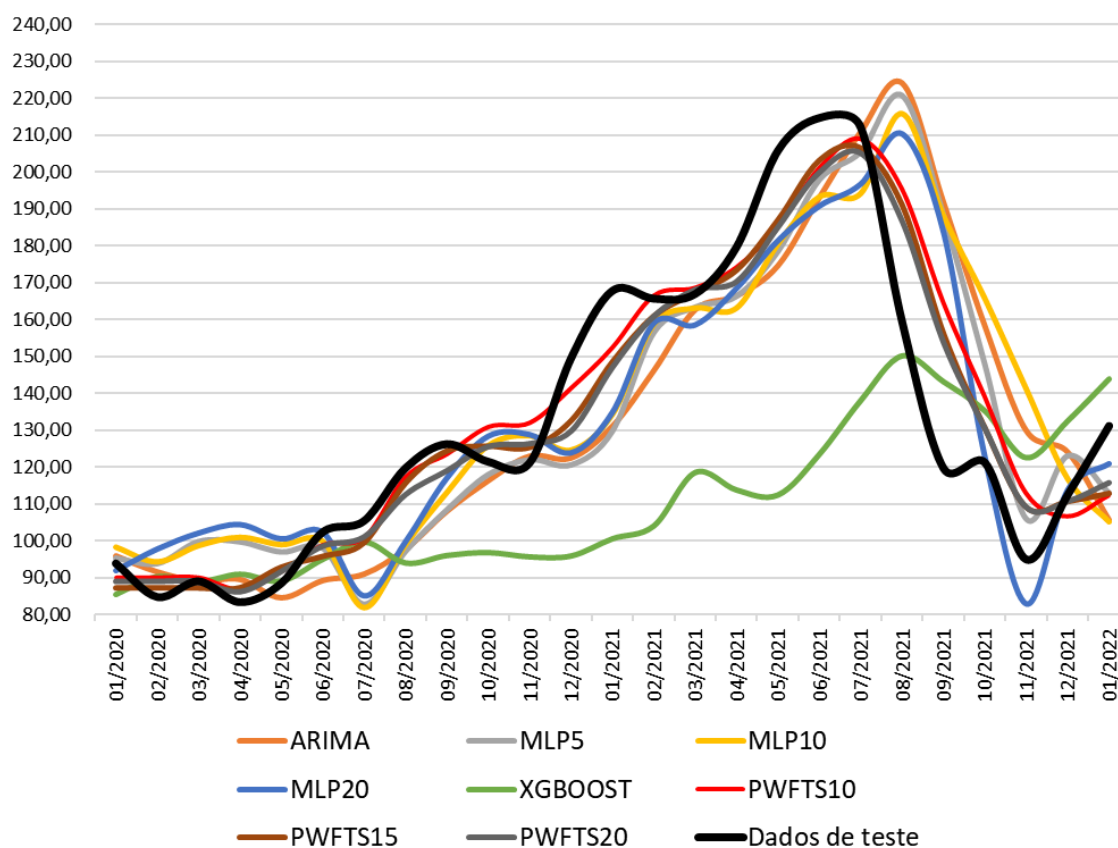


Figura 4.4: Comparação entre o valor real e o valor previsto para a série de teste univariada.

Fonte: Autor (2022).

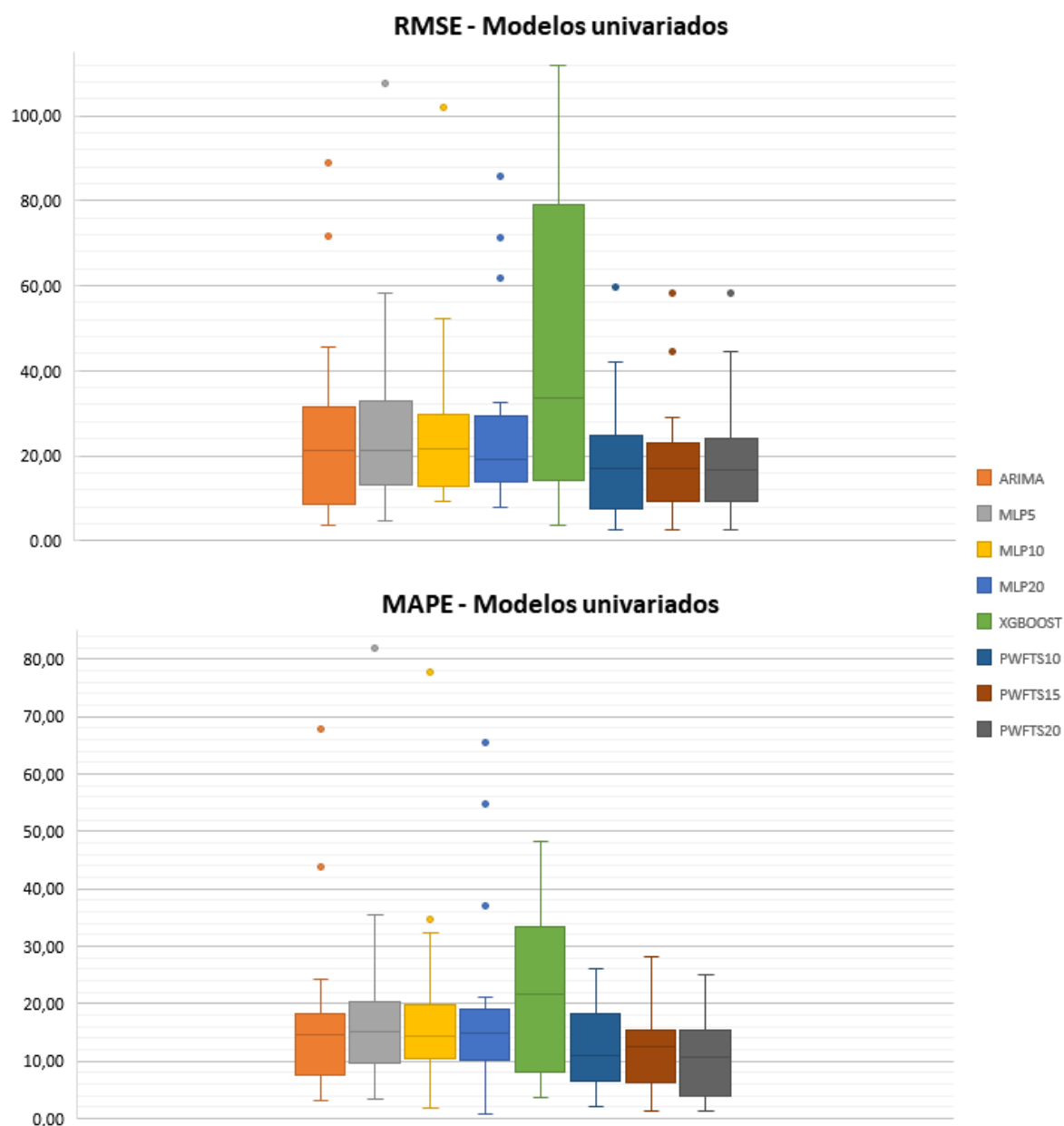


Figura 4.5: *Boxplots* para os subconjuntos de RMSE e MAPE das séries univariadas.

Fonte: Autor (2022).

4.4. Modelos preditivos multivariados

Para os modelos multivariados, a etapa de pré-processamento inclui diferenciação de primeira ordem e normalização em todo o conjunto de dados de treino. A Tabela 4.5 e a Tabela 4.6 exibem as estatísticas descritivas das métricas RMSE e MAPE, respectivamente, para os 23 subconjuntos de dados de teste da série multivariada, que também podem ser visualizadas nos *boxplots* da Figura 4.6.

Tabela 4.5: Estatísticas para o RMSE dos modelos multivariados.

Janela de testes	ARIMA	MLP5	MLP10	MLP20	XGBOOST	FDTFTS ID3	FDTFTS CART	FDTFTS RF
1	3,76	6,50	6,51	6,51	2,13	8,20	16,24	6,84
2	3,79	10,97	11,02	11,13	5,48	11,32	11,32	5,48
3	9,69	4,24	4,19	4,15	5,25	17,84	14,74	8,34
4	11,63	7,15	7,07	7,09	9,02	17,16	9,86	11,68
5	16,87	16,17	15,94	15,71	12,60	7,96	7,96	8,12
6	24,76	19,35	19,38	19,02	15,46	20,79	20,79	16,17
7	24,71	12,87	13,23	13,26	18,26	33,28	31,05	29,70
8	30,75	14,98	7,41	15,73	31,94	32,21	36,85	29,46
9	13,08	6,44	10,10	3,26	20,82	37,30	19,20	27,78
10	19,72	15,43	21,74	12,95	20,26	19,98	19,98	15,28
11	36,69	10,36	15,22	13,37	53,25	19,15	19,15	19,41
12	41,02	38,07	38,21	39,53	67,18	9,52	9,52	13,60
13	14,94	10,71	10,86	18,51	46,66	9,38	9,38	4,93
14	15,31	10,09	10,80	10,42	20,20	8,42	8,42	5,41
15	36,90	21,22	19,77	23,92	48,24	24,89	24,89	18,59
16	54,85	32,64	31,77	31,62	67,85	37,18	15,37	15,17
17	58,15	25,12	27,52	27,71	73,28	10,52	10,52	4,69
18	27,78	36,92	52,75	51,25	29,43	10,70	27,64	29,69
19	17,76	69,93	66,32	62,89	42,64	27,35	19,98	30,30
20	30,62	73,97	74,34	74,99	29,95	28,90	31,49	25,61
21	32,51	30,81	31,75	36,96	14,02	15,40	37,73	16,43
22	15,27	25,83	24,67	20,41	18,53	23,50	23,50	20,07
23	23,04	19,87	17,34	17,29	38,56	20,89	20,32	22,99
Mínimo	3,76	4,24	4,19	3,26	2,13	7,96	7,96	4,69
Mediana	23,04	16,17	17,34	17,29	20,82	19,15	19,20	16,17
Máximo	58,15	73,97	74,34	74,99	73,28	37,30	37,73	30,30
IQR	16,53	17,78	18,80	17,62	29,91	15,51	13,28	16,07

Fonte: Autor (2022).

Tabela 4.6: Estatísticas para o MAPE dos modelos multivariados.

Janela de testes	ARIMA	MLP5	MLP10	MLP20	XGBOOST	FDTFTS ID3	FDTFTS CART	FDTFTS RF
1	2,83	6,33	6,35	6,36	2,08	7,54	18,09	5,58
2	3,71	12,49	12,54	12,67	5,91	10,55	10,55	6,26
3	10,77	3,88	3,83	3,83	4,56	19,75	14,81	8,44
4	8,81	7,19	7,17	7,18	8,17	14,45	7,42	11,55
5	14,96	14,43	14,19	14,01	10,26	7,94	7,94	8,00
6	19,90	16,24	16,24	16,00	9,10	14,71	14,71	10,50
7	17,90	8,46	8,88	8,86	17,29	24,56	23,24	22,17
8	24,89	12,00	5,53	12,62	25,94	24,39	29,97	21,97
9	10,46	5,05	5,63	2,38	21,63	30,25	13,89	21,42
10	9,13	10,12	15,99	8,34	15,42	10,56	10,56	7,27
11	19,75	6,43	9,75	8,01	33,95	11,94	11,94	11,85
12	24,49	23,04	23,12	23,90	41,44	4,08	4,08	6,86
13	8,79	6,28	6,37	10,85	27,50	5,53	5,53	2,77
14	6,52	5,44	5,41	5,21	10,62	3,71	3,71	2,35
15	14,09	7,75	7,24	11,82	24,13	10,81	10,81	7,82
16	24,30	14,43	14,02	13,95	32,89	16,57	7,51	7,29
17	26,68	11,83	12,98	13,07	34,59	3,90	3,90	1,97
18	9,69	14,61	25,45	21,42	14,02	4,57	12,06	12,00
19	10,96	42,95	53,96	42,15	26,92	15,59	13,15	18,27
20	23,16	64,15	64,43	64,94	21,42	18,23	20,33	16,95
21	28,29	27,14	28,04	32,94	11,19	11,64	30,47	12,30
22	12,95	21,82	20,72	16,25	18,14	18,51	18,51	15,27
23	26,21	17,23	14,53	14,51	32,34	22,95	14,74	18,54
Mínimo	2,83	3,88	3,83	2,38	2,08	3,71	3,71	1,97
Mediana	14,09	12,00	12,98	12,67	18,14	11,94	12,06	10,50
Máximo	28,29	64,15	64,43	64,94	41,44	30,25	30,47	22,17
IQR	14,32	9,92	11,71	7,95	16,77	10,63	8,73	9,04

Fonte: Autor (2022).

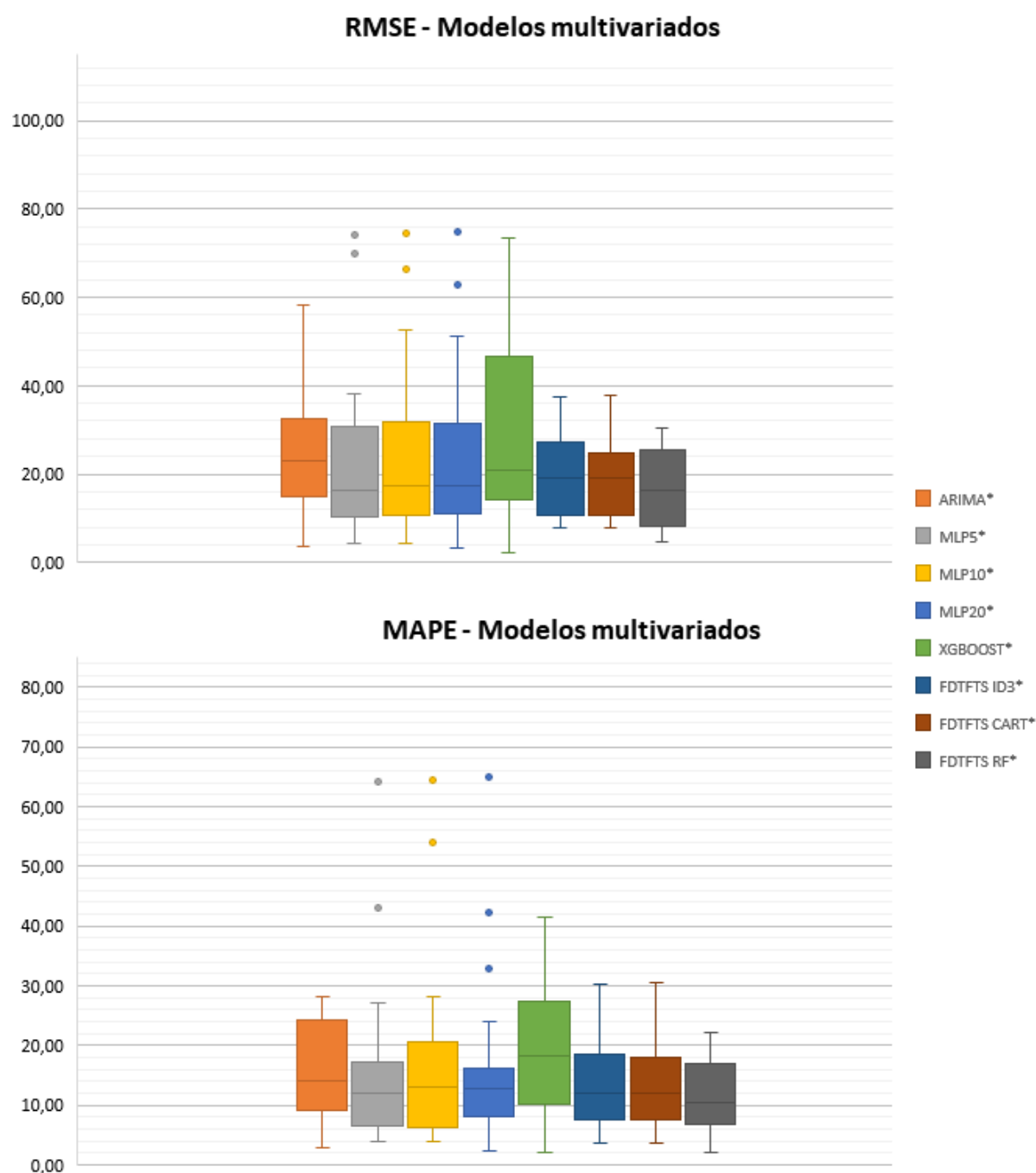


Figura 4.6: *Boxplots* para os subconjuntos de RMSE e MAPE das séries multivariadas.

Fonte: Autor (2022).

Acerca do modelo ARIMA, a função “*auto.arima()*” retornou os coeficientes $p = 3$, $d = 1$, $q = 0$ e $P = 1$. No contexto multivariado, a inclusão das variáveis explanatórias melhorou as métricas do ARIMA univariado, com menores RMSE e MAPE. Nos modelos MLP, o processo foi similar aos métodos univariados, considerando-se a variação do número de neurônios da camada oculta (*hd*) em 5, 10 e 20, e tomando a taxa de aprendizado (*eta*) como 0,4. Os *lags* 1,

2, 3 e 4 foram utilizados no treinamento dos modelos, sendo que cada modelo MLP passou por 10 rodadas de treino e teste. Para os modelos nebulosos FDT-FTS, considerou-se os parâmetros: número de árvores igual a 10 ($n_estimators$) e nível de profundidade igual a 20 (max_depth).

De modo geral, é possível notar que nem sempre a inclusão de variáveis exógenas à série *target* implica em melhores resultados. Contudo, a maior parte dos modelos testados aumentou significativamente o potencial de generalização após a introdução do conjunto de séries correlacionadas.

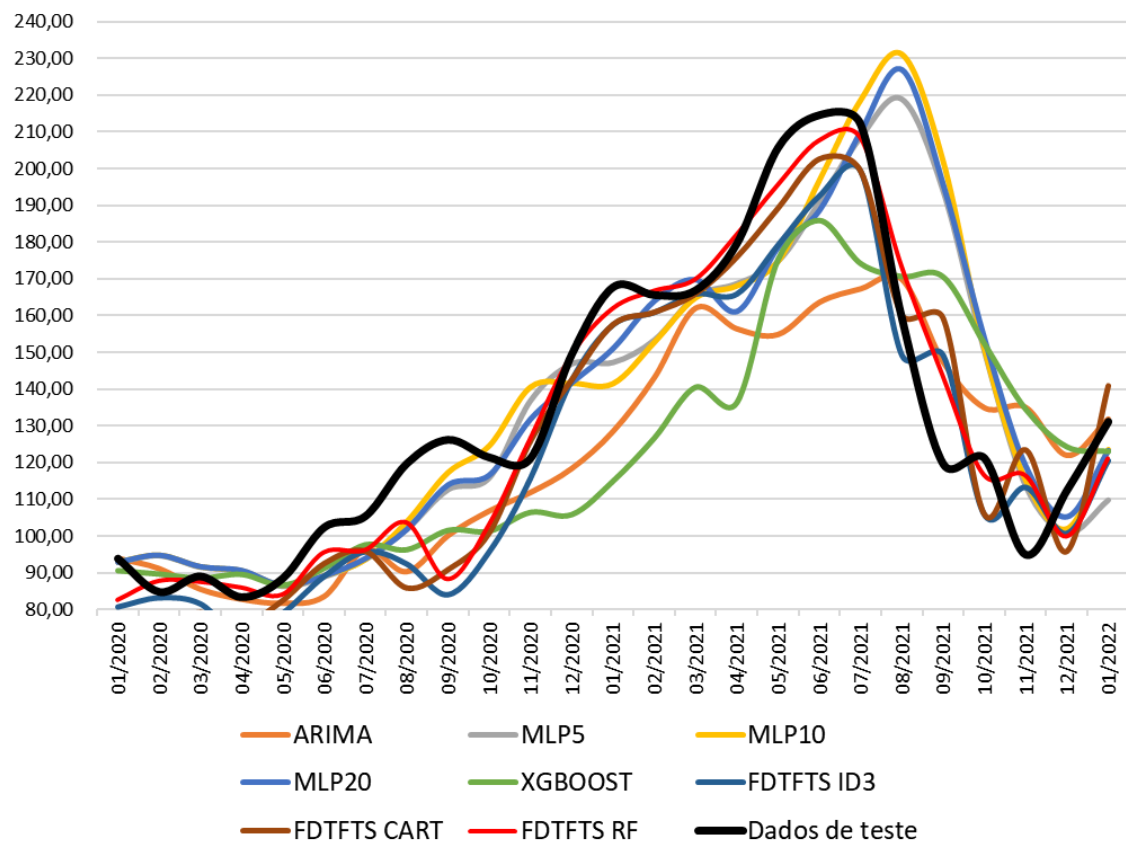


Figura 4.7: Comparação entre o valor real e o valor previsto para a série de teste multivariada.

Fonte: Autor (2022).

Em comparação às séries univariadas, o modelo ARIMA multivariado apresentou maior mediana na métrica RMSE, porém com menor dispersão ao redor, e ausência de *outliers*. Os modelos MLP, todavia, apresentaram comportamento inverso, reduzindo a mediana e o valor dos *outliers*, e aumentando a dispersão. O *Xgboost* multivariado foi o modelo que obteve maior salto de poder de generalização entre os testados, quando comparado ao *Xgboost* univariado, reduzindo o RMSE em 38% e o MAPE em 13%. Neste sentido, destaca-se novamente os

modelos nebulosos pela ausência de valores discrepantes, especialmente o FDTFTS_CART e o FDTFTS_RF, que apresentaram ainda a menor dispersão e a menor mediana, respectivamente, em ambas as métricas. A comparação entre os valores reais da série e os valores preditos por cada modelo para a série multivariada pode ser vista na Figura 4.7.

Os indicadores RMSE e MAPE permitem verificar o desempenho das previsões ao longo das janelas de treino e teste. Desta forma, se configuram como importantes parâmetros para comparação de acuracidade entre os modelos. Todavia, é desejável que os modelos testados tenham desempenho superior ao método ingênuo. Caso contrário, seria preferível optar por este, de fácil implementação. Para tanto, é estudado o coeficiente U de *Theil*.

A Tabela 4.7 mostra dispõe o coeficiente U de *Theil* para cada um dos modelos, considerando todas as janelas de dados. Neste caso, nota-se que nem todos os métodos superam a qualidade de previsão do método ingênuo. Em linhas gerais, o desempenho do modelo *Naive* decai à medida que o horizonte de previsão aumenta. Uma vez que este estudo trabalha para previsões de curto prazo, é esperado um desempenho razoável do *Naive*.

Os modelos univariados são inferiores ao *Naive*, com exceção dos nebulosos PWFTS de terceira ordem. De todo modo, nota-se que os modelos ARIMA, MLP e *Xgboost* de fato melhoram a qualidade das previsões com a introdução das variáveis preditoras. Ademais, destaca-se neste ponto as métricas alcançadas pelos modelos nebulosos, univariados e multivariados, que superam o modelo ingênuo, juntamente com os modelos MLP de 10 e 20 neurônios na camada oculta.

Tabela 4.7: U-*Theil* para os modelos testados.

Modelos univariados	ARIMA	MLP5	MLP10	MLP20	XGBOOST	PWFTS 10_3	PWFTS 15_3	PWFTS 20_3
U-Theil	1,145	1,128	1,066	1,026	1,180	0,892	0,887	0,878
Modelos univariados	ARIMA	MLP5	MLP10	MLP20	XGBOOST	FDTFTS ID3	FDTFTS CART	FDTFTS RF
U-Theil	1,025	1,110	0,976	0,962	1,150	0,879	0,894	0,858

Fonte: Autor (2022).

Considerando as métricas alcançadas pelos modelos nebulosos multivariados, verifica-se, na Figura 4.8, um recurso presente nos métodos FDT-FTS: o grau de importância das variáveis preditoras na construção das árvores de decisão dos modelos.

Nota-se que o primeiro *lag* da série principal é o atributo de maior contribuição nas previsões, o que já era esperado, considerando a natureza temporal da série, com forte autocorrelação entre os *lags* vizinhos. Nos três algoritmos, o excedente produtivo de minério de ferro da Vale S.A. aparece como a variável exógena de maior significância, atestando a importância desta variável como preditora da série *target*, dentro do conjunto de dados avaliado.

A Tabela 4.8 e a Tabela 4.9 mostram as estatísticas descritivas para as métricas de erro de previsão RMSE e MAPE ao longo de todas as 23 janelas deslizantes, desta vez desconsiderando o excedente de produção da Vale S.A. O intuito é confirmar o papel do excedente produtivo da Vale S.A. na qualidade das previsões, e verificar a acurácia dos modelos sem essa variável como preditora.

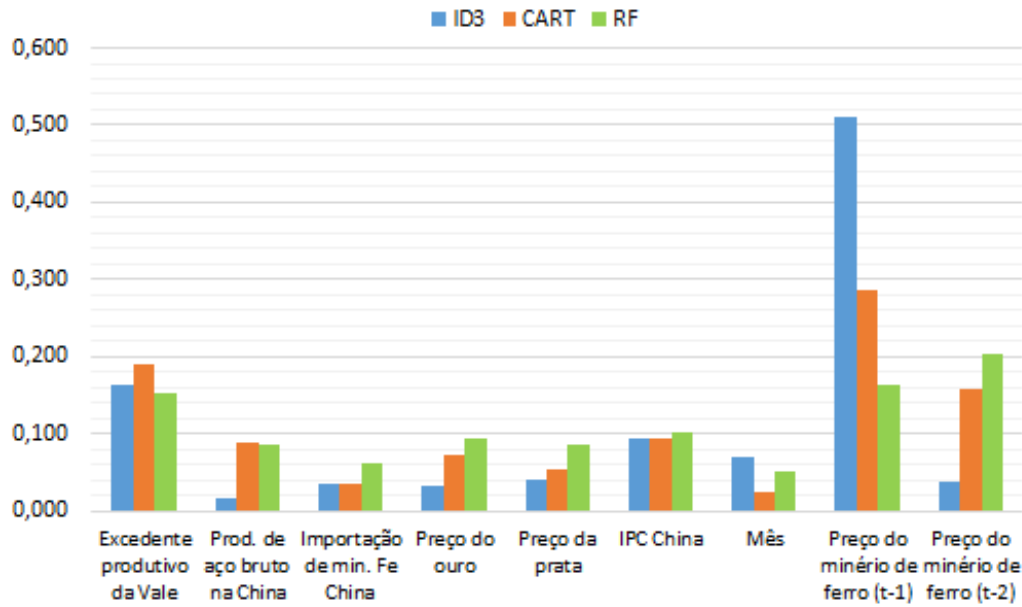


Figura 4.8: Grau de importância de cada variável preditora nos modelos FDT-FTS.

Fonte: Autor (2022).

Tabela 4.8: Estatísticas para o RMSE dos modelos sem o excedente produtivo da Vale S.A.

Modelo	ARIMA	MLP5	MLP10	MLP20	XGBOOST	FDTFTS ID3	FDTFTS CART	FDTFTS RF
Mínimo	3,82	4,21	4,23	4,19	2,13	11,40	5,31	4,66
Mediana	24,92	17,73	17,38	17,32	20,82	19,95	17,04	15,27
Máximo	58,06	110,55	96,41	99,38	92,06	37,27	38,73	43,28
IQR	22,09	17,91	17,91	17,94	31,01	11,74	15,58	13,19

Fonte: Autor (2022).

Tabela 4.9: Estatísticas para o MAPE dos modelos sem o excedente produtivo da Vale S.A.

Modelo	ARIMA	MLP5	MLP10	MLP20	XGBOOST	FDTFTS ID3	FDTFTS CART	FDTFTS RF
Mínimo	3,35	3,69	3,86	3,84	2,08	4,32	3,99	4,90
Mediana	17,19	12,58	12,99	13,01	18,14	12,56	12,54	10,06
Máximo	49,61	70,00	65,41	65,37	45,10	31,61	31,47	34,37
IQR	15,71	12,53	9,15	9,11	18,74	7,70	11,72	8,04

Fonte: Autor (2022).

Nas tabelas acima, os valores em negrito destacam os parâmetros que são maiores que os relacionados na Tabela 4.5 e na Tabela 4.6. Ser maior, nesse caso, indica piora no poder de generalização do modelo. É válido ressaltar, contudo, que este comportamento não é unânime. Observando inicialmente o RMSE, nota-se, por exemplo, que os modelos *Xgboost*, FDTFTS_CART e FDTFTS_RF obtiveram melhores medianas e mínimos. Em relação ao MAPE, o modelo MLP10, por exemplo, apresentou uma mediana menor, com menor dispersão nos erros ao longo das janelas de teste, similar ao ocorrido com o modelo FDTFTS_RF. De modo geral, avalia-se que a inclusão do excedente produtivo da Vale S.A. como variável preditora na previsão de preços de minério de ferro é positiva para a maior parte dos modelos testados, que apresentam ganhos no poder preditivo.

5. Conclusões

O presente trabalho propôs um estudo comparativo entre modelos de previsão para a série de preços de minério de ferro refinado em 62%, utilizando para tal técnicas de inteligência computacional, considerando modelos baseados em redes neurais, árvores de regressão e lógica nebulosa, com o intuito de contribuir na etapa de planejamento operacional das empresas mineradoras em um horizonte de previsão de curto prazo.

A análise de correlação e causalidade entre as variáveis exógenas permitiu identificar os melhores preditores para a construção dos modelos multivariados. Neste aspecto, é válido destacar o considerável grau de significância do excedente produtivo de minério de ferro da Vale S.A. nos modelos multivariados, que apresentaram melhorias no poder preditivo.

Ao comparar as métricas RMSE e MAPE nas janelas de previsão, conclui-se ainda que a inclusão de variáveis exógenas à série principal não implica necessariamente em maior acuracidade. Contudo, a maior parte dos modelos testados aumentou significativamente o potencial de generalização após a introdução do conjunto de séries correlacionadas. Vale ressaltar ainda que nem todos os modelos superaram o método ingênuo, especialmente os univariados. De certo modo, é esperado que modelos ingênuos tenham boas métricas em séries temporais de memória longa, dado o alto grau de autocorrelação entre os *lags*, como observado para a série de preços de minério de ferro utilizada neste estudo.

Em última análise, foi possível observar a superioridade dos modelos baseados em lógica nebulosa na previsão da série de preços de minério de ferro, com destaque para os modelos PWFTS univariados de terceira ordem com 15 e 20 partições, e os modelos FDT-FTS multivariados baseados nos algoritmos CART e *Random Forest*, que apresentaram boas métricas de acuracidade, e superioridade em relação à previsão ingênua e aos demais modelos.

6. Trabalhos futuros

Em trabalhos futuros focados na previsão da série de preços de minério de ferro, sugere-se o estudo preditivo em horizontes de previsão de médio prazo, a fim de que o resultado contribua não somente com decisões operacionais, mas também com o planejamento tático das empresas mineradoras.

Seguindo esta linha, indica-se também avaliar o impacto do excedente produtivo de outras empresas mineradoras além da Vale S.A. que também exercem influência no mercado de *commodities* minerais, especialmente de minério de ferro, e também verificar outras variáveis internas da Vale S.A., especialmente as de conhecimento prévio da equipe técnica.

Sugere-se, por fim, testar outros métodos de previsão de séries temporais, como os baseados em redes neurais recorrentes, por exemplo, ou ainda modelos convolucionais focados em *big data*.

Referências Bibliográficas

AMINROSTAMKOLAEI, B.; SCROGGS, J. S.; BORGHEI, M. S.; SAJDARI-VAIGHANI, A.; MOHAMMADI T.; POURKAZEMI, M. H. Valuation of a hypothetical mining project under commodity price and exchange rate uncertainties by using numerical methods. **Resources Policy**, v. 52, p. 296–307, 2017.

BAIN, C. Guide to commodities. **The Economist**. Hoboken, NJ: Wiley, 2013.

BAUER, A.; ZÜFLE, M.; HERBST, N.; ZEHE, A.; HOTH, A.; KOUNEV, S. Time Series Forecasting for Self-Aware Systems. **Proceedings of the IEEE**, v. 108, n. 7, p. 1068–1093, 2020.

BEASLEY, J. E. **OR-notes**. Disponível em: <<http://people.brunel.ac.uk/~mastjjb/jeb/or/forecast.html>>. Acesso em: 1 fev. 2021.

BONTEMPI, G.; TAIEB, S. B.; LE BORGNE, Y. A. **Machine learning strategies for time series forecasting**. European business intelligence summer school. **Anais...**2012

BROWNLEE, J. **Statistical methods for machine learning: Discover how to transform data into knowledge with Python**. [s.l.] Machine Learning Mastery, 2018.

CALUDE, C. S.; LONGO, G. The deluge of spurious correlations in big data. **Foundations of science**, v. 22, n. 3, p. 595-612, 2017.

CHEN, S. M. Forecasting enrollments based on fuzzy time series. **Fuzzy sets and systems**, v. 81, n. 3, p. 311–319, 1996.

CHENG, C. H.; CHEN, C. H. Fuzzy time series model based on weighted association rule for financial market forecasting. **Expert Systems**, v. 35, n. 4, p. e12271, 2018.

COMEXSTAT. **Ministério da Indústria, Comércio Exterior e Serviços. Exportação e Importação geral. Base de dados 2019-2020**. Disponível em: <<http://comexstat.mdic.gov.br/pt/home>>. Acesso em: 20 dez. 2020.

COSTA, F. R. B. Fuzzy time series aplicada na predição de QOS em redes de computadores. 2018.

EGRIOGLU, E.; BAS, E.; YOLCU, U.; CHEN, M. Y. Picture fuzzy time series: Defining, modeling and creating a new forecasting method. **Engineering Applications of Artificial Intelligence**, v. 88, p. 103367, 2020.

- EWEEES, A. A.; ABD ELAZIZ, M.; ALAMEER, Z.; YE, H.; JIANHUA, Z. Improving multilayer perceptron neural network using chaotic grasshopper optimization algorithm to forecast iron ore price volatility. **Resources Policy**, v. 65, p. 101555, 2020.
- GEMAN, H.; SMITH, W. O. Theory of storage, inventory and volatility in the LME base metals. **Resources Policy**, v. 38, n. 1, p. 18-28, 2013.
- GUMUS, M.; KIRAN, M. S. **Crude oil price forecasting using XGBoost**. 2017 International Conference on Computer Science and Engineering (UBMK). **Anais...**2017
- HAYKIN, S. **Redes neurais: princípios e prática**. [s.l.] Bookman Editora, 2007.
- HWANG, J.; CHEN, S. M.; LEE, C. H. Handling forecasting problems using fuzzy time series. **Fuzzy Sets and Systems**, v. 100, n. 1, p. 217–228, 1998.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: Principles and Practice**, 2020. Disponível em: <<https://otexts.com/fpp2/>>. Acesso em: 19 dez. 2020
- LI, D.; MOGHADDAM, M. R.; MONJEZI, M.; JAHED ARMAGHANI, D.; MEHRDANESH, A. Development of a group method of data handling technique to forecast iron ore price. **Applied Sciences**, v. 10, n. 7, p. 2364, 2020.
- LIVIERIS, I. E.; PINTELAS, E.; PINTELAS, P. A CNN–LSTM model for gold price time-series forecasting. **Neural Computing and Applications**, v. 32, n. 23, p. 17351–17360, 2020.
- MA, Y.; ZHEN, W. Market Fundamentals and Iron Ore Spot Prices*. **Economic Record**, v. 96, n. 315, p. 470–489, 2020.
- MARKETINDEX. **Iron Ore**. Disponível em: <<https://www.marketindex.com.au/iron-ore>>. Acesso em: 1 fev. 2021.
- NIELSEN, A. **Practical time series analysis: prediction with statistics and machine learning**. [s.l.] “O’Reilly Media, Inc.”, 2019.
- NIST/SEMATECH. e-Handbook of Statistical Methods. **e-Handbook of Statistical Methods**, 2012.
- PATRICIO, M. Correlações e causalidade: Breves reflexões numa perspectiva económica. **Revista Jurídica Luso-Brasileira**. Lisboa, ano 2, n. 4, p. 1347-1380, 2016.
- PENG, Z.; HUANG, Q.; HAN, Y. **Model Research on Forecast of Second-Hand House Price in Chengdu Based on XGboost Algorithm**. 2019 IEEE 11th International Conference

on Advanced Infocomm Technology (ICAIT). **Anais...**2019

ROKACH, L.; MAIMON, O. Decision Trees. In: **The Data Mining and Knowledge Discovery Handbook**. [s.l: s.n.]. v. 6p. 165–192.

SILVA, P. C. L. **Scalable Models for Probabilistic Forecasting with Fuzzy Time Series**. Tese (Doutorado) — Machine Intelligence and Data Science Laboratory, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 2019.

SILVA, R. R. C. **Previsão de Séries Temporais Multivariadas Utilizando Modelo Híbrido Interpretável Com Árvores de Decisão Fuzzy**. Dissertação (Mestrado) — Machine Intelligence and Data Science Laboratory, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 2021.

SETH, A. Granger causality. **Scholarpedia**, v. 2, n. 7, p. 1667, 2007.

SONG, Q.; CHISSOM, B. S. Fuzzy time series and its models. **Fuzzy Sets and Systems**, v. 54, n. 3, p. 269–277, 1993.

TRADEMAP. **Trade statistics for international business development**. Disponível em: <<https://www.trademap.org>>. Acesso em: 12 dez. 2020.

WANG, J. J.; WANG, J. Z.; ZHANG, Z. G.; GUO, S. P. Stock index forecasting based on a hybrid model. **Omega**, v. 40, n. 6, p. 758–766, 2012.

WOLPERT, D. H.; MACREADY, W. G. No free lunch theorems for optimization. **IEEE Transactions on Evolutionary Computation**, v. 1, n. 1, p. 67–82, 1997.

WORKING, H. The theory of price of storage. **The American Economic Review**, v. 39, n. 6, p. 1254-1262, 1949.

XU, M.; WATANACHATURAPORN, P.; VARSHNEY, P. K.; ARORA, M. K. Decision tree regression for soft classification of remote sensing data. **Remote Sensing of Environment**, v. 97, n. 3, p. 322–336, 2005.