

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**Previsão do Preço de Energia no Mercado de Curto Prazo:
Uma análise Combinada de Séries Temporais e Redes
Neurais Artificiais**

Yuri Belentani

Dissertação de Mestrado do Programa de Mestrado Profissional em
Matemática, Estatística e Computação Aplicadas à Indústria (MECAI)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Yuri Belentani

**Previsão do Preço de Energia no Mercado de Curto Prazo:
Uma análise Combinada de Séries Temporais e Redes
Neurais Artificiais**

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre – Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria.
VERSÃO REVISADA

Área de Concentração: Matemática, Estatística e Computação

Orientadora: Profa. Dra. Roseli Aparecida Leandro

**USP – São Carlos
Julho de 2023**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

B428p Belentani, Yuri
Previsão do Preço de Energia no Mercado de Curto
Prazo: Uma análise Combinada de Séries Temporais e
Redes Neurais Artificiais / Yuri Belentani;
orientadora Roseli Aparecida Leandro. -- São
Carlos, 2023.
83 p.

Tese (Doutorado - Programa de Pós-Graduação em
Mestrado Profissional em Matemática, Estatística e
Computação Aplicadas à Indústria) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2023.

1. Previsão Preço de Energia Mercado Livre (PLD).
2. Séries Temporais - SARIMA. 3. Previsão com LSTM.
4. Redes Neurais Artificiais. I. Aparecida Leandro,
Roseli, orient. II. Título.

Yuri Belentani

**Energy Price Forecast in the Short Term Market - A
Combined Analysis of Time Series and Artificial Neural
Networks**

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação - USP, in partial fulfillment of the requirements for the degree of the Master – Professional Masters in Mathematics, Statistics and Computing Applied to Industry. *FINAL VERSION*

Concentration Area: Mathematics, Statistics and Computing

Advisor: Profa. Dra. Roseli Aparecida Leandro

**USP – São Carlos
July 2023**

*Ao meu filho,
Henrique.*

AGRADECIMENTOS

Agradeço à minha família por me dar todas as condições necessárias para que eu pudesse ter o privilégio de sempre estudar e adquirir conhecimento, e assim seguir evoluindo em minha jornada como ser humano.

Agradeço especialmente à minha esposa Fabíola, que se mantém sempre presente ao meu lado e me apoiando incansavelmente.

Também gostaria de agradecer especialmente às professoras Roseli Aparecida Leandro e Clélia Maria de Castro Toloi por toda a colaboração durante este trabalho.

“Há épocas em que a sociedade, tomada de pânico, se desvia da ciência e procura a salvação na ignorância.”
(Mikhail Saltykov-Stcherdrine)

RESUMO

BELENTANI, Y. **Previsão do Preço de Energia no Mercado de Curto Prazo: Uma análise Combinada de Séries Temporais e Redes Neurais Artificiais**. 2023. 83 p. Dissertação (Mestrado – Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2023.

Atualmente o Brasil e o mundo passam por uma transição energética, cuja economia de baixo carbono tem se efetivado através de fontes de energia cada vez mais limpas e com operações industriais mais eficientes. Além dos ganhos para o meio ambiente, esse caminho deve trazer benefícios econômicos para os agentes do sistema elétrico brasileiro através das diversas oportunidades que serão geradas.

Neste contexto a possibilidade de prever os preços de energia elétrica no mercado de curto prazo (MCP), ambiente de contratação livre, pode contribuir significativamente para a otimização do planejamento de contratação da demanda e uma melhor avaliação dos riscos, seja para os agentes geradores ou para a indústria e grandes consumidores de modo geral.

Dessa maneira, buscando contribuir com o aumento de previsibilidade dos agentes que atuam no mercado livre de energia elétrica brasileiro, o presente estudo se propõe prever o valor do Preço de Liquidação das Diferenças (PLD) para quatro semanas operativas subsequentes. Para isso, a ideia é utilizar uma combinação entre modelos AutoRegressivos Integrados de Médias Móveis com Sazonalidade (SARIMA) e redes neurais artificiais recorrentes (LSTM). Primeiramente, serão previstos as quatro semanas operativas subsequentes para as variáveis predictoras e então alimentar uma Rede Neural Artificial (RNA) para produzir o PLD.

A metodologia traz como resultado previsões com níveis de acuracidade satisfatórios, principalmente pelo fator combinatório de cada técnica.

Palavras-chave: Preço da Liquidação das Diferenças, ambiente de contratação livre, mercado de energia elétrica brasileiro, inteligência artificial, redes neurais artificiais, LSTM, séries temporais, ARIMA, SARIMA, modelagem combinada.

ABSTRACT

BELENTANI, Y. **Energy Price Forecast in the Short Term Market - A Combined Analysis of Time Series and Artificial Neural Networks**. 2023. 83 p. Dissertação (Mestrado – Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2023.

Currently, Brazil and the world are going through an energy transition, whose low-carbon economy has been effected through increasingly clean energy sources and with more efficient industrial operations. In addition to the gains for the environment, this path should bring economic benefits to the agents of the Brazilian electrical system through the various opportunities that will be generated.

In this context, the possibility of forecasting electricity prices in the short-term market (MCP), a free contracting environment, can contribute to the optimization of demand contracting planning and a greater assessment of risks, whether for generating agents or for industry and large consumers in general.

In this way, seeking to contribute to the increase of predictability of the agents that act in the Brazilian electric energy free market, the present study proposes to forecast the value of the difference settlement price (acronym in portuguese PLD) for four subsequent operative weeks. For this, the idea is to use a combination between time series models (SARIMA) and recurrent artificial neural networks (LSTM). Firstly, the four subsequent operative weeks forecast for the predictor variables and then feed the artificial neural network (ANN) to produce the PLD.

The methodology results in prediction with strong levels of accuracy, mainly due to the combinatorial factor of each technique.

Keywords: Keywords: Price for Settlement of Differences, free contracting environment, market for Brazilian electricity, artificial intelligence, artificial neural networks, LSTM, time series, ARIMA, SARIMA, combined modeling..

LISTA DE ILUSTRAÇÕES

Figura 1 – Setor Industrial: Decomposição da Variação do Consumo Final	21
Figura 2 – Histórico PLD	22
Figura 3 – MCP - Mercado de Curto Prazo	23
Figura 4 – Evolução de unidades consumidoras no ACL por estado	24
Figura 5 – Mercado de Energia Elétrica Brasileiro	25
Figura 6 – Operação e Planejamento do Sistema Elétrico	26
Figura 7 – Metodologia	27
Figura 8 – Arquitetura do Sistema proposto	32
Figura 9 – Exemplo de Série Temporal	35
Figura 10 – Decompondo a Série Temporal	37
Figura 11 – Distribuição Resíduo	37
Figura 12 – Funções de Autocorreção e Autocorrelação Parcial	38
Figura 13 – Série de Dados Após Aplicar Diferenciação	41
Figura 14 – MMS - Previsto x Real	42
Figura 15 – Previsão Série Temporal Utilizando ARIMA	44
Figura 16 – Função de Autocorreção Residual	44
Figura 17 – Neurônio Biológico	48
Figura 18 – Perceptron	50
Figura 19 – Exemplo de classificação com Perceptron	51
Figura 20 – Representação Gráfica Classificação XOR	52
Figura 21 – MLP - Representação Gráfica	52
Figura 22 – MLP - Representação Gráfica	54
Figura 23 – MLP - Representação Gráfica	55
Figura 24 – Exemplo do Formato da Base de Dados PLD	58
Figura 25 – Histórico PLD	58
Figura 26 – Histograma do PLD SE (R\$/MWh)	59
Figura 27 – Comparativo entre Preços Máximos, Mínimos e Médio	60
Figura 28 – Correlação entre as Variáveis	61
Figura 29 – Relação Nível dos Reservatórios x Preço de Energia	62
Figura 30 – Demanda de Energia Máxima	63
Figura 31 – Retorno das Séries Utilizadas	64
Figura 32 – Boxplot CMO	64
Figura 33 – Decomposição da Série CMO	65

Figura 34 – Decomposição das Série Demanda Máxima	65
Figura 35 – Decomposição das Série Energia Armazenada	65
Figura 36 – Autocorrelação	66
Figura 37 – Demanda Máxima: Ajustamento x Real	67
Figura 38 – Demanda Máxima: Ajustamento x Real - Base de Dados Completa	67
Figura 39 – Análise Resíduos Ajustamento Demanda Máxima	68
Figura 40 – Energia Armazenada - EAR: Ajustamento x Real	68
Figura 41 – Energia Armazenada - EAR: Ajustamento x Real - Base de Dados Completa	69
Figura 42 – Análise Resíduos Ajustamento Demanda Máxima	69
Figura 43 – CMO: Ajustamento x Real	70
Figura 44 – CMO: Ajustamento x Real - Gráfico Completo	70
Figura 45 – Análise Resíduos Previsão CMO	71
Figura 46 – Representação da Estrutura dos Dados	74
Figura 47 – Representação da Estrutura dos Dados - Coluna PLD Deslocada	75
Figura 48 – Exemplo de Implementação LSTM em Python	75
Figura 49 – Rede LSTM x RMSE	76
Figura 50 – Erros Modelagem	76
Figura 51 – Previsão PLD: Real x Ajustamento - Dados Normalizados	77

LISTA DE TABELAS

Tabela 1 – Capacidade Instalada no SIN no final de maio de 2019	25
Tabela 2 – RMSE para diferentes configurações de rede neural	33
Tabela 3 – Análise descritiva PLD	58
Tabela 4 – Análise de Variabilidade	63
Tabela 5 – Estacionariedade das Variáveis Preditoras	66
Tabela 6 – Previsão das Variáveis Preditoras	71
Tabela 7 – Valores Previstos x Real	77

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Breve Explicação sobre o Setor Elétrico Brasileiro	24
1.2	Metodologia	27
1.3	Estrutura e Tópicos do Trabalho	28
2	REVISÃO BIBLIOGRÁFICA	31
2.1	Trabalhos Gerais com Séries Temporais	31
2.2	Trabalhos com Séries Temporais e Setor Elétrico	33
3	SÉRIES TEMPORAIS	35
3.1	Decomposição de Séries Temporais	36
3.2	Autocorrelação	38
3.3	Estacionariedade	39
3.3.1	<i>Teste de Dickey-Fuller Aumentando (ADF)</i>	39
3.3.2	<i>Diferenciação: Removendo a Tendência</i>	40
3.3.3	<i>Transformação Logarítmica</i>	40
3.4	Modelos de Previsão	41
3.4.1	<i>Médias Móveis Simples</i>	41
3.4.2	<i>Modelo ARIMA</i>	42
3.4.3	<i>Modelo SARIMA</i>	44
4	REDES NEURAIS ARTIFICIAIS	47
4.1	O Neurônio Biológico	48
4.2	Primeiro Modelo de Neurônio Artificial	49
4.3	Perceptron	49
4.4	O Perceptron Multi Camadas	51
4.5	Algoritmo de Retropropagação	52
4.6	Redes Neurais Recorrentes	54
4.7	Redes de Memória Longa de Curto Prazo	54
5	DESENVOLVIMENTO DO MODELO DE PREVISÃO DAS VARIÁVEIS PREDITORAS	57
5.1	Analisando os Dados	57
5.1.1	<i>PLD (R\$/MWh - Histórico de Preços de Energia)</i>	57

5.1.2	<i>Variáveis Predictoras</i>	59
5.2	Previsão das Variáveis Predictoras usando Modelos SARIMA	63
5.2.1	<i>Decompondo as Séries</i>	64
5.2.2	<i>Modelagem e Previsão</i>	67
6	PREVISÃO COM RNA	73
6.1	Explorando a Lógica da Análise Combinada	73
6.2	Modelagem da LSTM	74
6.3	Resultado Previsão	77
7	CONCLUSÃO	79
7.1	Trabalhos Futuros	80
	REFERÊNCIAS	81

INTRODUÇÃO

Atualmente há diversas ações no mundo e no Brasil, envolvendo iniciativa pública e privada, para promover um movimento de descarbonização (redução da emissão de dióxido de carbono e outros gases que causam efeito estufa) através da transição para uma matriz energética mais limpa que, a despeito dos recentes impactos nas cadeias de suprimento global devido à guerra Rússia-Ucrânia e o COVID-19, deve seguir avançando. Onde uma das rotas mais importantes desse avanço é a eletrificação (ex.: substituição de motores à combustão na mobilidade urbana, geração distribuída e redes inteligentes para gerenciamento).

Dessa maneira, é óbvio esperar que o consumo de energia elétrica nos próximos anos aumente além do crescimento esperado dos países, exigindo dos agentes do sistema elétrico um maior planejamento na operação, seja tanto para aqueles que geram quanto para os consumidores de médio e grande porte que compram energia no mercado livre.

Na Figura 1 tem-se a representação desse movimento pelo aumento do consumo de energia esperado no setor industrial brasileiro de acordo com as fontes entre os períodos de 2021 e 2031. A eletricidade deverá ter sua participação crescendo acima de todas as demais energias.

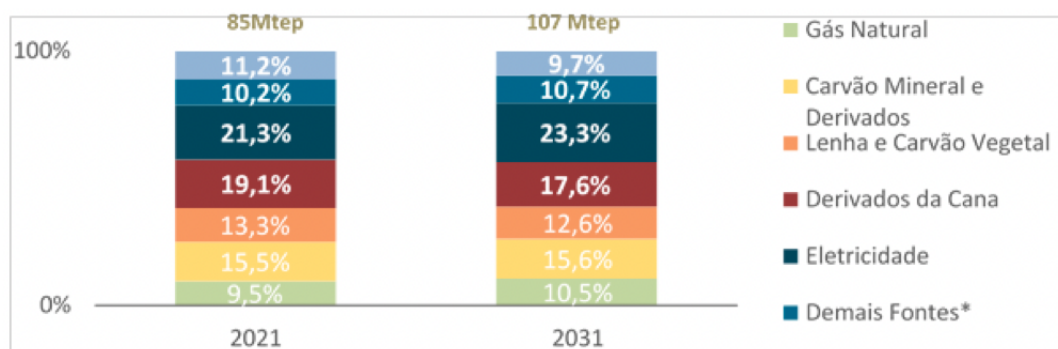


Figura 1 – Setor Industrial: Decomposição da Variação do Consumo Final

Fonte: (EPE, 2022)

Assim, existe uma busca por ferramentas e modelos que auxiliem a tomada de decisão dos agentes, principalmente em um cenário bastante volátil como o vivido nos últimos anos no Brasil e no mundo. Entender o comportamento dos preços de commodities, como petróleo, energia e minério, é muito valioso para a tomada de decisão refletindo na maior otimização de recursos e impactando diretamente o resultado operacional das companhias. Entretanto, a tarefa de prevê-los é realmente desafiadora, visto que há uma influência de muitas variáveis na composição do seu preço, o que resulta em um equacionamento não trivial.

Portanto, o presente trabalho tem como objetivo contribuir com o estudo da previsão do preço de energia no mercado livre, o PLD (Preço da Liquidação das Diferenças), fazendo uso de técnicas de previsão combinadas para assim obter uma melhor performance.

Primeiramente, serão utilizados os conceitos estatísticos de séries temporais para realizar a previsão das variáveis preditoras, a serem escolhidas de acordo com sua influência na variável target, e então alimentar uma rede neural recorrente, juntamente com os dados históricos do PLD, e finalmente prever o preço da energia para quatro semanas operativas subsequentes.

Na Figura 2 tem-se representado o histórico do PLD entre 06/01/2001 e 01/01/2021.

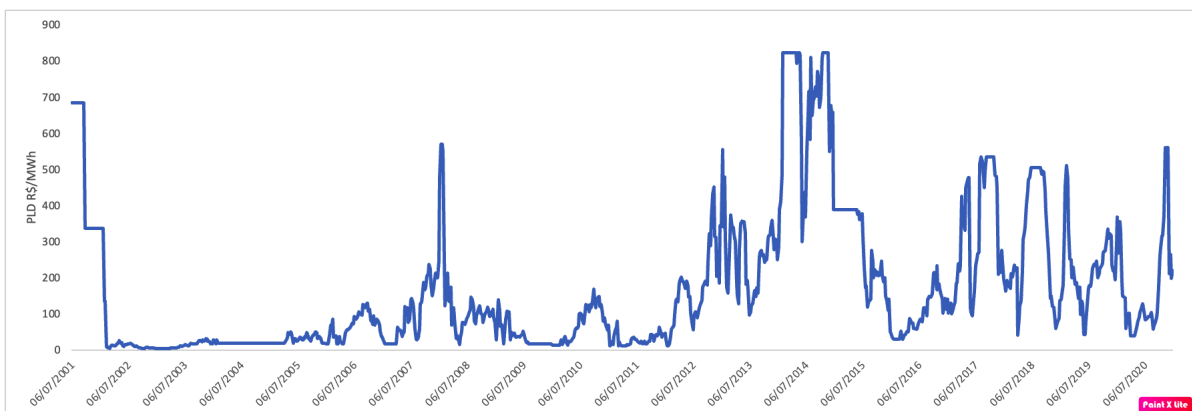


Figura 2 – Histórico PLD

Fonte: (EPE, 2022)

O trabalho revisita algumas teorias de séries temporais e redes neurais artificiais, além de obter uma análise exploratória dos dados em estudo e criar a modelagem necessária para a previsão. Por último, será feito uma comparação com os dados reais e avaliação sobre a confiabilidade do modelo proposto.

Entende-se que a aplicação das técnicas e o modelo proposto não estão limitados somente ao preço de energia e podem ser estendidos para outros ramos, desde previsão de commodities até o mercado de capitais, uma vez que seja utilizado outras variáveis preditoras de acordo com o que se espera prever.

O PLD é o preço base para o Mercado Livre de energia, onde os agentes podem comercializar entre si a compra e venda conforme suas necessidades e disponibilidades. Para

exemplificar sua importância, suponha-se hipoteticamente que um consumidor no mercado livre tenha um contrato de fornecimento de energia para suprir a sua demanda e no final do período contratual medido tenha extrapolado o que havia sido contratado por algum motivo específico, sazonalidade ou uma falha de projeção e planejamento. Haverá duas possibilidades quando se assume a obrigatoriedade de comercialização:

1. precificar a diferença via o PLD no mercado de curto prazo (MCP);
2. contratar a diferença com um agente gerador ou comercializadora de energia.

Outra situação pode ser a sobra de energia, ou seja, o agente consumiu menos que o contratado. Da mesma maneira, é possível negociar esse excedente de energia via MCP ou através de um agente comercializador.

A Figura 3 ilustra as possíveis situações descritas.

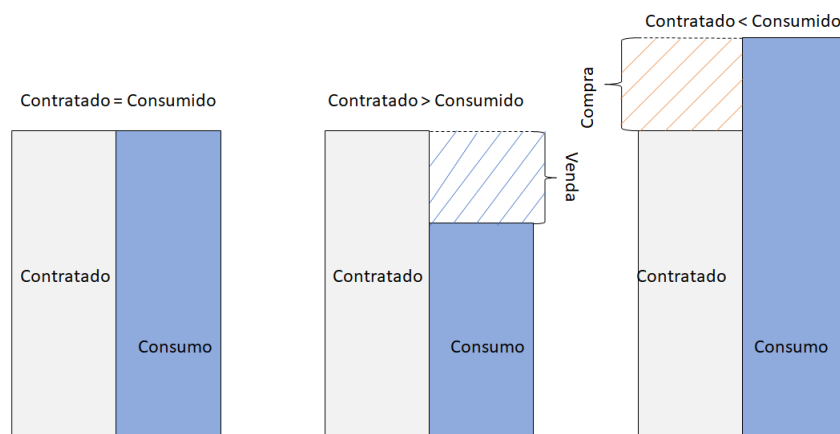


Figura 3 – MCP - Mercado de Curto Prazo

Fonte: Elaborado pelo autor

Além de ser fundamental o conhecimento do seu perfil de consumo, é de grande valia para o agente consumidor ter uma visão de médio prazo sobre o comportamento futuro do PLD. Por exemplo, considerando novamente a situação hipotética, conforme já descrito acima, onde a demanda contratada seja inferior à consumida, saber a tendência (aumento ou diminuição) do PLD colabora na tomada de decisão. Pois, o agente consumidor sabendo que há uma tendência de queda de preço de energia no MCP ele pode optar por contratar uma demanda menor e comprar a diferença via MCP e, portanto, criando uma economia muitas vezes expressiva para seu negócio. Ou ainda, simplesmente prever qual será o impacto do PLD no resultado das suas operações.

Para finalizar esta parte introdutória, uma vez já abordada a crescente eletrificação, e de forma geral o papel do PLD na dinâmica deste mercado, é importante complementar informando que existe uma crescente busca das empresas pelo ambiente de contratação livre, onde os agentes

têm um nível de liberdade maior para negociar, o que reforça a relevância da busca por métodos como o do presente estudo que fornecem subsídios para a tomada de decisão.

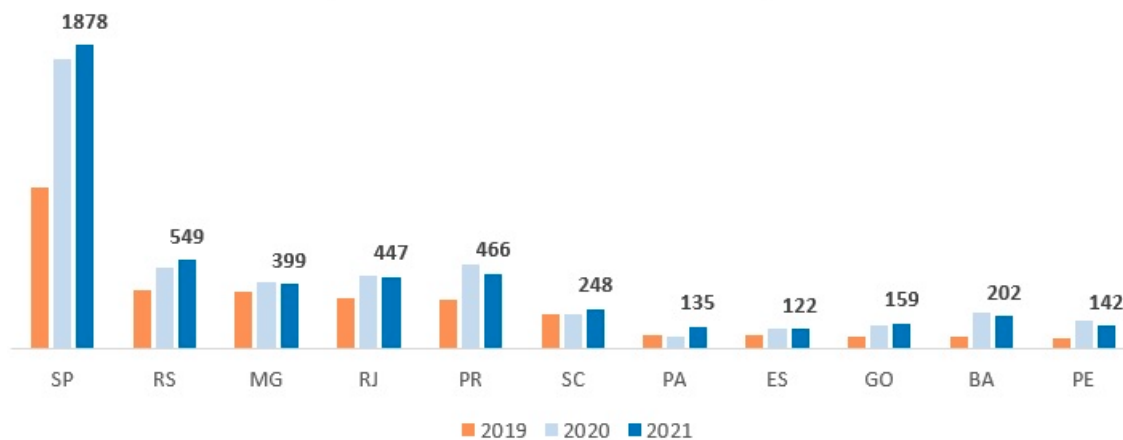


Figura 4 – Evolução de unidades consumidoras no ACL por estado

Fonte: (EPE, 2022)

1.1 Breve Explicação sobre o Setor Elétrico Brasileiro

Após a crise energética que o Brasil atravessou nos princípios dos anos 2000 foi identificada a necessidade de alteração da política energética do país, o que resultou em um novo modelo do sistema elétrico.

Para garantir a expansão da oferta, o governo (i) instituiu que cada consumidor do sistema deve estar 100% coberto por um contrato de suprimento de energia, o que significa que deve existir contratos para o suprimento de cada kWh provido ao sistema e (ii) segmentou o mercado de demanda em dois ambientes: o Ambiente de Contratação Regulada (ACR) e o Ambiente de Contratação Livre (ACL) (TOLMASQUIM, 2015, p. 105). Na Figura 5 tem-se um esquema simplificado dessa sistemática.

O ACR é regulado pela Agência Nacional de Energia Elétrica (ANEEL) e formado basicamente pelas concessionárias de energia e clientes cativos, sendo que a contratação da energia para essa demanda é realizada através de leilões reversos. Enquanto no ACL a regulação é feita pela Câmara de Comercialização de Energia Elétrica (CCEE) e as negociações entre os agentes geradores e consumidores são firmadas através de contratos de compra e venda de energia. É esperado que o interesse de contratações no ACL cresça, como na Figura 4 com a evolução desse tipo de contratação desde 2019 até 2021.

Ainda sobre o ACL, a diferença entre a energia gerada e consumida por cada agente e aquela registrada em contrato na CCEE é liquidada no mercado de curto prazo em uma janela de tempo predefinida, normalmente um mês. A valoração da diferença é feita através do PLD que é calculado pela CCEE semanalmente para cada submercado (NE, SE/CO, S, N) e para cada

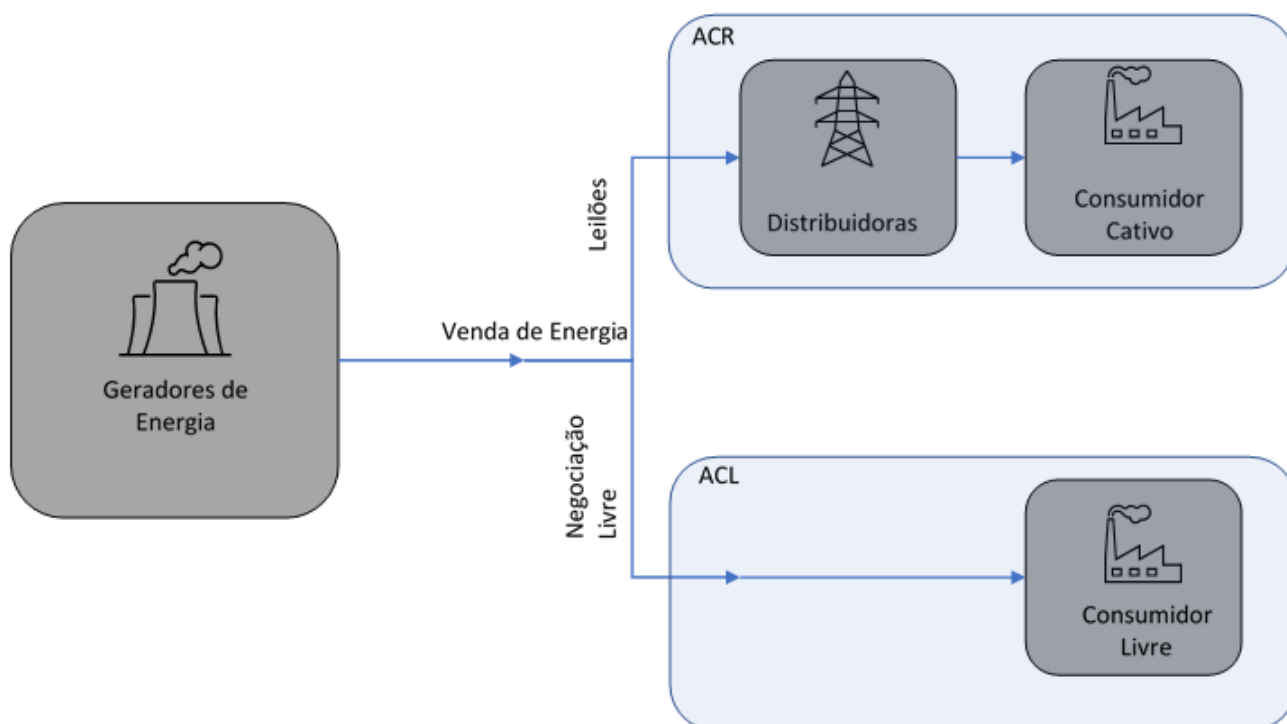


Figura 5 – Mercado de Energia Elétrica Brasileiro

Fonte: Elaborado pelo autor

patamar de carga (Leve, Média e Pesada), levando em consideração o custo marginal de operação (CMO). O CMO pode ser entendido como o custo necessário mínimo para se produzir energia no sistema. Apesar da participação das usinas hidrelétricas estar caindo enquanto outras fontes têm se desenvolvido, ainda é a fonte de energia predominante na matriz energética brasileira, conforme ilustrado na Tabela 1.

Tabela 1 – Capacidade Instalada no SIN no final de maio de 2019

	UHE	NUCL	SOL	PCH	ÓLEO	GÁS	EOL	DIES.	CARV.	BIOM.	IMPORT.
MW	97.630	1.990	2.072	6.309	3.697	12.921	14.968	1.047	2.672	13.335	7.000
%	60%	1%	1%	4%	2%	8%	9%	1%	2%	8%	4%

Fonte: (MME/EPE, 2019)

Nota: Dados retirados da fonte e elaboração da tabela pelo autor

"Nas condições atuais do Setor Elétrico Brasileiro, o valor da água depende do nível de armazenamento dos diversos reservatórios e da probabilidade de aflúncias futuras, o que torna o custo marginal de operação altamente volátil a médio e longo prazo. Logo, o despacho deve considerar este aspecto como um fator de restrição com características intertemporais, o que exige uma operação de forma centralizada e bastante complexa."(MELO *et al.*, 2015)

Assim o cálculo do PLD se baseia em modelos matemáticos do sistema que buscam por

uma condição de equilíbrio entre utilizar o recurso, ou seja, a água, no presente ou armazená-la em reservatórios para uso futuro. A geração mais barata é através da hidrelétrica, baixo valor do MWh, porém há o risco de ter um baixo nível dos reservatórios no futuro caso a gestão do recurso não seja eficiente e haver déficit de geração em relação à demanda necessária. Outra análise seria preservar os reservatórios para uso futuro e atender a demanda presente com uma maior utilização de geração térmica, mais cara, o que resulta em um maior custo de operação. A Figura 6 ilustra essa sistemática.

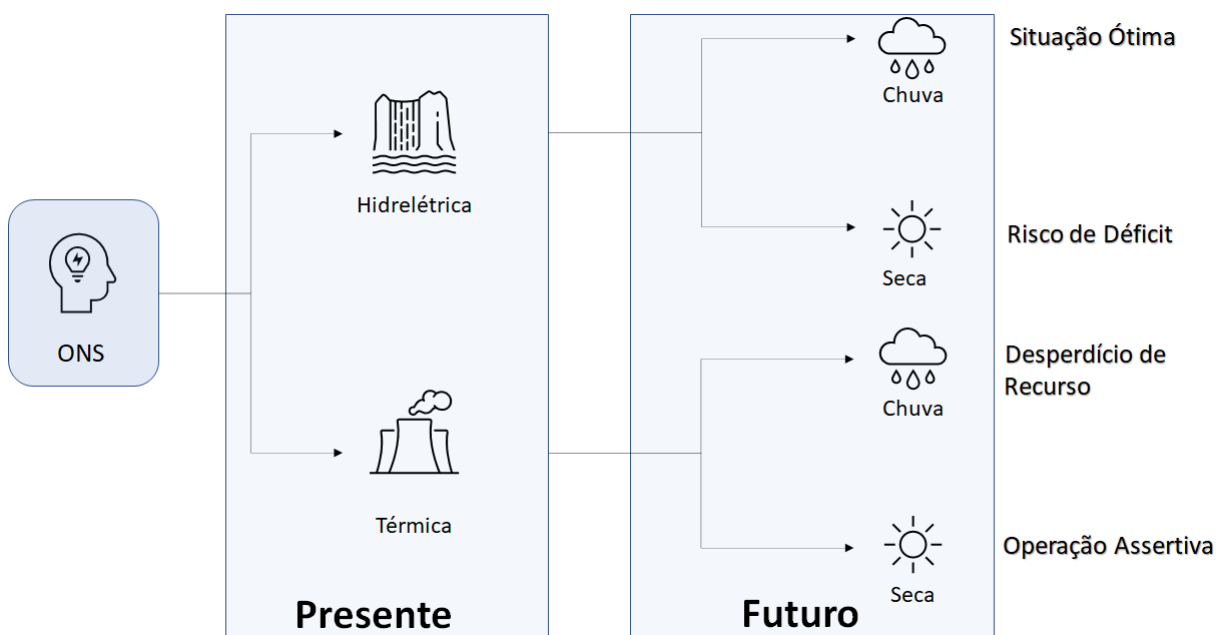


Figura 6 – Operação e Planejamento do Sistema Elétrico

Fonte: Elaborado pelo autor

Uma vez definido a importância do PLD para o mercado de curto prazo e a base para o seu cálculo vale ressaltar alguns dos que influenciam no preço do PLD.

1. Nível do Reservatório das Hidrelétricas - um período de estiagem pode resultar em uma diminuição dos níveis dos reservatórios e assim aumentar o risco de déficit de geração futura. Dessa maneira, o aumento no despacho das termelétricas irá provocar o aumento também do valor do PLD. Obviamente, o contrário também é aplicável, ou seja, níveis elevados dos reservatórios diminuem a necessidade de operação das termelétricas. No ano de 2021, o Brasil atravessou a maior crise dos últimos 91 anos
2. Aumento ou Diminuição da Demanda de Energia Elétrica – o período da pandemia do COVID-19 provocou uma diminuição muito abrupta na demanda do SIN (Sistema Interligado Nacional), o que ocasionou a queda do PLD em alguns momentos para o seu valor mínimo.

3. Previsão de Novas Fontes de Energia Elétrica - a entrada de novas fontes aumenta a oferta de energia elétrica no sistema e, conseqüentemente, tem impacto no PLD.

1.2 Metodologia

Na literatura há diversos trabalho buscando prever séries temporais, sejam dados econômicos, volume de vendas, ou quaisquer outros tipos de dados. As metodologias utilizadas são diversas, com previsão direta da série ou de forma indireta, utilizando-se variáveis correlacionadas com a que se busca prever.

Este trabalho se propõe em utilizar uma metodologia combinada, com o emprego de séries temporais e redes neurais artificiais. Após uma análise de três variáveis (Custo Marginal de Operação, Demanda Máxima de Energia e Energia Armazenada) espera-se concluir sobre qual a correlação com o PLD e, então, escolher dentre elas aquelas que deverão compor o que se convencionou chamar como primeira etapa.

Nesta primeira etapa serão realizadas as modelagens estatísticas das séries temporais das variáveis predictoras e, após chegar naquele que será o modelo mais adequado, os comparando através do Erro Quadrático Médio (sigla em inglês RMSE), será feito a previsão para os quatro passos subsequentes.

Enquanto na segunda etapa será a vez de realizar a modelagem da rede neural artificial recorrente, que será alimentada pelo histórico do PLD e pelas variáveis predictoras. Dessa maneira, após concluir sobre a rede neural com o número de camadas e neurônios que resultará em uma performance dentro do aceitável, avaliada com os dados de teste e comparando com base no RMSE, alimentá-la com as variáveis predictoras futuras previstas e produzir o preço do PLD para quatro semanas operativas subsequentes.

Resumidamente, a Figura 7 ilustra o que se espera modelar e produzir com o presente estudo.

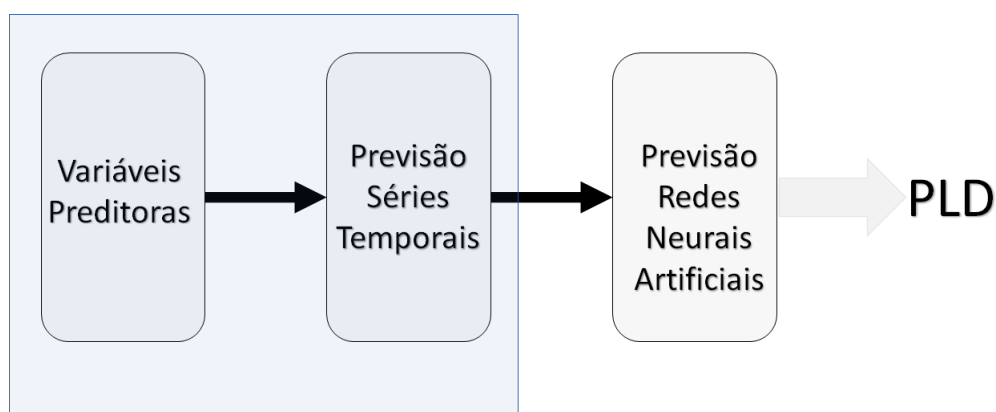


Figura 7 – Metodologia

Fonte: Elaborado pelo autor

1.3 Estrutura e Tópicos do Trabalho

Com intuito de criar uma visão geral sobre o trabalho, este item tem como objetivo descrever brevemente os próximos tópicos a serem abordados.

- Capítulo 2 - Revisão Bibliográfica
 - Verificar e apresentar os principais trabalhos relacionados com o tema de previsão de preços de commodities utilizando redes neurais, outras técnicas de machine learning e séries temporais;
 - Detalhar brevemente os últimos trabalhos e aqueles mais citados relacionados ao tema de comercialização de energia com foco no mercado livre;
- Capítulo 3 – Séries Temporais
 - Descrever os principais tópicos e a teoria base de séries temporais a ser utilizada na previsão futura das variáveis preditoras;
- Capítulo 4 – Redes Neurais Artificiais
 - Apresentação da linha temporal do desenvolvimento da teoria de redes neurais artificiais;
 - Apresentar os conceitos básicos e fundamentais para a utilização das RNA neste trabalho;
 - Introduzir conceito teórico sobre redes neurais recorrentes e Long Short-Term Memory (LSTM);
- Capítulo 5 – Desenvolvimento do Modelo Preditivo
 - Identificação das variáveis preditoras, análise e tratamento dos dados;
 - Análise exploratória das variáveis;
 - Análise e tratamento dos dados referentes ao histórico do PLD (variável target);
 - PLD e correlações
 - Implementação através da linguagem Python das técnicas tratadas nos capítulos 3 e 4;
 - Concluir sobre o modelo que apresentou o menor erro absoluto;
 - Realizar a previsão com origem no instante n para $h = n+1, n+2, n+3, n+4$ para as variáveis preditoras;
- Capítulo 6 – Apresentação dos Resultados? Análise e Avaliação do Modelo
- Capítulo 7 – Conclusão

- Principais conclusões sobre o tema;
- Sugestão de trabalhos futuros no mercado de energia ou outras áreas correlatas;

REVISÃO BIBLIOGRÁFICA

A previsão de séries temporais conta com diversos métodos estatísticos, tais como modelos SARIMA e suas particularizações por exemplo. Além desses, também há diversos trabalhos que utilizam métodos de inteligência artificial, como SVM (Support Vector Machine), redes neurais e outros modelos de regressão. Este capítulo se propõe revisitar alguns trabalhos que tratam de previsões de séries temporais com as principais técnicas utilizadas, inclusive com o emprego da combinação delas com intuito de elevar a acuracidade e tornar o modelo mais preciso.

2.1 Trabalhos Gerais com Séries Temporais

A evolução tecnológica recente tem aproximado a relação entre o mundo real e virtual com um aumento vertiginoso do fluxo de dados entre eles, seja através de equipamentos com conexão com a internet (aparelhos eletrodomésticos, carros, equipamentos industriais, etc), celulares ou qualquer outro dispositivo de IoT (Internert of Things). Toda gestão moderna, independentemente do tipo de negócio, caso não execute esse tipo de abordagem será praticamente obrigada a ter em sua área de marketing e planejamento especialistas que irão analisar dados, seja de produção, mercado, logística ou qualquer outra área que possa afetar seu negócio, para dar subsídios e embasamento aos próximos passos da empresa. Nessa linha o artigo ([ANDRADE; MIRANDA; ACHIDA, 2017](#)) reforça que a tendência de fabricação em larga escala na indústria para atender demandas futuras foi substituída por cálculos de previsões a fim de otimizar o recurso e, conseqüentemente, reduzir custos sem deixar de atender ao cliente. O trabalho ([QUEIROZ *et al.*, 2012](#)) aborda as séries temporais e a sua utilidade nos ramos industriais e objetivou-se em buscar um modelo para projeção de vendas mensais em uma dada indústria moveleira. Foram utilizados dados referentes a um período entre abril até julho de 2011 e, portanto, o modelo Holt-Winters se mostrou mais apropriado que o ARIMA, uma vez que não foi possível fazer uma análise sobre sazonalidade.

Apesar de não utilizada neste trabalho, uma metodologia que pode ser utilizada na aplicação de séries temporais é a decomposição em subséries com previsões dentro dessas, sendo que são aplicados filtros de ondaleta para particionar a série original em uma série de alta e outra de baixa frequência. Seguindo esta linha o trabalho (LIMA *et al.*, 2010) se propõe em prever preços de commodities, mais especificamente da saca de soja, utilizando a decomposição de séries temporais através de filtros de ondaleta em conjunto com modelos econométricos e redes neurais artificiais. Dessa maneira, são feitas previsões dentro das subséries e, portanto, a previsão da série original é composta por cada uma dessas partições. É demonstrado que a utilização da técnica de decomposição por ondaleta em conjunto com modelo ARIMA-GARCH resultou em uma melhora da acurácia em curto prazo e, ainda mais preciso, com redes neurais artificiais recorrentes como sendo uma excelente alternativa na predição de séries temporais do mercado financeiro.

As redes neurais têm uma grande presença em problemas relacionados ao mercado financeiro devido à sua capacidade de generalização conforme menciona (THIELE; ADAMI, 2016), onde é proposto um modelo combinado de três redes neurais para previsão da tendência (subida ou descida) do preço de ações. A primeira RNA tem como entrada os dados das variáveis dependentes do ativo, enquanto que a segunda RNA tem suas entradas sendo as variáveis exógenas e indicadores de análise técnica. Por fim, a terceira RNA utiliza os dados de saída da primeira e segunda RNA como dados de entrada. A Figura 8 ilustra a metodologia utilizada.

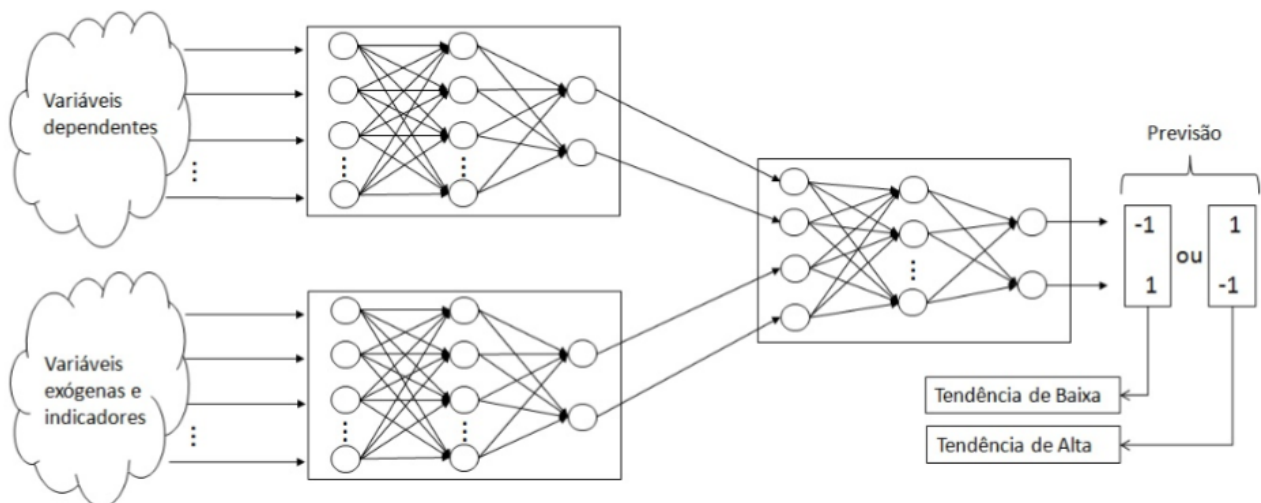


Figura 8 – Arquitetura do Sistema proposto

Fonte: (THIELE; ADAMI, 2016)

O resultado final detectou as tendências futuras em 59% das vezes e muito próximo dos resultados obtidos em outros trabalhos similares assim como citado pelos próprios autores, o que pode significar ganhos para um tomador de decisão.

Outro tema fundamental na utilização de redes neurais é a quantidade de neurônios na

camada oculta, visto que nem sempre um aumento de nós resulta em um melhor modelo de previsão, pois pode haver overfitting (ou sobreajuste) durante o treinamento da rede. Conforme tabela abaixo retirada de (FERREIRA *et al.*, 2012) é possível acompanhar a variação do RMSE (Root Mean Square Error) de acordo com a variação da quantidade de neurônios na camada intermediária em um modelo de previsão de preços das commodities (soja, boi gordo, milho, trigo).

Tabela 2 – RMSE para diferentes configurações de rede neural

Rede/ Produto	2-3-1	2-5-1	2-9-1	2-12-1
Soja	0,00359160	0,00360336	0,00346694	0,00355879
Boi gordo	0,00405874	0,00405874	0,00495829	0,00451003
Milho	0,00528765	0,00521627	0,00521928	0,00525511
Trigo	0,00237395	0,00229848	0,00230511	0,00232999
Média	0,00382799	0,00379421	0,00398741	0,00391348

Fonte: (FERREIRA *et al.*, 2012)

Nota: Dados tirados da fonte e elaboração da tabela pelo autor

Conforme mostrado há casos que cinco neurônios na camada intermediária resultou em uma melhor performance do que casos com nove, ou doze neurônios.

2.2 Trabalhos com Séries Temporais e Setor Elétrico

Em (REIS; SILVA, 2004) o tema é a previsão de carga no sistema elétrico a curto prazo sendo que a não estacionaridade das cargas é um problema para predição via redes neurais.

Conforme mencionado por (SHIRI *et al.*, 2015) a previsão de preços no mercado de energia pode ser feita de várias maneiras distintas, sendo que a maior parte das técnicas empregadas baseiam-se no histórico de preços e com o emprego de séries temporais, que são muito utilizadas em previsões de curto prazo. Entretanto, é citado que uma falha desses modelos é ignorar outros fatores que podem impactar o preço da energia. Dessa maneira, seu trabalho se propõe em aplicar Support Vector Machine (SVM) e incluir na análise outras variáveis que têm um impacto significativo no mercado considerado no estudo, tais como o histórico de preços do petróleo e do gás natural. A conclusão é que existem outras variáveis que possuem correlação com o preço da energia e podem ser incluídas na análise para aprimoramento do modelo.

O mercado livre tem como objetivo criar um ambiente em que haja concorrência no setor elétrico e, dessa maneira, criar benefícios aos consumidores industriais que podem escolher por quais empresas serão fornecedoras de sua energia, conforme cita o artigo (LEME *et al.*, 2008). Este trabalho tem um enfoque na volatilidade do preço da energia, citando por exemplo que variações climáticas têm impacto considerável devido à alta correlação entre eles por conta

da predominância de hidrelétricas na matriz energética brasileira. Assim, com o intuito de prever o preço e contribuir com o gerenciamento de risco neste mercado, o trabalho propõe uma metodologia prática utilizando modelo GARCH para estudar a volatilidade do preço em todos submercados entre os períodos de 03/2005 e 09/2007.

A previsão do PLD através do uso de redes neurais artificiais é abordado no trabalho de (RODRIGUES, 2009) para o preço de seis valores futuros dos quatros submercados (SE, NE, N, S) utilizando de forma indireta a influência de fontes geradoras hídrica, solar e eólica, além do histórico de preços e a demanda de energia. A RNA possui três camadas ocultas e foi simulada diversas vezes alterando a quantidade de neurônios em cada camada buscando assim a melhor combinação que resultaria em uma melhor performance. Considerando todos os cenários de previsão, para todos submercados e patamares de carga, o desvio padrão não excedeu R\$ 1,00, o que pode ser considerado como muito efetivo no atingimento de seu objetivo visto a natureza volátil da variável.

Ainda sobre a previsão de preços no mercado de energia elétrica, o artigo (LONGO *et al.*, 2016) faz uma comparação entre as performances utilizando previsão com redes neurais artificiais entre os modelos MLP (Multilayer Perceptron) e FCN (Fully Connected Neuron). O conjunto de dados foi dividido entre dados de entrada - indicativos do histórico de carga, condição climática, dias da semana, horários medidos e histórico das últimas 24 horas do preço do gás natural - e saída - o preço da energia (\$/MWh). A MLP utilizada consistiu em uma rede de duas camadas (entrada e saída), sem camada oculta, com a Sigmoid como função de ativação e com 30 neurônios de entrada e um de saída. Enquanto para a FCN foram 30 neurônios, sendo um de saída. Uma vez que o conjunto de dados para o treinamento das duas redes foi o mesmo, é possível comparar a eficiência de ambos os modelos e concluir que a FCN se saiu melhor na predição de valores.

Após a revisão bibliográfica conclui-se que existem vários trabalhos que se propuseram a prever o PLD ou outra variável de comportamento semelhante, como exemplo as commodities. Muitos trabalham se utilizam de séries temporais e redes neurais artificiais. Entretanto, devido a volatilidade deste tipo de variável e sua mudança de comportamento de tempos em tempos, entende-se que a busca pelo melhor modelo será contínua e irá gerar muitas alternativas que podem ser replicáveis para um determinado caso.

SÉRIES TEMPORAIS

As séries temporais são definidas de forma geral como observações realizadas em uma sequência ao longo do tempo, como exemplo a Figura 9. São diversos os exemplos e campos onde os dados estão organizados dessa forma, apenas para citar alguns: economia (taxa de desemprego, cotação de moedas e ações, histórico de preços e PIB, registro das exportações), engenharia (medidas realizadas por sensores em uma máquina, histórico da temperatura do vapor de saída de uma caldeira, registro do consumo de energia de uma fábrica), medicina (eletrocardiograma), etc. Dessa maneira, considerando todos os dados que são gerados neste padrão nos mais variados setores, é de extrema importância seu estudo e também a obtenção de técnicas para que seja possível extração de informações valiosas que resultem em conclusões efetivas.

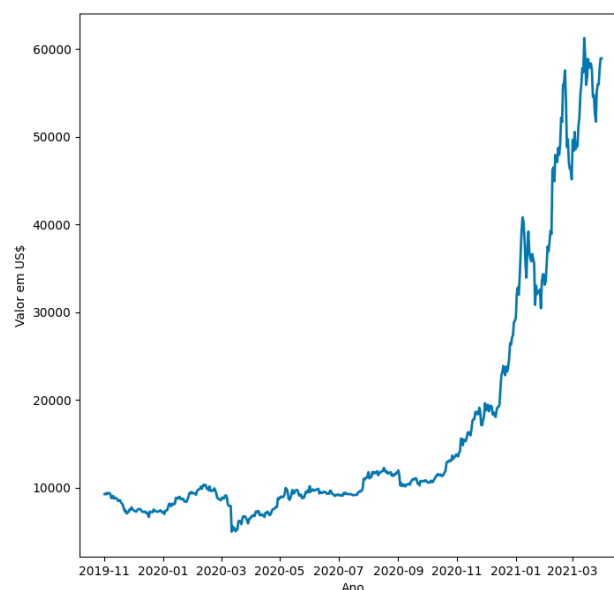


Figura 9 – Exemplo de Série Temporal

Fonte: Elaborado pelo autor

Portanto, este capítulo se propõe a revisitar alguns tópicos referentes à teoria de séries temporais para que assim seja possível seu entendimento e aplicação para a construção do modelo de previsão.

Quando se estuda uma série temporal é possível extrair informações estatísticas, verificar qual a relação com outra série e também prever valores futuros. Sendo que usualmente são encontrados padrões de comportamento que podem ser identificados através de técnicas que serão vistas mais adiante. Os principais padrões são: i) Tendência: quando é uma série decrescente ou crescente; ii) Sazonalidade: quando é observado um determinado padrão de repetição considerando uma periodicidade; iii) Ciclo: repete, porém, diferentemente da sazonalidade, não ocorre com regularidade.

3.1 Decomposição de Séries Temporais

Há um interesse em decompor uma série temporal, uma vez que é possível obter componentes que sejam mais simples de serem modeladas, além disso obter uma série sem sazonalidade e/ ou sem tendência. Uma forma bastante utilizada é a decomposição clássica que pode ser tanto da forma aditiva ou multiplicativa.

- Aditivo

Neste caso a decomposição é do tipo:

$$X = T + S + E \quad (3.1)$$

- Multiplicativo

É aconselhável que para casos onde haja uma amplitude da sazonalidade no tempo se utilize da decomposição do tipo multiplicativa, a qual possui a seguinte fórmula.

$$X = T \times S \times E \quad (3.2)$$

T - Termo referente à tendência

S - Termo referente à sazonalidade

E - Termo referente ao erro ou erro estimado

Normalmente, as séries são decompostas em suas componentes para que se tenha uma melhor visualização do seu comportamento, sendo: série original, tendência, sazonalidade e erro estimado. Conforme a Figura 10, uma série temporal gerada foi decomposta através da técnica aditiva, é possível constatar que há uma tendência de crescimento ao longo do tempo com uma

componente sazonal. Além disso, o gráfico do resíduo é a componente que não foi possível modelar na tendência e/ou sazonalidade, ou seja, uma estimativa do erro. Sendo um importante indicador sobre o sucesso ou não do processo de decomposição, pois caso seja possível notar a presença de tendência ou sazonalidade será necessário se utilizar de outras técnicas mais avançadas.

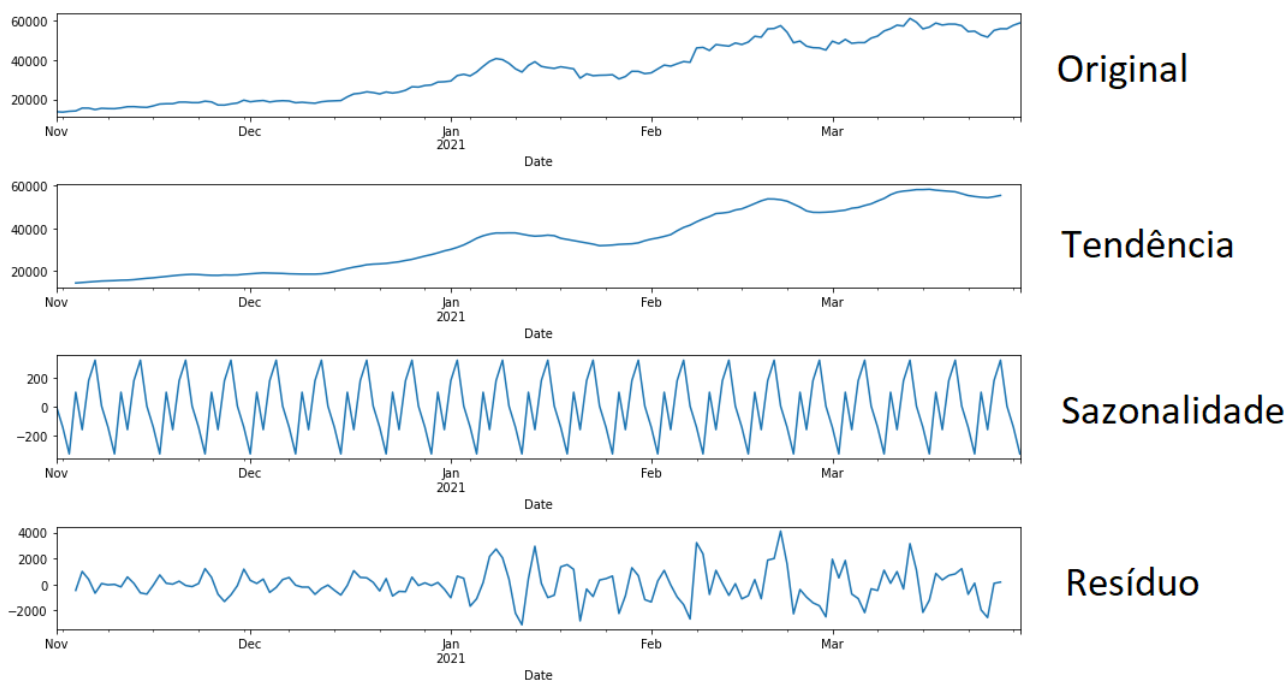


Figura 10 – Decompondo a Série Temporal

Fonte: Elaborado pelo autor

Apresenta-se na Figura 11 um histograma do resíduo da série da Figura 9.

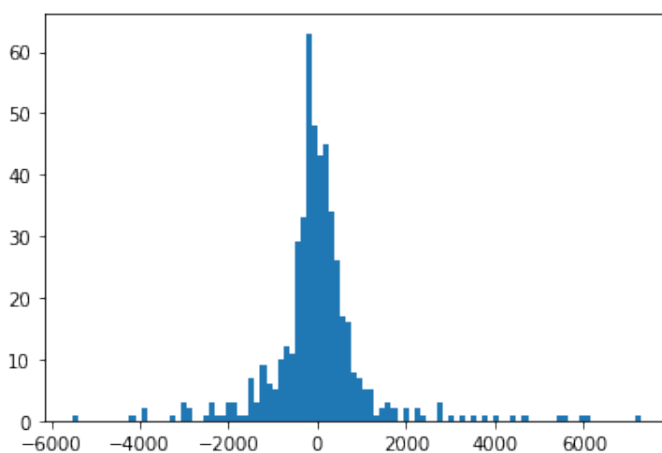


Figura 11 – Distribuição Resíduo

Fonte: Elaborado pelo autor

3.2 Autocorrelação

Em uma série temporal as observações podem possuir uma relação de dependência entre si, cujo negligenciamento poderá resultar em um modelo de previsão com viés. Para que seja possível entender qual é essa relação pode-se fazer uso da autocorrelação, o que irá dizer como uma observação em um determinado instante se relaciona linearmente com as observações passadas. Uma ferramenta utilizada para isso é a Função de Autocorrelação (ACF – Autocorrelation Function) e Autocorrelação Parcial (PACF - Partial Autocorrelation Function).

Segundo a definição estatística tem-se que autocorrelação mede quanto o valor de uma variável aleatória é capaz de influenciar seus vizinhos, a autocorrelação varia entre -1 e 1, sendo que indica a correlação negativa perfeita e correlação positiva perfeita, respectivamente. Para uma valor nulo (zero) tem-se uma ausência total de correlação. Considerando uma variável aleatória X_t discreta com média μ , tem-se a sua autocorrelação dada por:

$$\rho_X^2(k) = \frac{\text{Cov}(X_t, X_{t+k})}{\text{Var}(X_t)} \quad (3.3)$$

Na Figura 12, à esquerda, tem-se um exemplo de uma função de autocorrelação de uma série temporal. O decaimento gradual mostra que há uma correlação entre as observações, diferentemente do gráfico à direita, em que os dados são não correlacionados. A faixa azul na figura trata-se do intervalo de confiança, o que por padrão é de 95%, onde as informações que estão fora dessa faixa são dadas como correlacionadas estatisticamente.

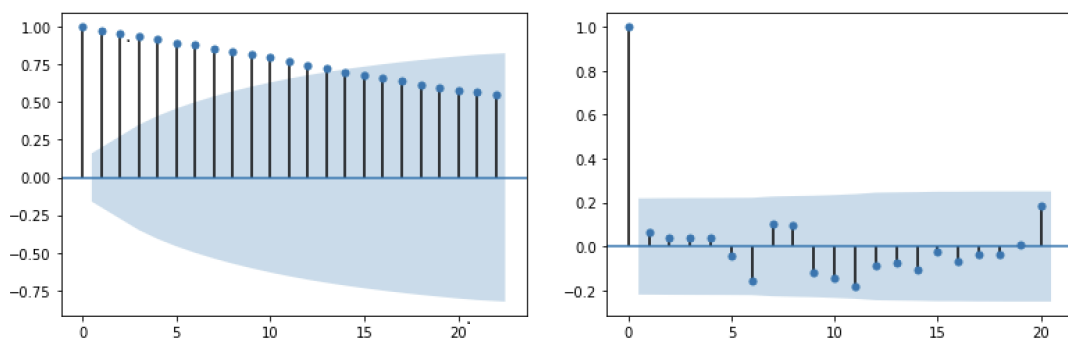


Figura 12 – Funções de Autocorreção e Autocorreção Parcial

Fonte: Elaborado pelo autor

3.3 Estacionariedade

Uma série é considerada estacionária uma vez que sua variância e média não mudem ao longo do tempo, o que implica em uma série sem tendência e/ ou sazonalidade. Dessa maneira, a função de autocorrelação vista na última seção é muito importante como ferramenta para inferir sobre a estacionariedade. Caso a série não seja estacionária as suas características estatísticas irão mudar ao longo do tempo, o que torna a tarefa de prevê-la no futuro mais difícil. Portanto, nesta seção serão mostradas algumas das maneiras de tratar a não estacionariedade.

Conforme mostra (FERREIRA *et al.*, 2017, p. 129) normalmente uma série temporal é não estacionária devido à presença de uma tendência determinística ou estocástica. Sendo que uma tendência determinística pode ser descrita da seguinte maneira:

$$x_t = \alpha + \beta t + \varepsilon \quad (3.4)$$

O termo βt é responsável por criar uma tendência determinística, o que resultará na não estacionariedade. Portanto, para torná-lo estacionário deve-se eliminar a parte βt .

Enquanto que para um processo aleatório discreto pode-se ter um passeio aleatório, o que é definido como:

$$x_t = x_{t-1} + \varepsilon_t \quad (3.5)$$

O que ao longo do tempo tem-se:

$$x_t = x_0 + \sum_{j=1}^t \varepsilon_j \quad (3.6)$$

Assim, a variância e a média serão dependentes de t e, portanto, a série temporal não será estacionária. Ou seja, tanto uma tendência determinística quanto estocástica resulta em processos similares.

3.3.1 Teste de Dickey-Fuller Aumentado (ADF)

Conforme mostra (FERREIRA *et al.*, 2017, p. 140), uma das formas mais utilizadas para avaliar a estacionariedade de uma série temporal é através do teste de raiz unitária, sendo que para isso há diversos tipos de testes possíveis. Nesta seção será abordado o teste de Dickey-Fuller Aumentado devido à sua aplicação ser mais usual, apesar de haverem outros testes que também podem ser utilizados para o mesmo propósito.

Considerando uma série temporal x_t tem-se que, caso a raiz da sua equação característica seja unitária, o processo é não estacionário. Enquanto que, para casos onde a raiz está dentro do círculo unitário, o processo é estacionário.

Dessa maneira, considerando a série temporal a seguir:

$$x_t = \phi x_{t-1} + \varepsilon_t \quad (3.7)$$

os testes de hipótese para a raiz unitária são descritos da seguintes maneira:

$$H_0 : |\phi| = 1 \rightarrow x_t, \text{ não estacionário} \quad (3.8)$$

$$H_1 : |\phi| < 1 \rightarrow x_t, \text{ estacionário} \quad (3.9)$$

3.3.2 Diferenciação: Removendo a Tendência

Uma das técnicas utilizadas para remover a tendência da série e estabilizar a variância, com o objetivo de torná-la estacionária, é a diferenciação. Basicamente, a técnica consiste na diferença do período T com o valor anterior (T-1). Sendo que pode ser aplicada mais de uma vez para buscar a estacionariedade.

Buscando uma série temporal qualquer e aplicando diferenciação, com auxílio da função `.diff()` do Python, tem-se como resultado um conjunto de dados conforme a Figura 13.

Em uma primeira análise já possível notar que os valores estão variando em torno de um valor médio e sem tendência, porém com a variância perdendo estabilidade próximo a 2021. Entretanto, para que se aplique um critério técnico na análise e possa concluir, novamente se utilizará do teste de Dickey-Fuller para verificar se estatisticamente é possível considerar a série como estacionária. O resultado de p é 0,000023 (ou 0,0023%), ou seja, menor que 5% (intervalo de confiança) e mostrando que é possível rejeitar a hipótese nula. Dessa maneira, fica evidente que apenas uma diferenciação foi suficiente e que a nova série transformada é estacionária.

3.3.3 Transformação Logarítmica

Conforme cita (MORETTIN; TOLOI, 2006, p.8), é comum séries temporais da área de finanças e econométrica a existência de tendências cuja variância da série (ou suas diferenças) podem aumentar durante o tempo. Para estes casos é aconselhado utilizar a transformação logarítmica para estabilizar a variância e tornar os dados mais próximos de uma distribuição normal.

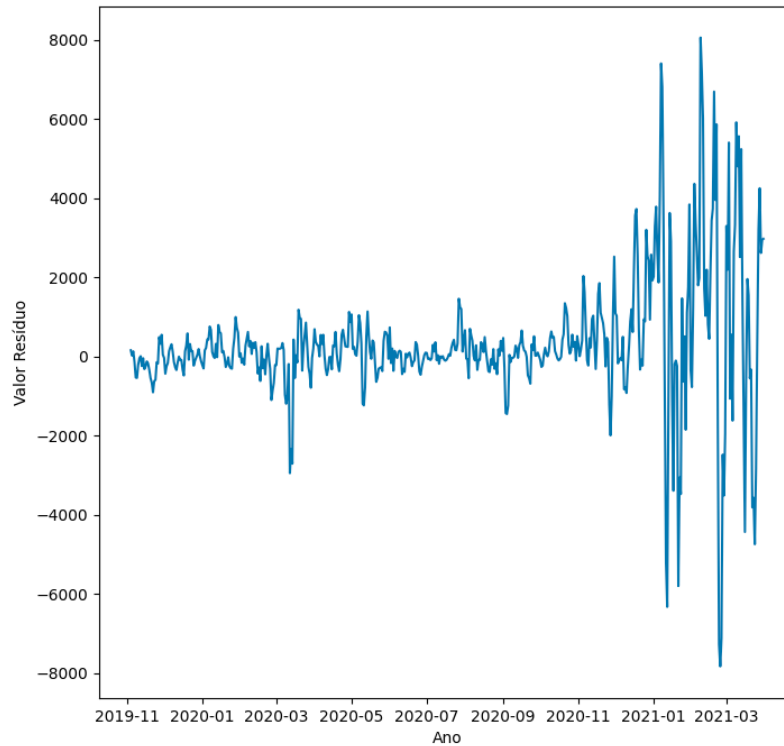


Figura 13 – Série de Dados Após Aplicar Diferenciação

Fonte: Elaborado pelo autor

De forma mais ampla é possível tratar esse tipo de dados através de uma transformação Box-Cox, a qual pode ser descrita de forma generalizada pelas igualdades abaixo, cuja transformação logarítmica é apenas um caso particular.

$$Z_t^\lambda = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & \text{se } \lambda \neq 0, \\ \log Z_t & \text{se } \lambda = 0, \end{cases} \quad (3.10)$$

A transformação de Box-Cox reúne um conjunto de funções de transformação exponencial, onde encontra o melhor valor λ com o objetivo de estabilizar a variância.

3.4 Modelos de Previsão

Mais uma vez recorrendo a (MORETTIN; TOLOI, 2006, p.83), a maior parte dos modelos de previsão de séries temporais utiliza-se uma combinação linear de observações anteriores para explicar o caso presente.

3.4.1 Médias Móveis Simples

Essa técnica é bastante simples e consiste em calcular a média aritmética das n observações mais recentes, sendo n definido para cada caso. A previsão subsequente será a última média

móvel calculada. A fórmula 3.11 mostra como realizar o cálculo da média móvel.

$$M_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-n+1}}{n} \quad (3.11)$$

Na Figura 14 tem-se um exemplo de um ajustamento utilizando-se média móvel, sendo que a linha vermelha é o valor previsto e o azul é o real.

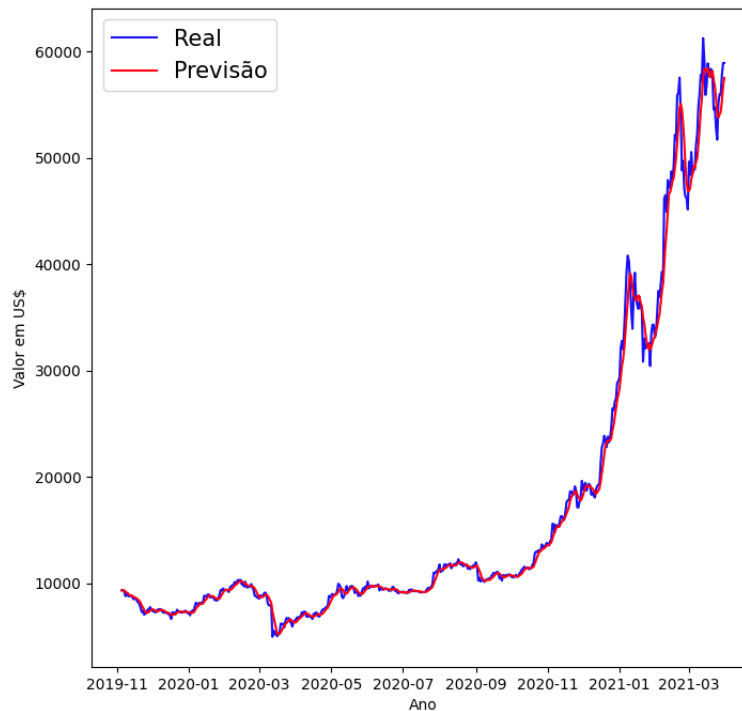


Figura 14 – MMS - Previsto x Real

Fonte: Elaborado pelo autor

Entretanto, a média móvel tem como uma desvantagem o fato de atribuir o mesmo peso para todos os valores anteriores. Ou seja, caso a janela escolhida seja de 20 valores, ou outro qualquer, todos terão a mesma participação no cálculo do valor futuro, o que normalmente não é o correto, pois dependendo da série os valores mais atuais têm uma influência maior na previsão.

3.4.2 Modelo ARIMA

Existem séries que são naturalmente estacionárias, entretanto, este tipo de série é muito difícil de ser observada sendo que a grande maioria são séries de natureza não estacionárias. Para essas últimas é muito comum se utilizar de modelos estatísticos como ARIMA (Auto-Regressive Integrated Moving Averages). Nesta seção será tratado em mais detalhes o modelo ARIMA, cuja aplicação para séries temporais trata-se basicamente de uma equação linear que depende dos parâmetros p , d e q do modelo.

- Modelo autoregressivo de ordem p: AR(p)

Da mesma forma que uma regressão linear, o modelo AR se utiliza de valores defasados para que seja possível prever valores futuros, sendo que a variável p apresentada anteriormente determina quantos valores anteriores serão utilizados na predição. Assim, quanto maior o valor de p mais os valores passados são levados em conta para prever os dados futuros.

Na equação 3.12, tem-se um forma geral para representar o modelo, onde c é uma constante e ε_t é o ruído branco, isto é, variáveis não correlacionadas de média constante e variância constante.

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t \quad i = 1, 2, \dots, p \quad (3.12)$$

- Modelo de média móvel de ordem q: MA(q)

O modelo de média móvel utiliza uma combinação linear de ruídos brancos. Assim, o valor q do modelo MA indicará quantos valores de ruídos brancos anteriores serão utilizados. Abaixo, tem-se a fórmula onde X_t é como uma média ponderada dos ruídos passados.

$$X_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \quad (3.13)$$

- Número de Diferenças (d)

Por fim, a componente d indica a ordem do grau de diferenciação, ou seja, o número de vezes em que os dados tiveram valores passados subtraídos.

Analisando-se a série temporal da Figura 14 e ajustando um modelo ARIMA com os parâmetros $p = 1$, $q = 1$ e $d = 1$, obteve-se a curva da Figura 15. Da mesma forma que feito anteriormente com o ajustamento por meio das médias móveis, é possível verificar quando comparado com os dados reais. É muito provável que alterando os parâmetros seja possível melhorar a precisão do modelo e medir a performance através do RMSE, inclusive criar um modelo de machine learning para otimização do modelo visando o menor erro possível. Como este capítulo não tem o objetivo de realizar o ajustamento, mas mostrar as bases teórica sobre o tema, o trabalho de ajuste para procurar o melhor modelo será feito no capítulo 5 com a série temporal em estudo.

Uma forma de checar a eficácia do modelo é através do cálculo da função de autocorrelação residual. Na Figura 16 tem-se um histograma dos resíduos, onde a maior parte da distribuição concentra-se próximo ao zero.

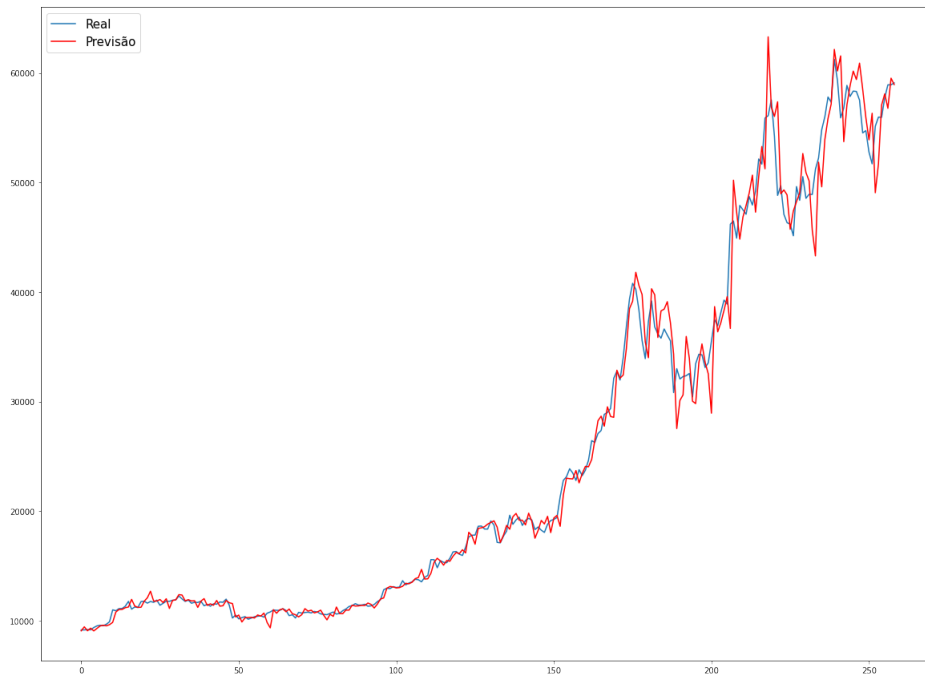


Figura 15 – Previsão Série Temporal Utilizando ARIMA

Fonte: Elaborado pelo autor

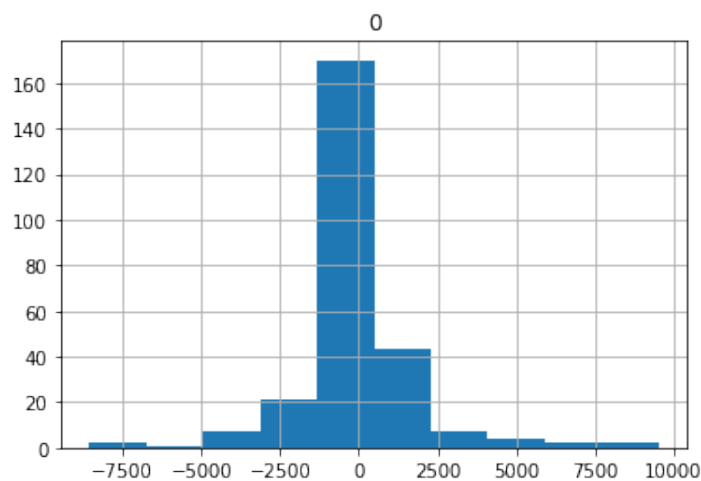


Figura 16 – Função de Autocorreção Residual

Fonte: Elaborado pelo autor

3.4.3 Modelo SARIMA

Por fim, para que seja possível completar toda a teoria necessária para a previsão das variáveis preditoras, devido à característica sazonal presente em cada uma delas, é importante que seja abordado o modelo $SARIMA(p,d,q)(P,D,Q)_s$.

Conforme citado por (SAMOHYL, 2013), a sazonalidade está presente em diversas séries temporais e, para que seja possível uma modelagem que busque melhores níveis de performance,

Box e Jenkins (1976) implementaram uma generalização do modelo ARIMA incluindo a componente sazonal, o que resultou no SARIMA (Autoregressive Integrated Moving Average), que adiciona três parâmetros adicionais para modelar a componente sazonal, sendo representado por: $ARIMA(p,d,q)(P,D,Q)$. Os três parâmetros adicionais modelam a parte sazonal da série através da ordem da componente autorregressivo, a diferenciação e a média móvel, assim como é feito para a parte não sazonal.

REDES NEURAIIS ARTIFICIAIS

Conforme mencionado em (GÉRON, 2019), muitas invenções e descobertas ao longo da história foram baseadas e inspiradas na própria natureza. Com as redes neurais artificiais (RNA) não foi diferente, seu conceito básico leva em conta o funcionamento de células do cérebro humano e busca, através de algoritmos, criar uma lógica para que seja possível reconhecer padrões em conjuntos de dados previamente submetidos à RNA e depois identificá-los ao expor novos dados à rede.

Hoje, as RNA's estão presentes nas mais diferentes aplicações diárias, como por exemplo: aplicativos de reconhecimento de imagem (ex.: ao destravar o celular utilizando reconhecimento facial), reconhecimento de voz, previsão de séries temporais, a lista de filmes sugeridos pela Netflix ou YouTube, escrita manual em tablet (ex.: que converte para texto), aplicações na agricultura (ex.: reconhecimento de uma determinada praga através da cor da folhagem em uma cultura específica), prevenção de fraude de cartão de crédito baseado no padrão de compra de um cliente, entre muitas outras.

Para cada uma das aplicações citadas anteriormente existe uma RNA mais apropriada a ser utilizada, assim como comenta o autor (FELTRIN, 2020), citando apenas algumas das mais comumente utilizadas tem-se: Redes Neurais Artificiais, Profundas, Convolucionais, Recorrentes, Auto-Organizáveis, Boltzman- Machines, Algoritmo Genético, entre outras.

O processo de previsão utilizando RNA's, assim como outros algoritmos de previsão, têm em uma de suas fases a de modelagem, o que basicamente é onde a rede aprende os padrões dos dados. E, utilizando uma parábola para explicar o processo de aprendizagem, o autor (HAYKIN, 2009) criou um paralelo com a vida real onde existe o aprendizado sob a tutela, ou não, de um professor. No mundo da inteligência artificial o processo é semelhante e ocorre através da aprendizagem supervisionada ou não supervisionada. Sendo que a primeira expõe a RNA a um conjunto de dados e busca ajustar os parâmetros do sistema almejando uma resposta conhecida, assim permanece em um processo que se repete até que o erro esteja dentro da faixa aceitável.

Enquanto que para a segunda não existem dados rotulados e a rede busca por padrões por conta própria.

O conceito de RNA é relativamente antigo e remonta há mais de 50 anos atrás, entretanto a sua aplicação só veio a ganhar maior utilidade na década de 90 após o desenvolvimento computacional, principalmente das GPU's (Graphic Processing Unit) que nasceram inicialmente focadas para o mercado de jogos e processamento de imagem, porém, devido ao alto poder de processamento, permitiram que grandes bancos de dados fossem usados para treinar as RNA's.

Nas próximas seções deste capítulo serão abordados temas com o objetivo de revisitar a teoria básica das RNA's, entender sua concepção, explorar seu modelo matemático e também lançar as bases para que seja possível a aplicação na previsão dos preços através do modelo mais apropriado para essa finalidade. O campo de RNA's é muito vasto e se estende por diversas aplicações, assim, com o objetivo de não alongar muito esta seção, as bases teóricas estarão restritas ao conteúdo apropriado a este trabalho.

4.1 O Neurônio Biológico

Conforme mencionado anteriormente, as RNA's tiveram como base o funcionamento do neurônio biológico, portanto é importante entender como este sistema funciona antes de aprofundar nos conceitos matemáticos.

Neurônios são células compostas basicamente por três partes: 1) corpo celular, cuja principal função é receber e integrar os sinais, 2) dendritos - que recebem os estímulos e levam até o corpo celular, 3) axônios, que conduzem o impulso de um neurônio para outro. Na Figura 17, retirada de (GÉRON, 2019), tem-se a representação simplificada do neurônio biológico.

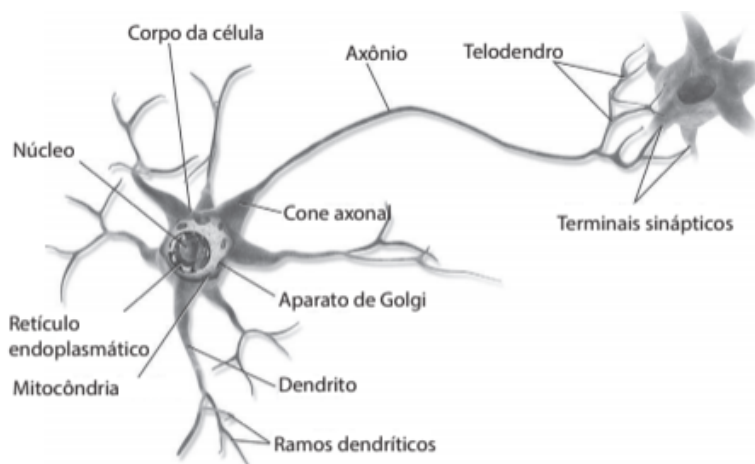


Figura 17 – Neurônio Biológico

Fonte: (GÉRON, 2019)

Outro ponto importante, e que tem bastante analogia na aplicação da RNA, é que nem

todos os sinais são processados, pois dependendo do estímulo excitatório o sinal pode ser bloqueado.

Obviamente que o aprofundamento no funcionamento detalhado destas células, que são a base para o sistema nervoso, exigiria um desvio do foco deste trabalho. Portanto, o que mais importa nesta seção é entender a ideia principal que é: os neurônios são responsáveis pela transmissão de impulsos elétricos e que toda a comunicação se dá através de milhões de elementos deste tipo que produzem uma resposta final. Assim todo o sistema depende das informações processadas em cada uma dessas células individualmente que, ao se comunicarem, tem-se uma resposta final.

4.2 Primeiro Modelo de Neurônio Artificial

Conforme (HAYKIN, 2009), foi em 1943 que o primeiro modelo proposto para um neurônio artificial foi amplamente aceito, ou seja, um modelo matemático que se baseava nos princípios básicos do neurônio biológico, foi criado por Warren S. McCulloch e Walter H. Pitts. A ideia base é que um neurônio é um sistema binário, onde as entradas são binárias e combinadas por uma soma ponderada resultando em uma saída única e também binária.

Além disso, o modelo também tem algumas outras premissas: i) deve-se ter um número mínimo de sinapses excitadas para que o neurônio dispare a informação, ii) o único atraso significativo é o atraso sináptico, iii) uma atividade sinapse inibitória impede o disparo do neurônio no instante associado, iv) a estrutura não se altera ao longo do tempo.

4.3 Perceptron

Posteriormente, nos anos 50, um novo modelo baseado naquele proposto anteriormente mas com alguns ajustes foi proposto por Frank Rosenblatt e chamado de Perceptron. O Perceptron é a forma mais simples de uma rede neural usada para classificação de padrões ditos linearmente separáveis (HAYKIN, 2009).

A ideia básica do Perceptron consiste em que entradas (x_i) são multiplicados por pesos (w_i) e uma unidade de processamento recebe esses sinais realizando a somatória e adicionando um elemento denominado *Bias* (b), que é uma constante com a finalidade de melhorar o ajuste do modelo aos dados que entram na rede.

Então, é produzido uma resposta (v) que é comparada com um determinado limiar através da função de ativação (φ) pré estabelecida, de tal forma que quando exceder esse limiar irá produzir uma saída (y_i) (*output*), ou seja ativando o modelo, conforme equação 4.1.

$$output = \begin{cases} 0 & \text{se } \sum_i w_i x_i \leq threshold \\ 1 & \text{se } \sum_i w_i x_i > threshold \end{cases} \quad (4.1)$$

A Figura 18 sintetiza o modelo do perceptron.

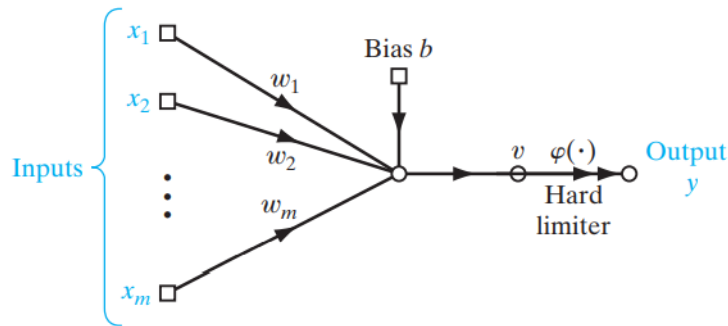


Figura 18 – Perceptron

Fonte: (HAYKIN, 2009)

O ajuste dos pesos sinápticos e do Bias é feito através de um algoritmo que, conforme mostra o teorema da convergência do Perceptron, irá sempre convergir em um número finito de iterações quando houver duas variáveis linearmente separáveis, sendo que a separação será através de um hiperplano conforme descrito na equação 4.2 e graficamente mostrado na Figura 19.

$$v = \sum_m^{i=1} w_i x_i + b \quad (4.2)$$

O treinamento do Perceptron segue um algoritmo que pode ser resumido pelos seguintes passos: inicia-se os pesos com valores randômicos, escolhe uma taxa de aprendizado η entre 0 e 1, aplica uma entrada determinada e compara a saída s_i com valor esperado t_i , dessa maneira calcula-se o erro encontrado. Caso o erro seja igual a zero não há necessidade de ajuste nos pesos, entretanto se o erro for diferente de zero os pesos são atualizados conforme a equação 4.4, que é somado com o peso anterior. Caso seja escolhido um valor η muito próximo a 1 a rede tende a convergir mais rapidamente, entretanto pode-se afastar do ponto ótimo, enquanto que um η próximo de zero tende a demorar mais para convergir, contudo pode aumentar a precisão. Normalmente, escolhe como um ponto de partida η igual a 0,5.

$$E_i = t_i - s_i \quad (4.3)$$

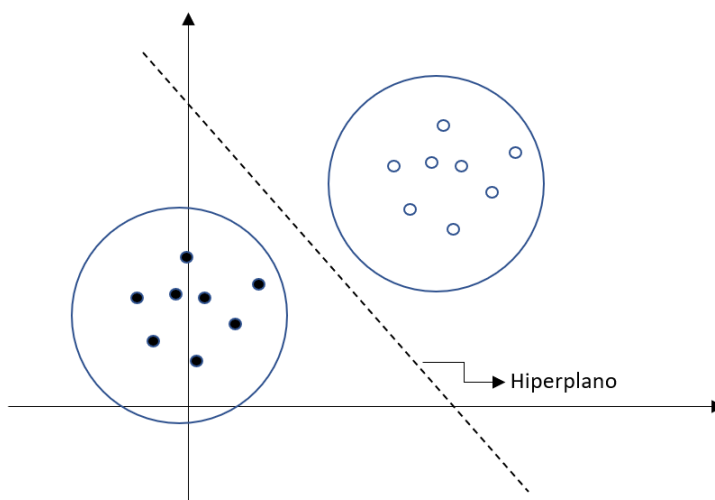


Figura 19 – Exemplo de classificação com Perceptron

Fonte: Elaborado pelo Autor

$$\Delta W_{ij} = \eta X_i E_j \quad (4.4)$$

Para o erro não se espera que em todos os casos seja possível alcançar um valor igual a zero, assim é comum definir um valor diferente de zero, porém que seja aceitável.

4.4 O Perceptron Multi Camadas

Conforme já mencionado, o Perceptron possui uma saída linear, sendo incapaz de classificar problemas que tenham natureza não linear. Nessa linha um caso que ficou muito famoso na história do desenvolvimento das RNA's foi o problema de classificação XOR (Exclusive OR). A não possibilidade de resolver tal problema criou um desapontamento na comunidade científica da época. Entretanto, posteriormente chegou-se a conclusão que seria possível resolver tais problemas se fosse utilizado uma aplicação combinada com vários perceptrons. Assim, conforme mostra graficamente na Figura 20, é possível classificar o problema do XOR utilizando dois hiperplanos.

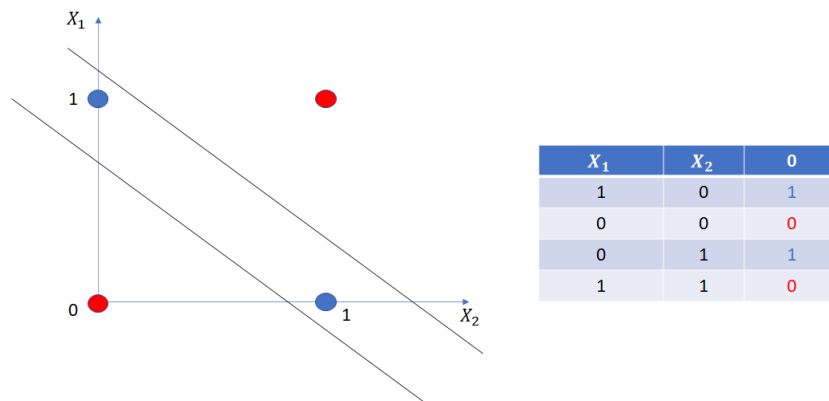


Figura 20 – Representação Gráfica Classificação XOR

Fonte: Elaborado pelo Autor

Assim, expandindo para outros problemas semelhantes, inclusive com a necessidade maior de hiperplanos para resolvê-los, criou-se o Perceptron Multicamada (sigla em inglês, MLP), cuja definição dada por (GÉRON, 2019) é que são compostos de uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída/ final. E, ainda, quando a MLP possui mais de uma camada oculta é chamada de RNA Profunda.

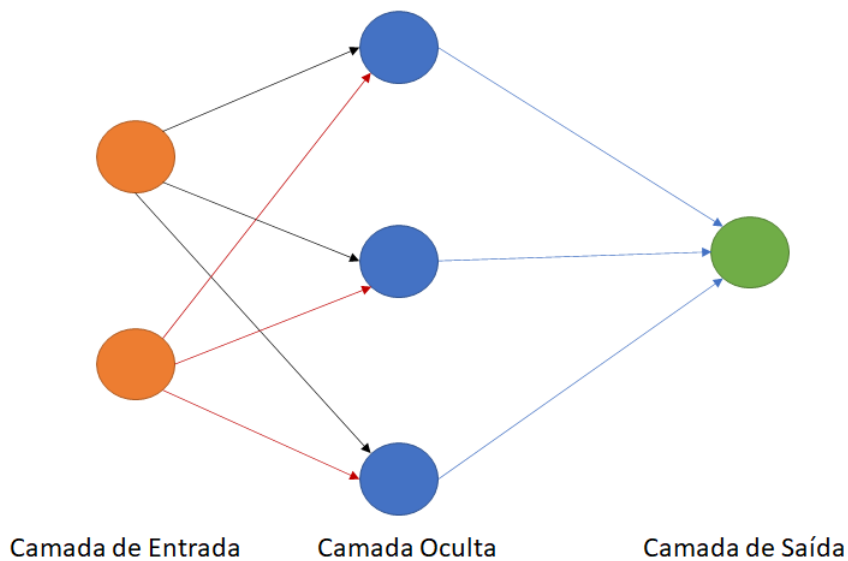


Figura 21 – MLP - Representação Gráfica

Fonte: Elaborado pelo Autor

4.5 Algoritmo de Retropropagação

O algoritmo de retropropagação (em inglês, *backpropagation*) é um método de treinamento supervisionado de redes neurais artificiais. O objetivo de retropropagação é modificar os

pesos de modo a treinar a rede neural para mapear corretamente entradas arbitrárias para suas saídas (AMRUTHA; AJAI, 2018).

Os passos do algoritmo para o treinamento podem ser definidos da seguinte forma:

1. Iniciar os pesos de forma aleatória;
2. Apresentar um padrão à rede e seguir com o processamento direto na direção da entrada para a saída (denominado como *forward pass*);
3. Calcular todos os valores nas camadas ocultas e também o resultado na saída da rede. Assim, calcular o erro obtido ao comparar com a resposta esperada para determinada entrada;
4. Fazer o cálculo reverso da saída para entrada calculando os novos pesos para cada neurônio;
5. Repetir o segundo passo até que o critério estabelecido seja atendido;

Conforme já definido anteriormente, o erro dado pela equação 4.3, é possível expandir o conceito para o erro quadrático que é dado por:

$$E_i = \frac{1}{2}(t_i - s_i)^2 \quad (4.5)$$

Considerando que a intenção é buscar a minimização do erro total da rede, é preciso somar o erro em cada neurônio da MLP, portanto:

$$E_i = \frac{1}{2} \sum_{i=1}^N (t_i - s_i)^2 \quad (4.6)$$

Em (SILVA; SPATTI; FLAUZINO, 2016) é demonstrado que ao aplicar o método do gradiente descendente para minimizar o erro total tem-se o ajuste dos pesos da seguinte forma:

$$\Delta W_{ij} = -\eta \frac{\partial E_t}{\partial W_{ij}} = -\eta \frac{\partial E_t}{\partial Y_i} \frac{\partial Y_i}{\partial W_{ij}} \quad (4.7)$$

Calculando a derivada chega-se à equação 4.8 que define como o ajuste dos pesos em relação à taxa de aprendizado, erro no neurônio e input de entrada:

$$\Delta W_{ij} = \eta E_i X_i \quad (4.8)$$

Sendo,

i = representa i-ésimo neurônio de uma determinada camada (L);

j = representa j-ésimo neurônio de uma determinada camada anterior (L-1);

4.6 Redes Neurais Recorrentes

Até o momento o conceito de redes neurais explanado tratava de entradas e saídas independentes. Entretanto, quando se estuda séries temporais, como as que são objeto principal deste trabalho, existe uma relação de interdependência e, portanto, é necessário introduzir um modelo de redes neurais que seja capaz de mapear os dados captando essas informações prévias. Isso é possível através das redes neurais recorrentes (sigla em inglês, RNN), que ao contrário da RNA, uma unidade da RNN recebe como entrada seu estado oculto anterior e a entrada atual, e produz um novo estado oculto (HUA *et al.*, 2019).

Dessa maneira, as RNN's utilizam laços de loop para que a informação de um momento anterior seja considerado no seguinte e também mantendo de alguma forma a informação de todos aqueles dados que precederam, como se fosse uma memória. A Figura 22 mostra essa dinâmica de forma simplificada com a rede sendo desdobrada em uma rede completa.

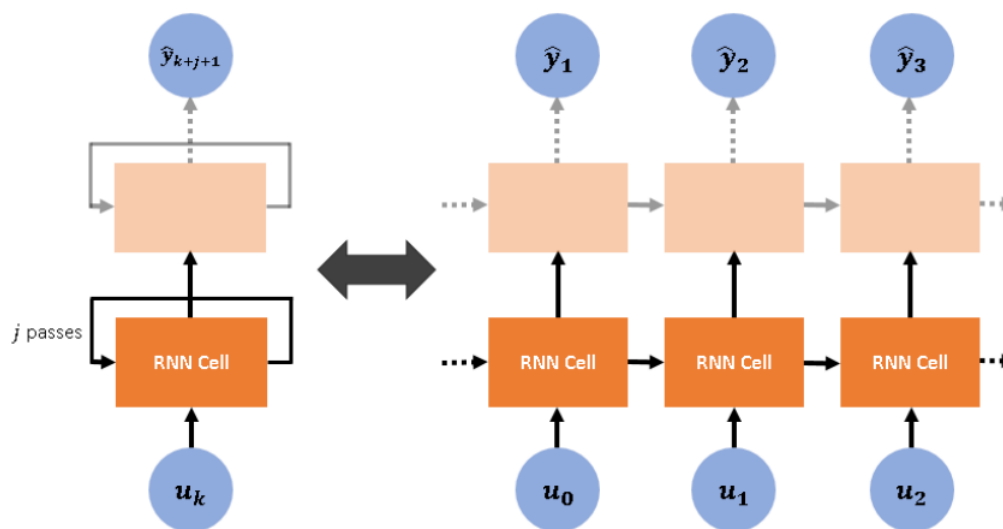


Figura 22 – MLP - Representação Gráfica

Fonte: (WONG *et al.*, 2018)

4.7 Redes de Memória Longa de Curto Prazo

As Redes Neurais Recorrentes (sigla em inglês, RNN) são eficientes em previsão de dados sequenciais, como exemplo, um texto ou uma série numérica, entretanto quando a lacuna a ser prevista cresce e são necessários dados de longo prazo as RNN's tem baixa performance. Assim, em casos como o descrito são empregados um outro modelo de rede, que basicamente é um caso especial de RNN's com alguns recursos adicionais, são as *Long Short Term Memory* (LSTM).

As LSTM's possuem a mesma estrutura em cadeia das RNN's sendo que a principal diferença é a estrutura de repetição mostrada na Figura 23.

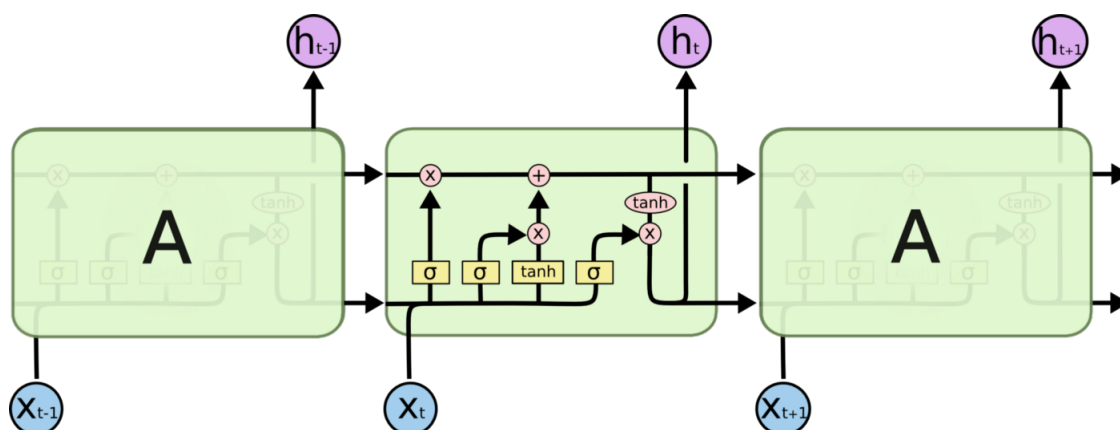


Figura 23 – MLP - Representação Gráfica

Fonte: (JUNIOR, 2019)

A linha horizontal superior da Figura 23 irá receber iterações lineares que serão controladas por filtros. Basicamente é possível definir a função de cada portão da seguinte forma:

1. *Forget Gate*: Responsável por determinar quais informações serão descartadas, sendo que o controle é realizado pela função sigmóide;
2. *Input Gate*: Seleciona quais informações de longo prazo devem ser adicionadas através de um filtro combinado com a função sigmoide e tangente hiperbólica;
3. *Output Gate*: Filtra qual parte do estado de longo prazo será produzida;

Adicionalmente, é importante mencionar que o treinamento das RNN's utiliza-se do mesmo conceito básico de retropropagação apresentando anteriormente, entretanto, devido à memória o conceito é estendido para o que se define como *backpropagation through time*, cuja a ideia é definir o tempo por uma série ordenada.

DESENVOLVIMENTO DO MODELO DE PREVISÃO DAS VARIÁVEIS PREDITORAS

Neste capítulo será desenvolvida toda a modelagem para aplicação das técnicas de séries temporais estudadas e uma análise exploratória das variáveis mostrando um pouco sobre a estatística descritiva de cada uma, seu comportamento nos últimos anos e qual a correlação entre as variáveis. Finalmente, a série será particionada em dados para ajustamento e dados para teste, para verificar a eficácia do ajustamento realizando uma previsão e comparando com os valores conhecidos.

5.1 Analisando os Dados

5.1.1 PLD (R\$/MWh - Histórico de Preços de Energia)

Através do portal ([CCEE...](#)), acessado no dia 21 de novembro de 2022, foi obtido o histórico de preços semanais do PLD (R\$/MWh), conforme Figura 24. Os dados selecionados apresentam os preços com sazonalidade de sete dias. Além disso, como o sistema elétrico é dividido em quatro submercados (SE, NE, NO, S), que representam as regiões do Brasil, constam quatro PLD's, sendo um para cada submercado. Entretanto, com o objetivo de simplificar o trabalho e com foco na teoria, foi escolhido somente o submercado (SE) para o desenvolvimento da modelagem.

Na Figura 25, para exemplificar a base de dados, tem-se o gráfico do PLD entre 08/01/2010 e 01/11/2020. Como uma primeira análise exploratória para que seja possível entender a natureza e comportamento da variável target, o PLD, as estatísticas encontram-se na Tabela 3.

ANO	MES	SEMANA	DATA_INICIO	DATA_FIM	SUDESTE (R\$/MWh)	SUL (R\$/MWh)	NORDESTE (R\$/MWh)	NORTE (R\$/MWh)
2019	11	966	23/11/2019	29/11/2019	317,33	317,33	317,33	317,33
2019	11	967	30/11/2019	06/12/2019	234,99	234,99	234,99	234,99
2019	12	968	07/12/2019	13/12/2019	217,57	217,57	217,57	217,57
2019	12	969	14/12/2019	20/12/2019	226,59	226,59	226,59	226,59
2019	12	970	21/12/2019	27/12/2019	195,13	195,13	195,13	195,13
2019	12	971	28/12/2019	03/01/2020	290,33	290,33	290,33	290,33
2020	1	972	04/01/2020	10/01/2020	368,51	368,51	368,51	368,51
2020	1	973	11/01/2020	17/01/2020	268,32	268,32	268,32	268,32
2020	1	974	18/01/2020	24/01/2020	355,42	355,42	355,42	355,42
2020	1	975	25/01/2020	31/01/2020	333	333	332,28	332,28
2020	2	976	01/02/2020	07/02/2020	190,96	191,01	190,96	190,96
2020	2	977	08/02/2020	14/02/2020	150,67	174,24	150,67	149,89
2020	2	978	15/02/2020	21/02/2020	146,13	171,14	144,93	132,63
2020	2	979	22/02/2020	28/02/2020	143,85	180,77	125,8	39,68
2020	2	980	29/02/2020	06/03/2020	58,6	66,78	58,6	57,34
2020	3	981	07/03/2020	13/03/2020	86,62	158,2	84,67	39,68
2020	3	982	14/03/2020	20/03/2020	100,86	224,12	98,97	39,68
2020	3	983	21/03/2020	27/03/2020	101,98	201,83	39,68	39,68

Figura 24 – Exemplo do Formato da Base de Dados PLD

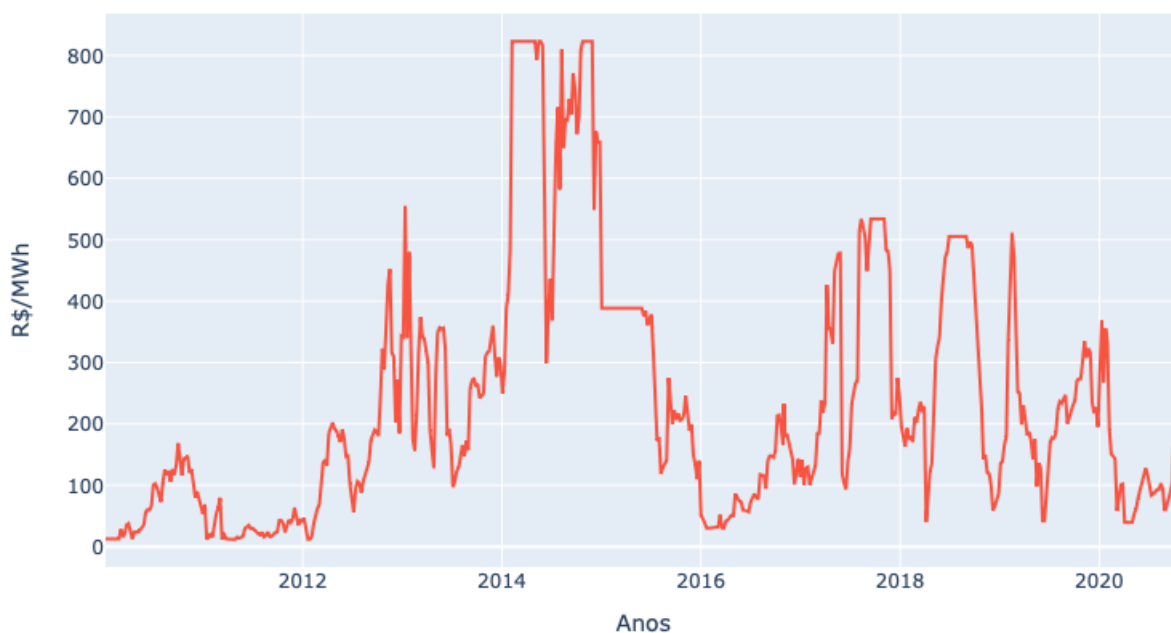


Figura 25 – Histórico PLD

Tabela 3 – Análise descritiva PLD

	PLD SE (BRL/MWh)
Tamanho Amostral	567
Média	235,24
Des. P	204,76
Min.	12,08
25%	82,91
50%	177,28
75%	337,99
95%	711,62
Max	822,83

A Tabela 3 mostra que 95% dos valores observados de PLD estão concentrados abaixo de R\$ 711,62 e através da Figura 26 nota-se uma frequência menor para valores acima de R\$600/MWh. Dessa maneira, entende-se que casos acima desta faixa de preço são menos comuns, ou ainda, podendo refletir situações que estão fora de um contexto mais esperado. Para exemplificar na prática essa constatação, na Figura 27, no ano de 2014, o valor máximo foi acima de R\$ 800,00, o que é explicado por uma grave falta de chuvas naquele período elevando o valor a máxima recorde já registrado até então.

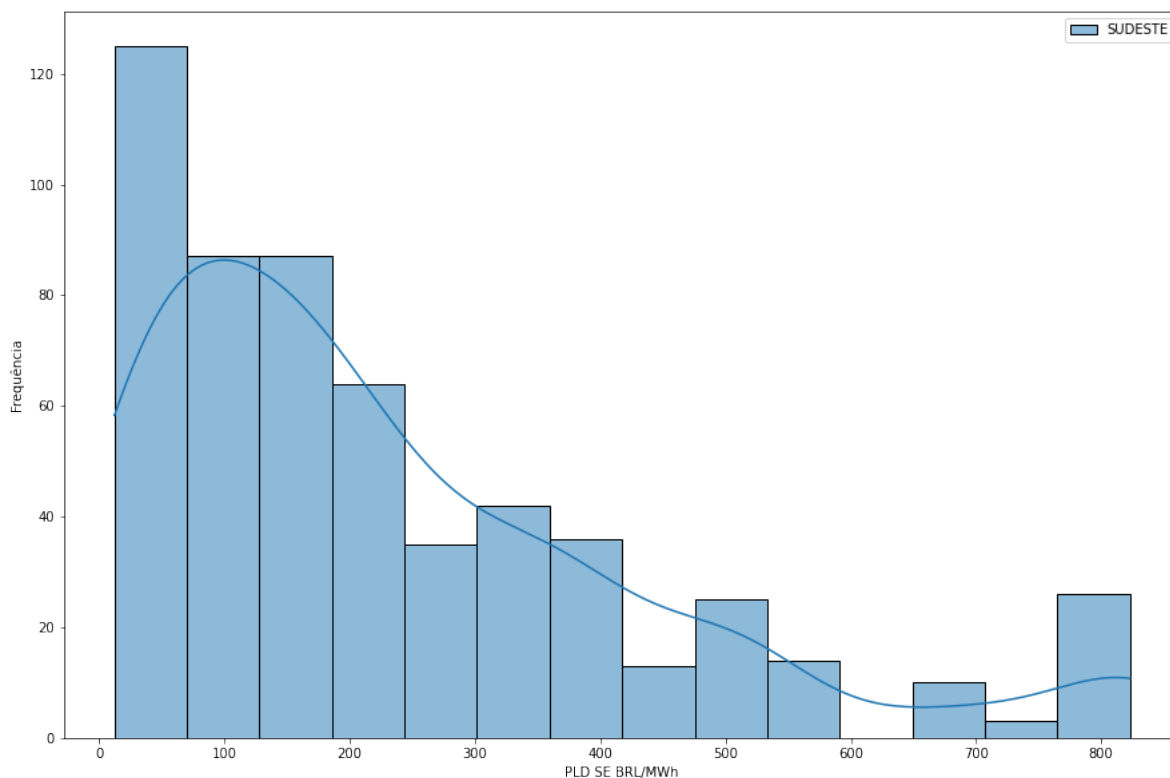


Figura 26 – Histograma do PLD SE (R\$/MWh)

5.1.2 Variáveis Predictoras

O site do Operador Nacional do Sistema (ONS) disponibiliza diversos dados relacionados à operação do Sistema Interligado Nacional (SIN), dessa maneira buscando representatividade e simplificação, foram escolhidos dados que tenham correlação significativa com o PLD SE, sendo apenas três variáveis que refletem bastante os impactos no preço de energia, as quais são: i) CMO (Custo Marginal de Operação); ii) Energia Armazenada GWh (responsável por informar sobre a energia armazenada nos reservatórios das hidrelétricas); iii) Demanda Máxima Horária MWh/h (energia máxima consumida por hora em uma semana operativa). Através destas variáveis espera-se refletir diversas situações que impactam no preço final de energia, como exemplo: o crescimento econômico do país, quantidade de novos investimentos em geração/transmissão de energia, quantidade de chuvas e aumento/ diminuição do custo de energia (ex.: necessidade de despacho das usinas térmicas).

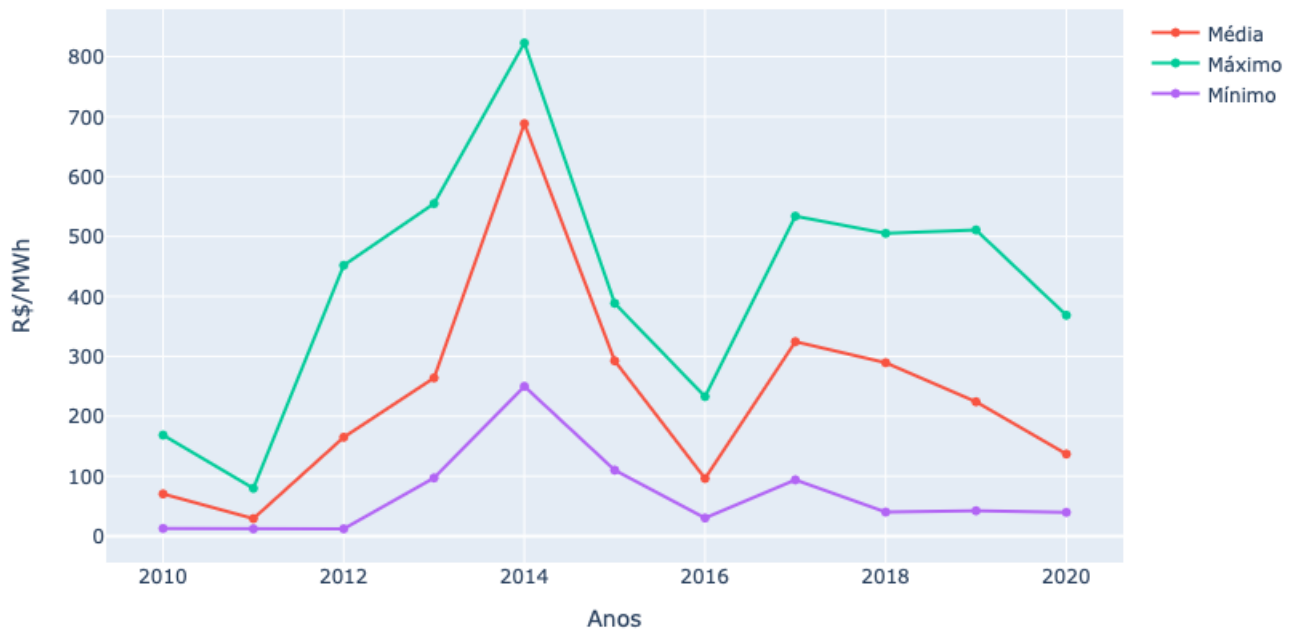


Figura 27 – Comparativo entre Preços Máximos, Mínimos e Médio

Para o processo de previsão é muito importante ter certeza sobre a escolha correta das variáveis predictoras, ou seja, entender se realmente há uma correlação a ser explorada. Outro ponto muito importante é que as variáveis predictoras estão fortemente correlacionadas, o que pode vir a ser um problema em aplicações de regressão, uma vez que podem aumentar a variância dos coeficientes de regressão e torná-los instáveis. Dessa maneira, na Figura 28, no *heat map*, é possível analisar sobre a correlação de cada uma das variáveis em relação ao preço, sendo que fica evidente que todas têm uma contribuição. Onde a maior correlação está com o Custo Marginal de Operação, enquanto que a correlação mais fraca é com a demanda máxima horária.

Conforme já mencionado em outros capítulos deste trabalho, a energia hidrelétrica, aquela gerada através do recurso hídrico, representa ainda uma parcela considerável de toda energia gerada no Brasil. Assim, é possível esperar que momentos de crises hídricas deverão ter um impacto muito considerável no preço de energia. Pois, nesses momentos a estratégia nacional energética prioriza a geração através das fontes térmicas, normalmente contratadas previamente em leilões de capacidade, para suprir esse déficit, o que encarece a matriz e impacta os consumidores através do aumento do preço. Na Figura 29, o gráfico ilustra a relação inversamente proporcional entre a energia armazenada nos reservatórios (nível de água) e o preço final.

A demanda máxima está também totalmente relacionada com o preço de energia, apesar dos agentes governamentais (ANEEL, ONS, EPE, etc) estarem sempre atentos realizando o planejamento do sistema, através de novas fontes entrantes e linhas de transmissão, para que possa suprir a demanda esperada pelo crescimento econômico previsto sem que haja uma pressão nos preços de energia, além de minimizar o risco de desabastecimento. Contudo, como esse tipo de

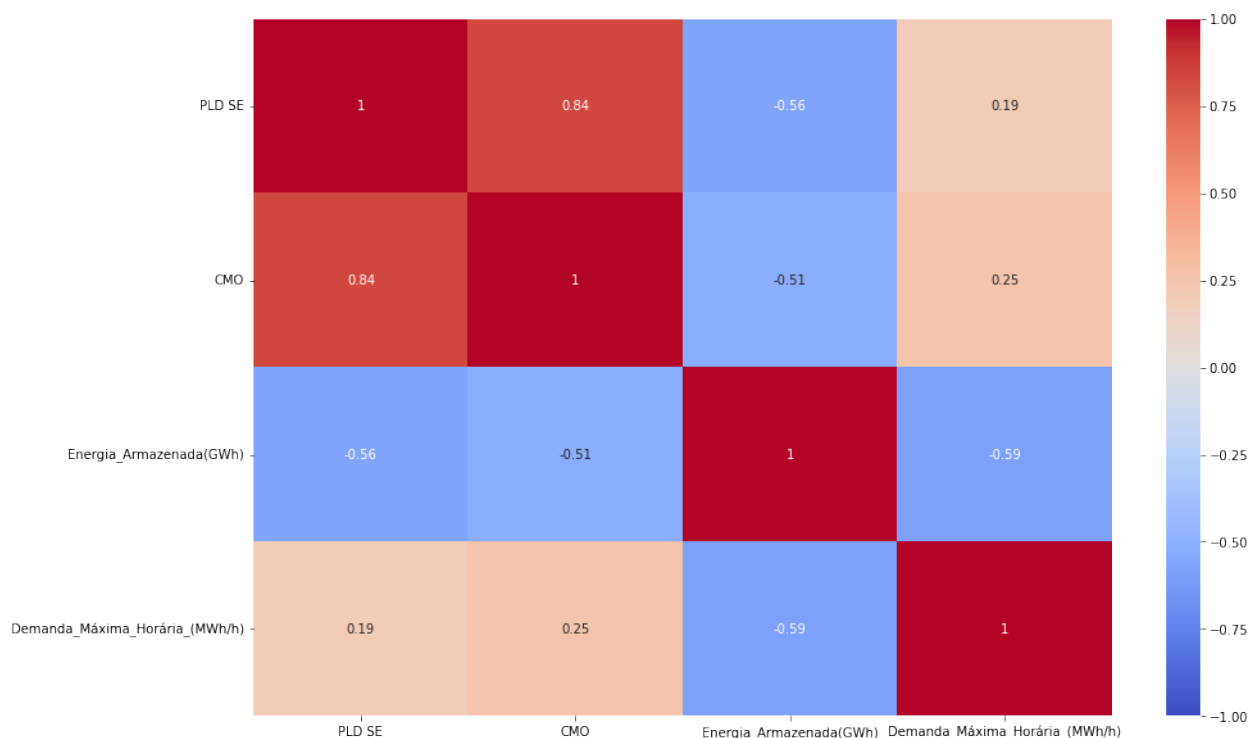


Figura 28 – Correlação entre as Variáveis

obra de infraestrutura tem um tempo longo para ser concluída, essa previsão precisa ser feita com antecedência e visa prever com base em dados o comportamento da carga no futuro, entretanto, como qualquer previsão, está sujeito a falhas e também a eventos inesperados.

Para exemplificar o comportamento inesperado da demanda máxima do sistema nacional, na Figura 30, representando a demanda máxima ao longo do tempo, há uma evidência clara sobre como o consumo de energia do país cresce ano após ano, apesar das oscilações de curto prazo. Porém, em meados de 2020 houve uma queda bastante acentuada, que foi provocada pelo início da pandemia do COVID-19, refletindo as políticas de *lockdown* e outras medidas sanitárias que impactaram na economia de forma geral. Nesse período houve uma sobre contratação de energia diante da queda da demanda, o que provocou queda nos preços de energia e impactos nas receitas dos agentes do sistema elétricos. O mesmo efeito de queda abrupta na demanda ocorreu em maio/2018 com o evento que ficou conhecido como greve geral dos caminhoneiros, cujo impacto se refletiu na diminuição do ritmo econômico naquele período. Evidenciar tais acontecimentos e seus impactos são importantes para entender que toda previsão tem suas imprecisões e está sujeito a eventos não esperados.

Uma análise sobre a variabilidade das variáveis a serem previstas ajuda a entender a sua natureza e comportamento, além de obviamente traduzir a dificuldade em prevê-las. Dessa maneira, foi utilizado um conceito bastante aplicado no mercado financeiro, conforme mostra (NETO, 2018), o retorno de um ativo calculado através das diferenças logarítmicas.

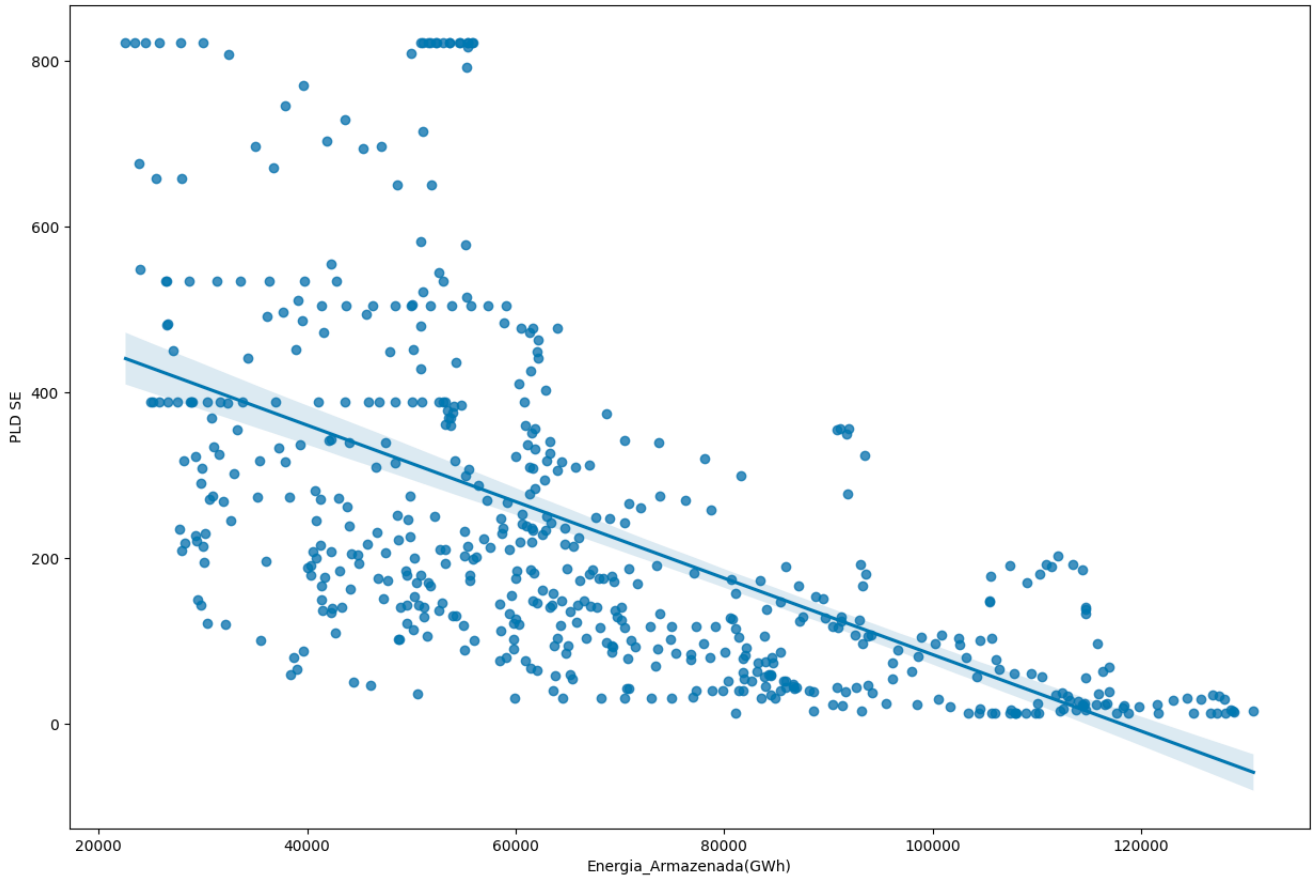


Figura 29 – Relação Nível dos Reservatórios x Preço de Energia

$$\text{Log Retorno} = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (5.1)$$

No caso das aplicações financeiras o conceito consiste em calcular a diferença entre o preço anterior (t-1) e o atual (t), sendo possível utilizar modelos estatísticos para quantificar a variabilidade do ativo. No caso específico deste trabalho, cuja natureza das variáveis divergem entre elas, pois são medidas específicas que quantificam por exemplo custos de produção e consumo de energia, a ideia é utilizar a mesma aplicação e comparar o coeficiente de variação para inferir sobre a variabilidade relativa entre elas. Lembrando que o coeficiente de variação é dado pelo desvio padrão dividido pela média, conforme equação 5.2, considerando que a média seja diferente de zero.

$$CV = \frac{\sigma}{\mu} \quad (5.2)$$

Dessa maneira, foi aplicada a metodologia em uma parcela dos dados (de 16/12/2016 até 18/01/2019), sendo que nada impediria de realizar em sua totalidade. Na Tabela 4 tem-se o resultado, sendo CMO a variável com maior variabilidade e a Energia Armazenada com a menor durante o período.

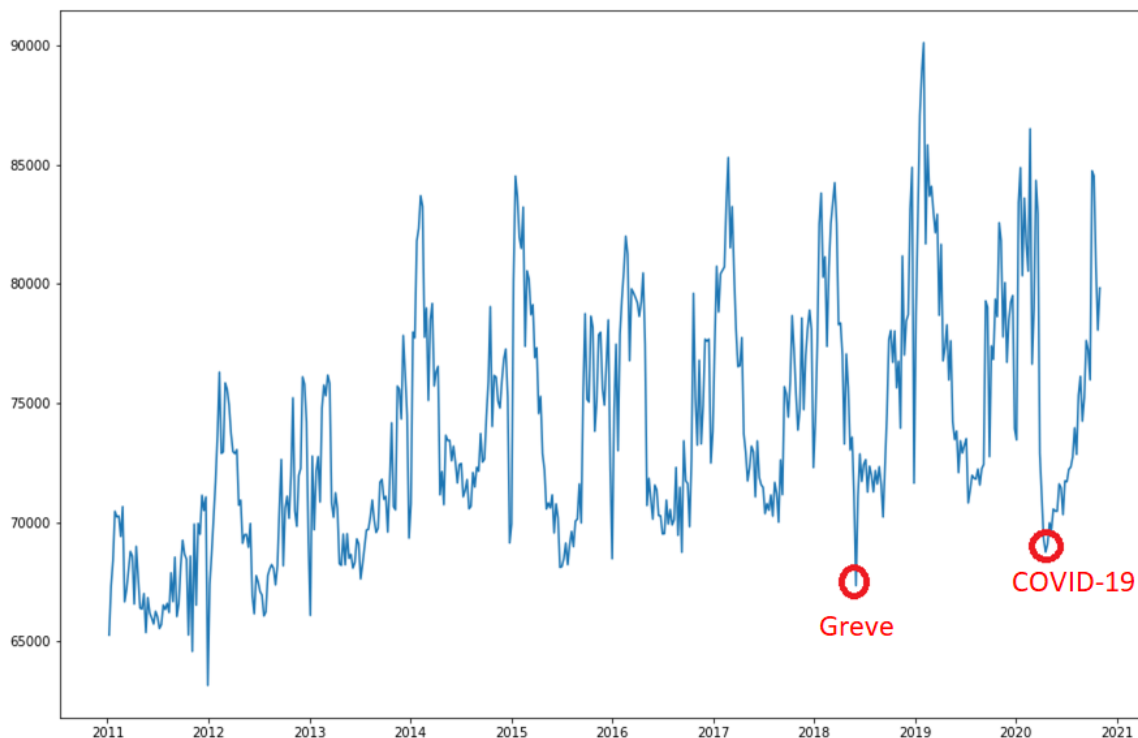


Figura 30 – Demanda de Energia Máxima

Tabela 4 – Análise de Variabilidade

	Ret CMO	Ret Energia Armazenada	Ret Demanda Máxima
Coeficiente de Variação	38,46	0,52	37,45

Na Figura 31 tem-se a função retorno com os dados obtidos para uma semana, ou seja, qual foi a variação para cada variável analisada entre uma semana subsequente. A CMO, conforme mostrado anteriormente, variou consideravelmente mais do que as outras, sendo que a linha vermelha mostra um intervalo de -2,5% até 2,5% enquanto que a linha verde indica -0,2% até 0,2%.

A Variabilidade da variável CMO pode ser explicada pela dispersão dos seus dados, conforme Figura 32 com o boxplot incluindo todas as observações desde 07/01/2011 e mostrando uma elevada quantidade *outliers*.

5.2 Previsão das Variáveis Predictoras usando Modelos SARIMA

Uma vez analisadas todas as variáveis predictoras sob ótica da estatística descritiva, agora se inicia o processo de modelagem, a constatação sobre sua eficácia e, finalmente, os valores

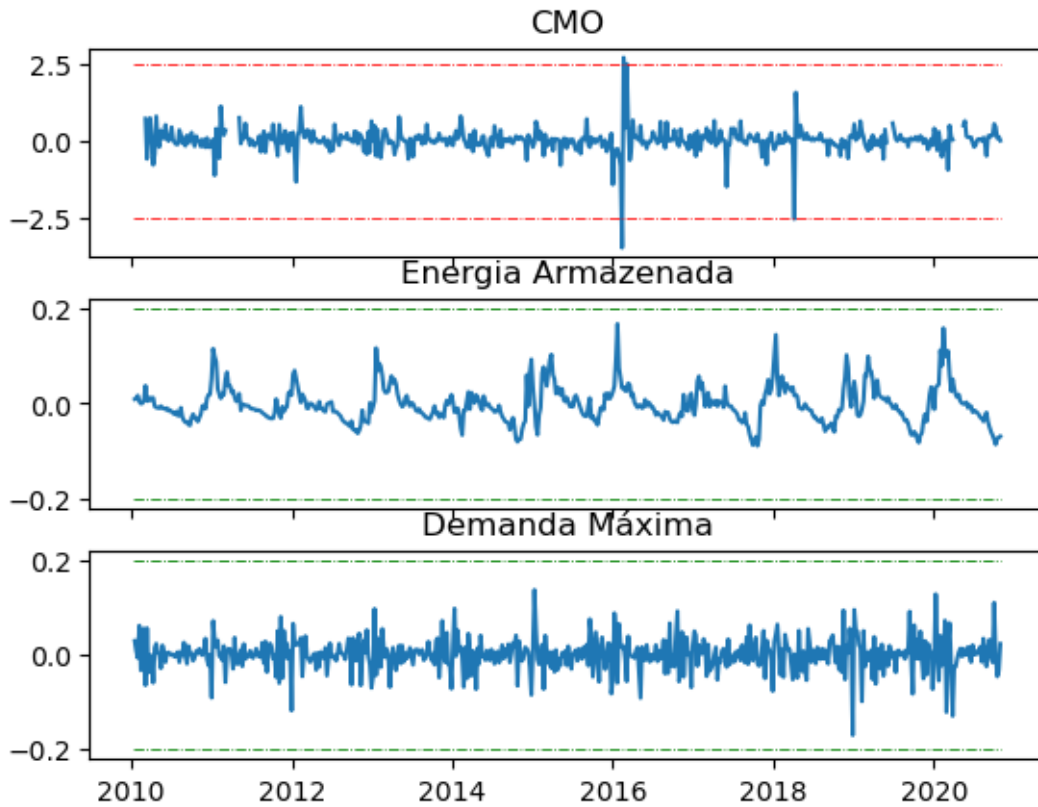


Figura 31 – Retorno das Séries Utilizadas

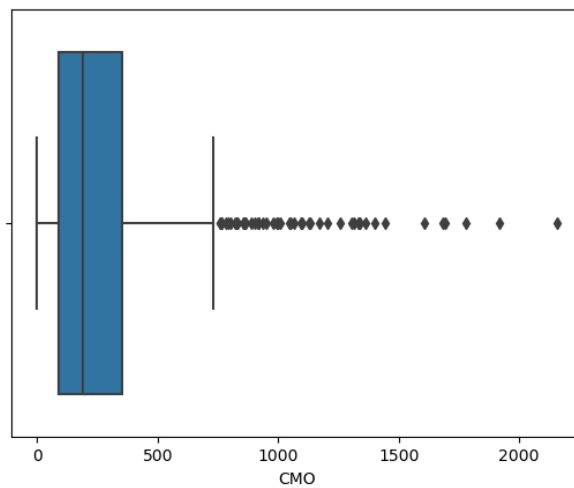


Figura 32 – Boxplot CMO

futuros que serão base para alimentar o modelo de redes neurais.

5.2.1 Decompondo as Séries

Ao decompor as séries, conforme as Figuras 33, 34 e 35, nota-se que a Demanda Máxima é crescente ao longo do tempo, enquanto que a Energia Armazenada e CMO têm uma tendência decrescente e estável, respectivamente.

Outro ponto é que as ambas possuem uma componente sazonal presente que deve ser refletida no desenvolvimento do modelo.

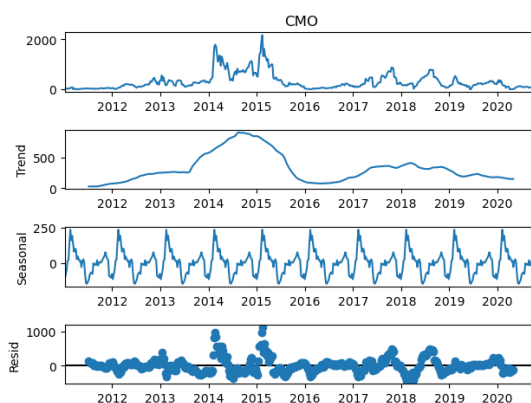


Figura 33 – Decomposição da Série CMO

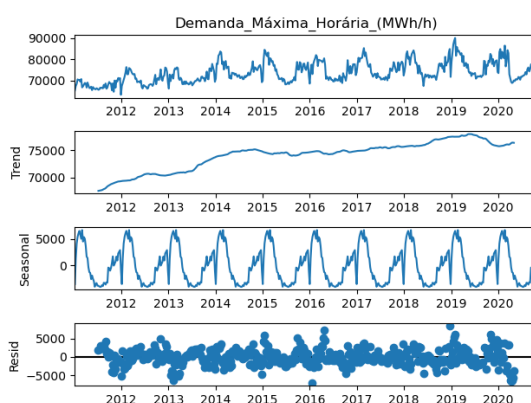


Figura 34 – Decomposição das Série Demanda Máxima

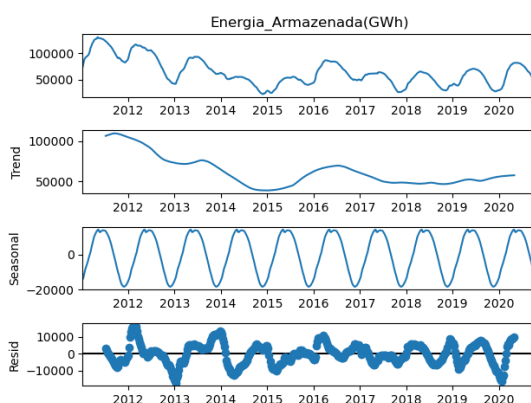


Figura 35 – Decomposição das Série Energia Armazenada

Aplicando o teste de Dickey Fuller, para verificar a estacionariedade ou não de cada série, conforme Tabela 5 nota-se que somente para a variável de Energia Armazenada não se pode rejeitar a hipótese nula ($p\text{-value} > 5\%$) e, portanto, entende-se que neste caso a série não

é estacionária. Dessa maneira, lembrando que o parâmetro d do modelo ARIMA refere-se a quantidade de diferenciação, é possível já concluir que a energia armazenada deverá passar pelo processo de diferenciação pelo menos uma única vez ($d = 1$) para eliminar a estacionariedade.

Tabela 5 – Estacionariedade das Variáveis Predictoras

	CMO	Demanda Máxima	Energia Armazenada
p-value	0,0060%	0,239%	8,1%

Na sequência tem-se os gráficos das funções de autocorrelação e autocorrelação parcial para entender como é o comportamento e influência dos dados mais antigos em relação aos mais recentes. Também através do comportamento de cada um, seja por ter uma decaimento exponencial ou lento, ou ainda verificando seu truncamento, é possível estimar os parâmetros do modelo (pdq).

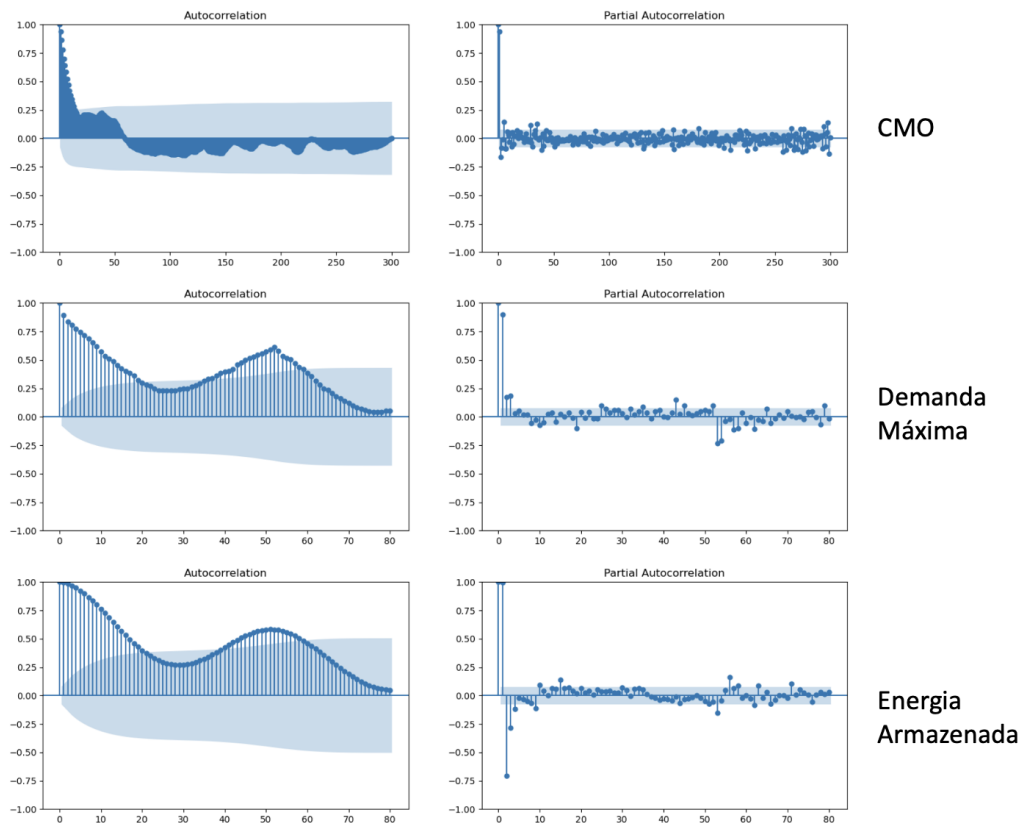


Figura 36 – Autocorrelação

Outra informação bastante relevante que a Figura 36 fornece é referente à sazonalidade das variáveis. Indicando que os dados completam ciclos após 52 semanas, o que significa que possui uma sazonalidade de um ano, uma vez que as amostras são semanais.

5.2.2 Modelagem e Previsão

Primeiramente, foi modelada a série referente à Demanda Máxima Horária (MWh/h) através do SARIMA obtendo como o melhor modelo o SARIMA(1,0,0)(0,1,2)₅₂, ou seja cujo o elemento auto-regressivo foi de ordem 1, a parte sazonal como P=0, D=1, Q=2 e a sazonalidade de 52, o que representa a quantidade de semanas em um ano calendário (52).

Para que seja possível testar a eficiência do modelo os dados foram divididos em dois grupos, teste (24,4%) e treino (75,6%), sendo que os de treino incluem todos os dados de 01/01/2012 até 01/01/2019 enquanto que os de teste são de 01/01/2019 até 02/04/2021. O resultado é mostrado na Figura 37, com a linha azul sendo os dados reais e a linha vermelha aqueles ajustados.

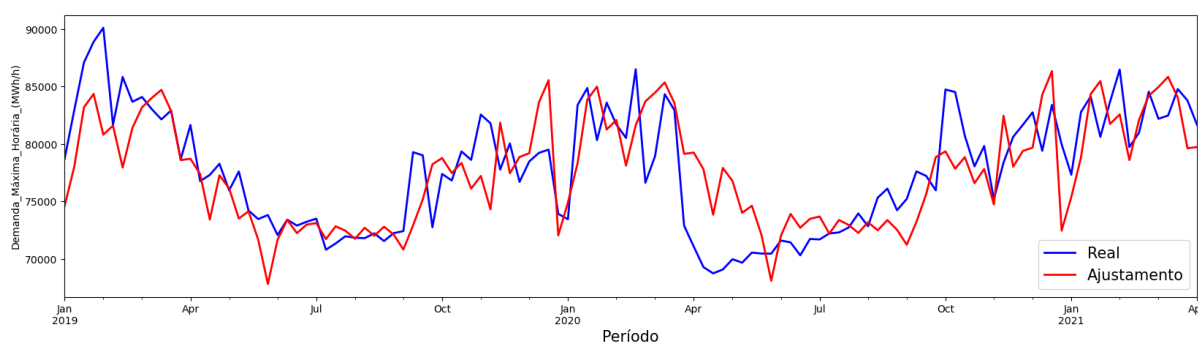


Figura 37 – Demanda Máxima: Ajustamento x Real

Nota-se que o modelo conseguiu captar as principais características da série, como a sazonalidade, tendência e pontos máximos e mínimos, conforme mostrado na Figura 38.

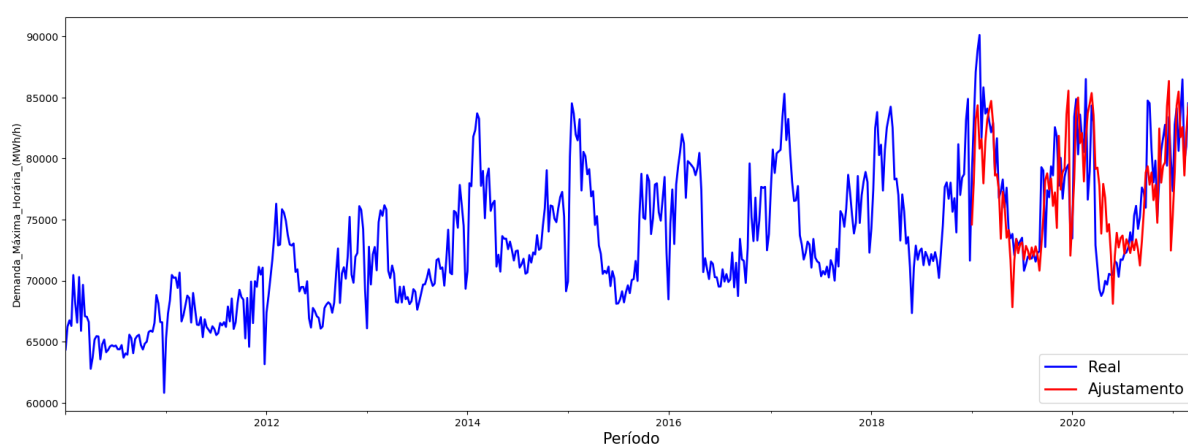


Figura 38 – Demanda Máxima: Ajustamento x Real - Base de Dados Completa

Verificando a análise residual conforme a Figura 39 tem-se que o resíduo não é ruído branco uma vez que a função de autocorrelação amostral no lag 1, intervalo de tempo entre observações em uma série temporal, é significativa, além do que nas caudas esses resíduos se afastam de uma distribuição normal. Dessa maneira, é possível assumir que a modelagem foi bastante

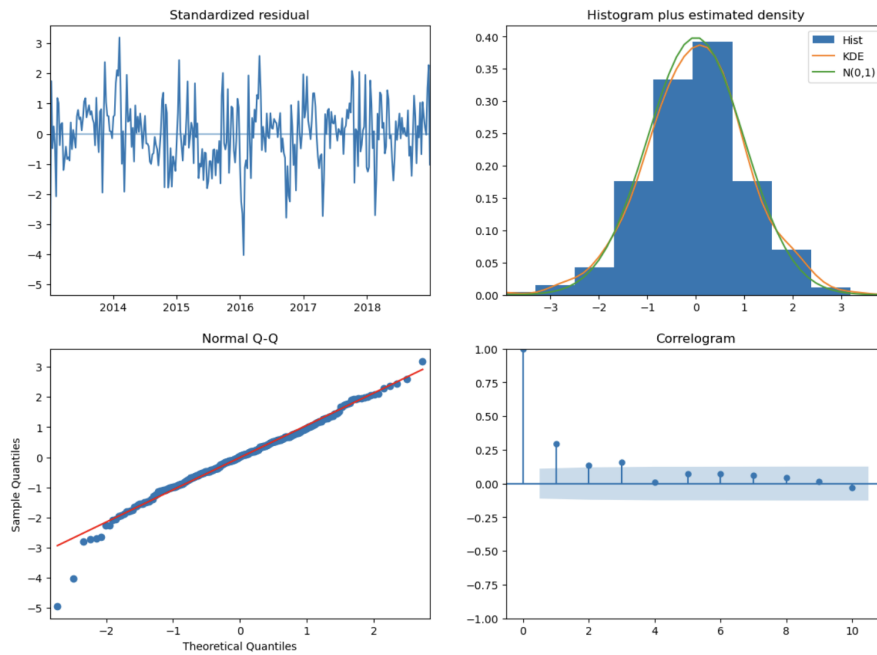


Figura 39 – Análise Resíduos Ajustamento Demanda Máxima

efetiva ao capturar as principais características da série temporal em estudo. Ainda na Figura 39, a função densidade estimada dos resíduos, representada por KDE, e a distribuição normal padrão, representado por $N(0,1)$, estão muito próximas indicando o sucesso da modelagem.

Enquanto que para a variável Energia Armazenada(GWh) também foram separados dados para treino (63,2%) e teste (36,8%) e o modelo que obteve a melhor performance foi o SARIMA(1,1,1)(0,1,1)₅₂, com o elemento auto-regressivo é de ordem 1, sendo necessário diferenciar uma vez e a ordem do elemento de média móvel é de ordem 1.

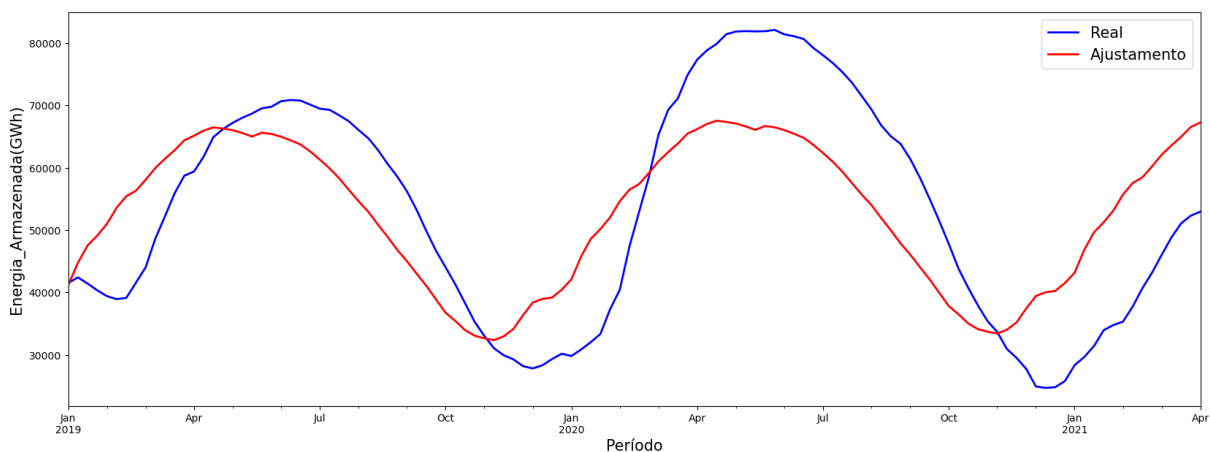


Figura 40 – Energia Armazenada - EAR: Ajustamento x Real

Por fim, ao analisar a série completa e a previsão realizada, conforme Figura 41, conclui-se que a modelagem estatística foi capaz de captar as características da série temporal.

Ao verificar a base de dados completa (linha azul) observa-se que há uma mudança

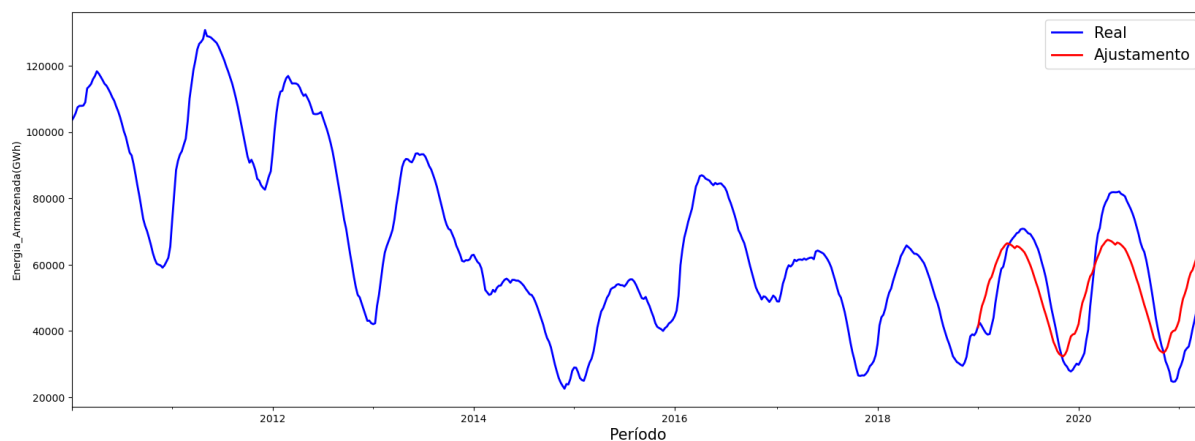


Figura 41 – Energia Armazenada - EAR: Ajustamento x Real - Base de Dados Completa

de padrão sazonal, principalmente após 2015 com mínimos e máximos se comportando de forma distinta do histórico. Dessa maneira, para o ajustamento do modelo utilizou-se os dados a partir de 01/01/2015 para que assim fosse possível reduzir o ruído na previsão, apesar de ainda constatar que há um atraso (linha vermelha) na previsão quando comparada com os dados reais.

Realizando a mesma análise dos resíduos, também fica evidente, conforme Figura 42, que a modelagem para a variável Energia Armazenada atingiu um nível de eficácia aceitável, uma vez que não há nenhum tipo de viés constatado, apesar da autocorrelação residual de ordem 1 e 2 serem insignificantes.

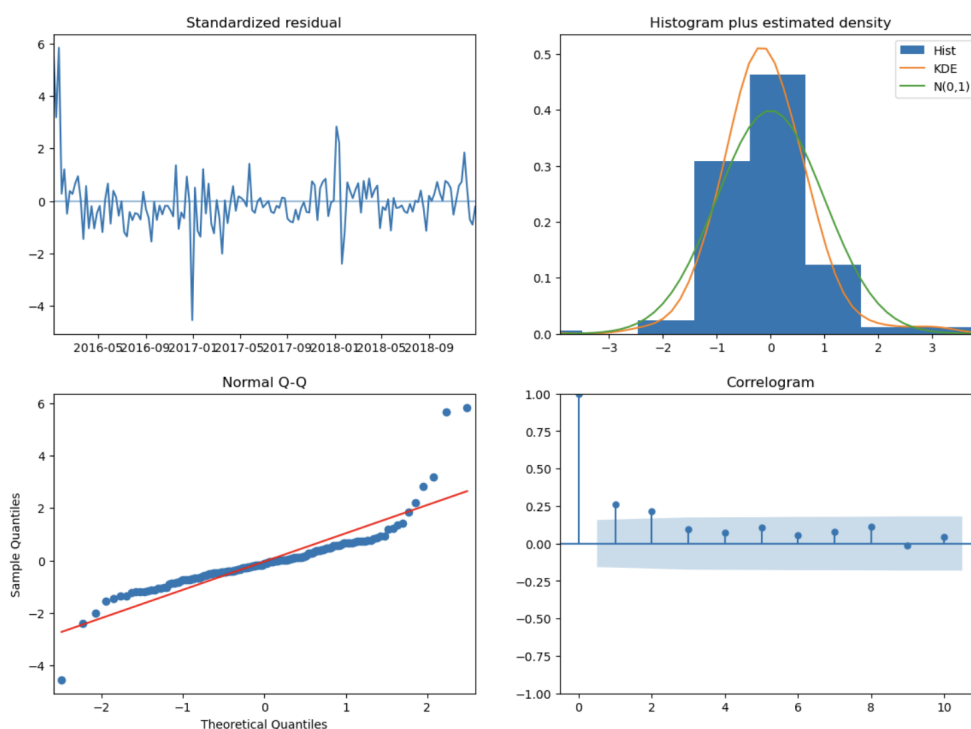


Figura 42 – Análise Resíduos Ajustamento Demanda Máxima

Por último, para encerrar a previsão das séries das variáveis preditivas, tem-se a mo-

delagem do SARIMA para a CMO. E, de forma similar ao realizado nas séries anteriores, os dados foram separados em teste (31,7%) e treino (68,2%). Com isso o modelo com o melhor desempenho foi o SARIMA(1,0,0)(2,1,0)₅₂, ou seja, cujo o elemento auto-regressivo foi de ordem 1, a parte sazonal como P=2, D=1, Q=0 e a sazonalidade de 52, o que representa a quantidade de semanas em um ano calendário (52).

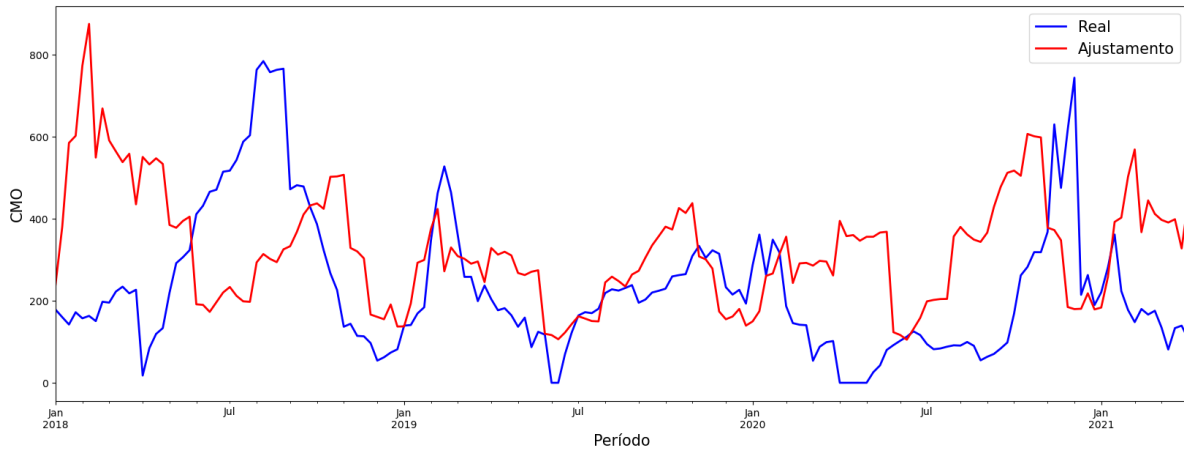


Figura 43 – CMO: Ajustamento x Real

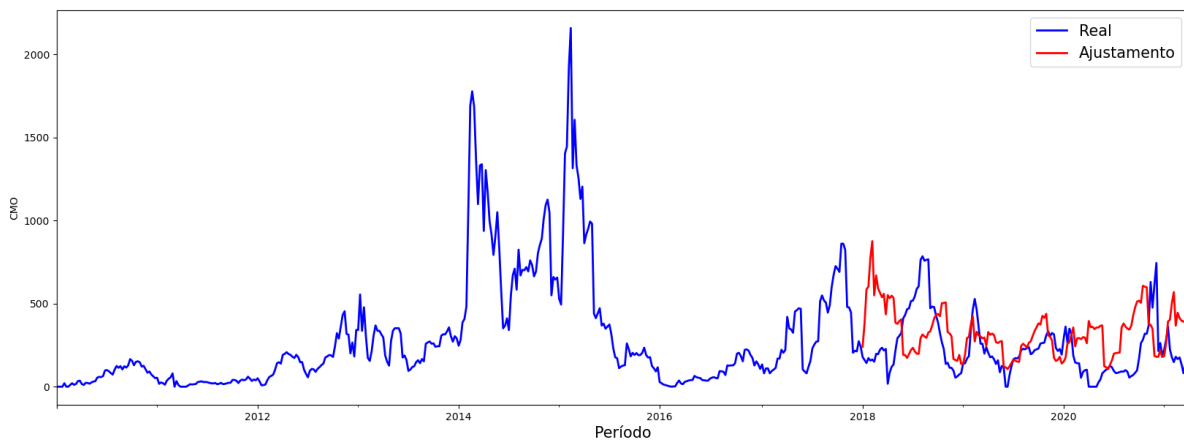


Figura 44 – CMO: Ajustamento x Real - Gráfico Completo

Finalmente, após toda a modelagem desenvolvida, é possível prever os quatro valores subsequentes, tanto para a Demanda Máxima, Energia Armazenada e CMO, que serão as entradas da rede neural que irão produzir o PLD, a ser abordado no capítulo 6. Na Tabela 6 tem-se os valores previstos para cada variável preditora.

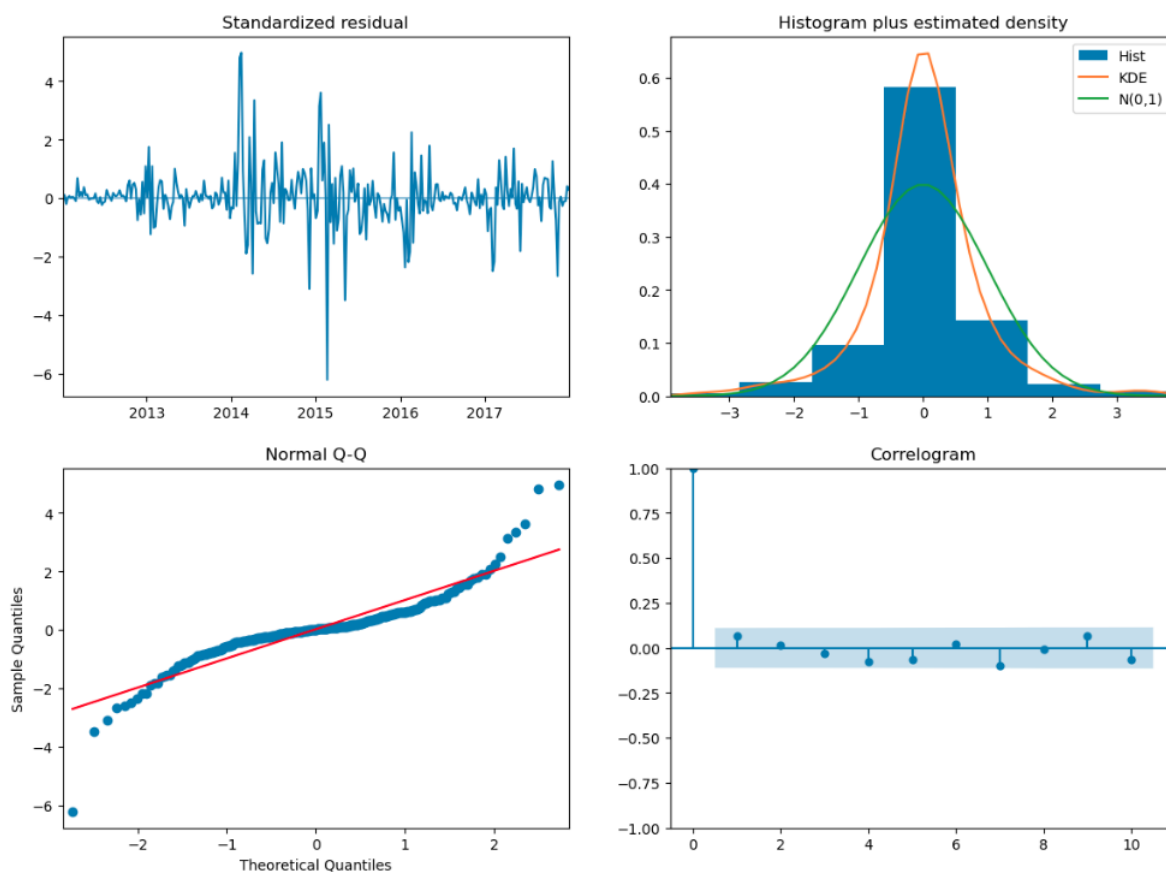


Figura 45 – Análise Resíduos Previsão CMO

Tabela 6 – Previsão das Variáveis Predictoras

Semana Operativa	Demanda Máxima (MWh/h)	Energia Armazenada (GWh)	CMO (R\$/MWh)
09/04/2021	73.452,8	53.887,2	118,3
16/04/2021	70.604,9	54.925,8	127,6
23/04/2021	70.156,3	55.723,4	125,7
30/04/2021	70.803,2	55.998,0	144,0

Sumarizando, todo processo de modelagem deve passar por uma fase de análise dos dados, onde se espera consolidar o conhecimento sobre o comportamento das variáveis e já neste momento tirar algumas conclusões que contribuirão para a interpretação e previsão, além de identificar valores atípicos (outliers), que a depender da variável pode indicar desde possíveis erros de medição até eventos inesperados. Em um segundo momento, é feita a modelagem escolhendo alguma métrica como parâmetro comparativo (neste caso foi o RMSE) para que se comparem os resultados obtidos para cada modelo e escolha aquele que melhor se comportou. Por fim, é importante analisar o resíduo do processo e entender se há algum viés ainda que poderá indicar necessidade de uma investigação sobre a modelagem.

Assim, tem-se finalizada a primeira etapa do modelo proposto para a previsão do PLD. No segundo passo, os dados resultantes da previsão deste capítulo serão as entradas da rede neural a ser modelada no próximo capítulo.

PREVISÃO COM RNA

Partindo para a etapa final do trabalho, que deverá modelar a rede neural LSTM tendo como entradas os dados históricos do PLD, Energia Armazenada, CMO e Demanda Máxima, serão utilizados os valores previstos no capítulo 5 para então produzir o preço da energia para os quatro passos seguintes a partir da data de 02/04/2021.

Finalmente, será feita uma análise final entre os dados obtidos e os reais para concluir a eficácia do modelo completo proposto neste trabalho.

6.1 Explorando a Lógica da Análise Combinada

Para prever o preço de energia poderia utilizar-se de técnicas que não fossem combinadas, isto é, aplicando por exemplo o modelo ARIMA direto na série do PLD, ou mesmo a redes neurais recorrentes. Entretanto, a ideia central deste trabalho é aliar duas poderosas técnicas e com isso obter uma previsão que seja mais precisa. Para isso, nesta segunda etapa deve-se tomar os valores previstos no capítulo 5 e inseri-los na lógica de aprendizado da LSTM para que assim contribua com a previsão do PLD.

A Figura 46 representa como os dados estão estruturados, de tal forma que as variáveis X, Y e K influenciam na Z e podem ser divididas de tal maneira para que a rede seja treinada. Entretanto, para as quatro valores previstos no capítulo 5 (representados como X_n até X_{n+3} e Y_n até Y_{n+3}) o que se busca é justamente os valores de Z, que representam os preços futuros do PLD.

Para que seja possível utilizar os dados histórico do PLD e a influência das variáveis predictoras para o treinamento do modelo, a coluna PLD será deslocada uma posição para trás (n-1), assim ao treinar a rede espera-se que modelo possa captar qual é a influência da Demanda Máxima, CMO e Energia Armazenada no preço de energia. A Figura 47 tem os dados que são representados novamente, porém com os valores do PLD deslocados uma posição.

Predictoras			Target
CMO	EAR	Demand Máx.	PLD
K_1	X_1	Y_1	Z_1
K_2	X_2	Y_2	Z_2
K_3	X_3	Y_3	Z_3
K_4	X_4	Y_4	Z_4
K_5	X_5	Y_5	Z_5
·	·	·	·
·	·	·	·
·	·	·	·
K_n	X_n	Y_n	???
K_{n+1}	X_{n+1}	Y_{n+1}	???
K_{n+2}	X_{n+2}	Y_{n+2}	???
K_{n+3}	X_{n+3}	Y_{n+3}	???

Figura 46 – Representação da Estrutura dos Dados

As LSTM's consideram uma janela de dados anteriores a ser configurada livremente para que se busque a melhor modelagem, conforme Figura 47 com $K_2 X_2 Y_2 Z_2$ até $K_4 X_4 Y_4 Z_4$. Dessa maneira, pode-se criar um algoritmo para prever o próximo valor de Z (Z_6). E, seguindo a mesma lógica, é possível treinar os dados dessa maneira ajustando os parâmetros da LSTM para que sejam previstos os quatro passos seguintes ($Z_{n+1} Z_{n+2}, Z_{n+3}, Z_{n+4}$).

6.2 Modelagem da LSTM

Para elaboração de modelos em Python é muito comum a utilização da API específica Keras em conjunto com o modelo Sequential. Sendo que os principais parâmetros são:

1. Camadas: adicionar a quantidade de camadas intermediárias;
2. Neurônios: definir a quantidade de neurônios para cada camada;
3. Loss Function: função perda, ou seja, o erro. É necessário definir qual será a forma de minimização do erro. Para este trabalho foi utilizado método do Erro Quadrático Médio e o otimizador Adam;
4. Épocas: quantidade de iterações que serão realizadas no conjunto de dados de treino;
5. Batch: define o tamanho do lote que passará na rede;

CMO	EAR	Demand Máx.	PLD
K_1	X_1	Y_1	Z_2
K_2	X_2	Y_2	Z_3
K_3	X_3	Y_3	Z_4
K_4	X_4	Y_4	Z_5
K_5	X_5	Y_5	Z_6
.	.	.	.
.	.	.	.
.	.	.	.
K_n	X_n	Y_n	Z_{n+1}
K_{n+1}	X_{n+1}	Y_{n+1}	Z_{n+2}
K_{n+2}	X_{n+2}	Y_{n+2}	Z_{n+3}
K_{n+3}	X_{n+3}	Y_{n+3}	Z_{n+4}

Figura 47 – Representação da Estrutura dos Dados - Coluna PLD Deslocada

Na Figura 48 tem-se o trecho da programação em *Python* com a formação da LSTM utilizando a biblioteca *Keras*.

```

✓[2145]model = Sequential()
15 model.add(LSTM(100, return_sequences = True, input_shape = (X_train.shape[1],3)))
model.add(LSTM(100, return_sequences = True))
model.add(LSTM(80, return_sequences = False))
model.add(Dense(50))
model.add(Dense(1))

✓[2146]model.compile(optimizer = 'adam', loss = 'mean_squared_error')
05

[2147]model.fit(x=X_train, y=y_train, batch_size=10, epochs=100, shuffle=True, validation_split = 0.1)

```

Figura 48 – Exemplo de Implementação LSTM em Python

Outro parâmetro importante, apesar de não fazer parte diretamente do *Keras*, é a janela deslocável, que definirá quantos dias anteriores está sendo considerado na previsão do próximo valor. Sendo que o melhor resultado para a modelagem foi uma janela de três períodos, o que significa três semanas operativas anteriores ao valor previsto.

Então, os dados foram divididos entre treino (75%) e teste (25%) e, após diversas iterações para encontrar a melhor modelagem, chegou-se a um resultado com os parâmetros com uma performance bastante satisfatória, conforme Figura 49 com um resumo de onze redes e seus respectivos RMSE.

Foram feitas diversas simulações de rede alterando seus parâmetros como indicado na Figura 49. Como medida de performance considerou o RMSE para os dados normalizados.

Dessa maneira, a rede número 3 obteve o melhor desempenho com RMSE sendo o menor dentre os demais. Assim, a modelagem foi utilizada para a previsão final dos dados.

Rede	Neurônio 1° Camada	Neurônio 2° Camada	Neurônio 3° Camada	Camada Densa	Janela	Épocas	Batch	RMSE (dados normalizados)
1	100	100	80	50	4	100	10	0,0491
2	100	100	90	50	4	100	10	0,0048
3	100	100	90	50	4	70	10	0,0011
4	100	100	120	50	4	100	10	0,0311
5	100	90	100	50	4	100	10	0,0426
6	100	90	80	100	4	100	10	0,0150
7	200	200	80	50	8	100	10	0,02916
8	50	50	40	50	8	100	10	0,04243
9	100	100	80	50	3	100	8	0,03151
10	100	100	80	50	3	100	15	0,03900
11	100	100	80	50	3	100	20	0,00932
12	100	100	80	50	3	150	10	0,02103

Figura 49 – Rede LSTM x RMSE

Na Figura 50 tem-se a evolução do erro após cada iteração, sendo o erro referente ao conjunto de dados de treino representado pela linha azul e medindo a capacidade da rede neural se adaptar aos dados utilizados para treiná-la. Enquanto que a linha em laranja refere-se ao erro de validação, cujo objetivo é medir a performance da rede dado um conjunto de validação. As duas curvas são decrescentes com um compartimento semelhante à uma exponencial e atingindo valores muito próximos de uma estabilidade, ou seja, conforme mostrado na Figura 49 o aumento do número de épocas (simulação 12) não se traduziu em uma melhora das performances.

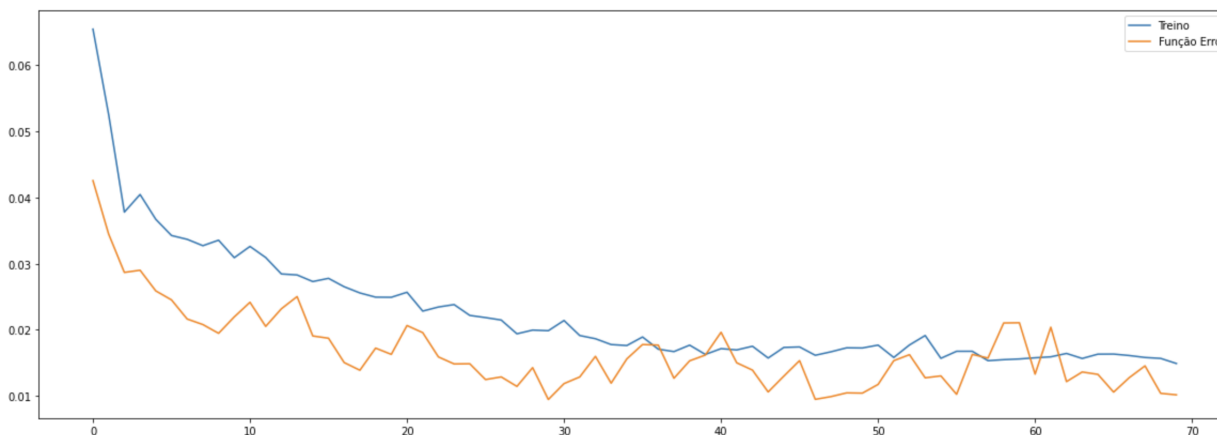


Figura 50 – Erros Modelagem

A Figura 51 condensa os valores ajustados, em vermelho, e os reais utilizados para teste, em azul. Dessa maneira, fica muito evidente que o modelo realmente captou as principais características com valores de teste muito próximos dos reais.

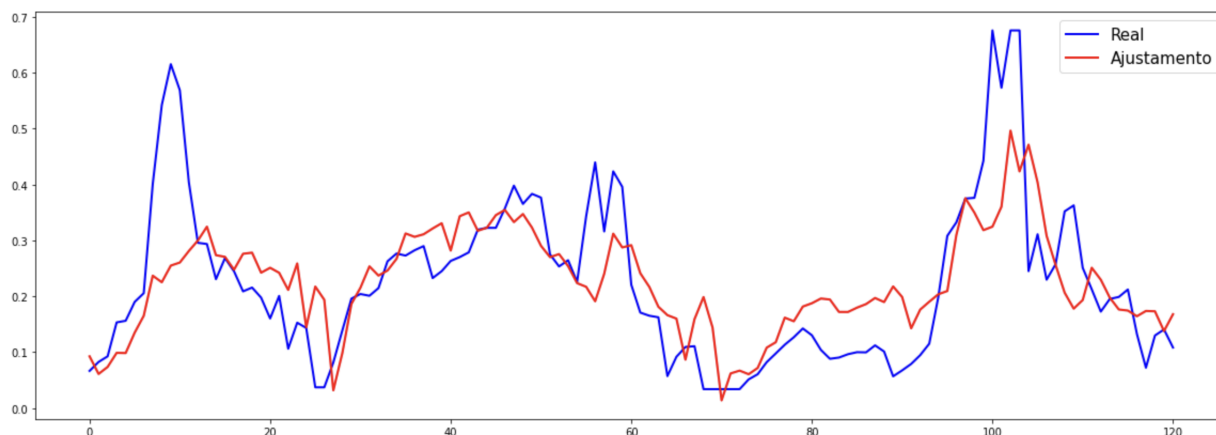


Figura 51 – Previsão PLD: Real x Ajustamento - Dados Normalizados

6.3 Resultado Previsão

Após a modelagem completa e finalizada é possível fazer a previsão para os dados futuros, ou ainda não expostos para à rede. Na Tabela 7 são apresentados os valores previstos, os reais e o erro absoluto entre eles.

Tabela 7 – Valores Previstos x Real

Semana Operat.	Real	Previsto	Erro Absoluto
09/04/2021	125,43	157,68	32,25
16/04/2021	108,77	135,01	26,24
23/04/2021	161,94	134,66	27,28
30/04/2021	146,95	138,54	8,41

Assim, ao observar que o modelo foi capaz de produzir informações de valor, uma vez que mostrou em sua previsão uma variação dentro de uma faixa de preço próxima à dos preços da energia para o próximo mês, além de alguns valores previstos serem bem próximos dos reais.

A diferença mostrada pelo erro absoluto pode ser considerada como aceitável, pois conforme havia sido mostrado na Tabela 3, na análise sobre o histórico do PLD, o seu preço variou desde o mínimo de 12,08 até o valor máximo de 822,83. Além disso, é uma variável com alta variabilidade (desvio padrão de 204,76). Portanto, considerando que a erro absoluto médio foi de apenas 23,55, o resultado foi bastante satisfatório.

CONCLUSÃO

Na tarefa de propor uma análise combinada de técnicas para previsão do preço de energia, cujo objetivo é gerar uma maior transparência na tomada de decisão para os agentes do sistema elétrico, utilizou-se das técnicas de séries temporais para previsão das variáveis preditoras escolhidas e de redes neurais recorrentes, especificamente LSTM, para então prever o PLD.

Foram escolhidas como variáveis preditoras a energia armazenada, o que reflete o nível dos reservatórios das hidrelétricas do país, a demanda máxima de energia, normalmente impactada pela atividade econômica e o CMO representando o custo de produção de energia em um determinado momento. Então, aplicando as técnicas de séries temporais foi possível obter uma previsão de tais variáveis com um grau de acerto dentro de valores aceitáveis considerando o modelo SARIMA, aquele que considera a sazonalidade dos dados em sua previsão, cuja característica para todas as séries foi uma sazonalidade anual.

O estudo de previsão de preços, seja commodities ou ações, tem produzido diversos trabalhos pois envolve tanto o potencial de impacto financeiro como a dificuldade de tal tarefa. E, analisando os dados históricos do PLD, conforme demonstrado no capítulo 5, tem-se uma série temporal com alta variabilidade, o que torna a sua previsibilidade mais difícil. Isso se deve a diversos fatores e eventos inesperados que influenciam no preço de energia, como o caso do COVID-19. Entretanto, ao alimentar a rede neural modelada com os dados históricos e com os valores previstos de energia armazenada, CMO e demanda máxima de energia, foram produzidos valores que indicam sobre a tendência para o curto prazo. Obviamente que tal previsão isolada não fornece toda a certeza necessária para um gestor tomar decisão devido ao risco ligado à variabilidade, porém em um contexto maior espera-se que possa auxiliar neste processo de tomada de decisão.

Em suma, foi fornecida uma transparência sobre a tendência que o preço deverá seguir no horizonte futuro e atendendo o objetivo que o trabalho se propôs. Pois, apesar do desejo cada vez maior de obter previsões mais precisas através do avanço das técnicas utilizadas e

do processamento de dados, não há uma expectativa de que não haja erros neste processo. Afinal, quando se tenta enxergar o futuro o que se espera é obter uma visibilidade adicional para que possa ser incluída no processo de tomada de decisão juntamente com outros fatores não relacionados neste trabalho e assim fornecer subsídios para que se tenha uma decisão cada vez mais assertiva no processo de gerenciamento de recursos.

7.1 Trabalhos Futuros

Ao relacionar outros fatores de impacto no preço de energia é possível também sugerir como próximos trabalhos estudar outras variáveis preditoras que tenham correlação com o PLD. Apenas citando algumas delas tem-se: i) Preço do barril de petróleo: influenciará o valor da energia produzida pelas térmicas; ii) Atividade econômica: influenciará a demanda de energia; iii) Expectativa de chuvas: influenciará o nível dos reservatórios; iv) novas fontes de energia; vi) atividade da bolsa de valores; vii) IGMP-M: refletindo preços de commodities que impactam no CAPEX de investimento do setor elétrico (ex.: alumínio para construção de linhas de transmissão). Talvez, ao acrescentar, ou trocar, outras variáveis preditoras o modelo poderá responder de forma mais assertiva produzindo previsões mais precisas no médio e longo prazo.

Referente ao modelo é possível estudar outras maneiras para implementar melhorias na previsão, como exemplo a alteração da combinação das técnicas adotadas. Nesta linha talvez investigar qual é o modelo de previsão que melhor se encaixe para a variável preditora adotada, como exemplo trocar o SARIMA por uma outra técnica de machine learning (ex.: regressão através de SVM - Support Vector Machine) ou variáveis de estado. Pois, conforme comentado na seção 5.2.2, o ajustamento do modelo para a variável Energia Armazenada apresentou um atraso quando comparado com os dados reais, ou seja, seria interessante buscar ainda outros modelos específicos para cada variável, uma vez que têm comportamentos distintos.

Como uma sugestão para melhorar a precisão e deixar o processo mais automatizado, pode-se desenvolver um algoritmo para rodar em *looping*, inclusive com auxílio de algoritmos de machine learning, até que se chegue ao erro mínimo/ modelo ideal. Enquanto neste trabalho o processo de busca pelo melhor modelo foi realizado de forma manual, ou seja, para cada ajuste dos parâmetros o algoritmo simulado e assim sucessivamente para minimizar o erro.

Outra técnica que pode ser aliada na análise da tendência do preço, e quais seriam os cenários futuros, é a Simulação de Monte Carlo. Com esta técnica é possível tentar prever o comportamento estatístico do preço, se vai subir ou cair em um determinado intervalo de tempo futuro com suas respectivas probabilidades. Dessa forma, juntamente com a previsão gerada neste trabalho, pode-se entender qual seria o risco probabilístico de escolher um caminho, baseado na volatilidade do preço no passado, e assim produzir um cenário de tomada de decisão mais completo para os agentes.

REFERÊNCIAS

AMRUTHA, J.; AJAI, A. S. R. Performance analysis of backpropagation algorithm of artificial neural networks in verilog. In: **2018 3rd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)**. [S.l.: s.n.], 2018. p. 1547–1550. Citado na página 53.

ANDRADE, W. A.; MIRANDA, L. de J.; ACHIDA, T. A. C. TIPOS DE PREVISÃO DE DEMANDA E A APLICAÇÃO EM INDÚSTRIA AUTOMOTIVA PAULISTA. **Revista Fatec Zona Sul**, v. 3, p. 1 – 17, 2 2017. ISSN 2359-182X. Citado na página 31.

CCEE Painel de Preços. Disponível em: <<https://www.ccee.org.br/web/guest/precos/painel-precos>>. Citado na página 57.

EPE, E. de P. E. **Plano Decenal de Expansão de Energia 2031**. 2022. Urlhttps://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/Documents/PDE%202031_RevisaoPosCP_rvFinal_v2.pdf. Citado nas páginas 21, 22 e 24.

FELTRIN, F. B. **Redes Neurais Artificiais**. 1º. ed. [S.l.]: Uniorg, 2020. Citado na página 47.

FERREIRA, L.; MOURA, G. L. d.; BORENSTEIN, D.; FISCHMANN, A. A. Utilização de redes neurais artificiais como estratégia de previsão de preços no contexto de agronegócio. **INMR - Innovation amp; Management Review**, v. 8, n. 4, p. 6–26, jan. 2012. Disponível em: <<http://www.revistas.usp.br/rai/article/view/79235>>. Citado na página 33.

FERREIRA, P. G. C.; BARROS, A. C.; MATTOS, D. M. de; OLIVEIRA, I. C. L. de; DUCA, V. E. L. de A. **Análise de Séries Temporais em R: Curso Introductório**. 2º. ed. [S.l.]: GEN Atlas, 2017. ISBN 978-8535290875. Citado na página 39.

GÉRON, A. **Mãos à obra: aprendizado de máquina com Scikit-Learn TensorFlow**. 1º. ed. [S.l.]: Alta Books, 2019. ISBN 8550803812. Citado nas páginas 47, 48 e 52.

HAYKIN, S. **Neural Networks and Learning Machines**. 3º. ed. [S.l.]: Pearson, 2009. ISBN 0-13-147139-2. Citado nas páginas 47, 49 e 50.

HUA, Y.; ZHAO, Z.; LI, R.; CHEN, X.; LIU, Z.; ZHANG, H. Deep learning with long short-term memory for time series prediction. **IEEE Communications Magazine**, v. 57, n. 6, p. 114–119, 2019. Citado na página 54.

JUNIOR, J. R. F. **Redes Neurais Recorrentes — LSTM**. 2019. Url<https://medium.com/@web2ajax/redes-neurais-recorrentes-lstm-b90b720dc3f6>. Citado na página 55.

LEME, R.; TURRIONI, J.; BALESTRASSI, P.; SOUZA, A. Zambroni de; SANTOS, P. A study of electricity price volatility for the brazilian energy market. In: . [S.l.: s.n.], 2008. p. 1 – 6. ISBN 978-1-4244-1743-8. Citado na página 33.

- LIMA, F. G.; KIMURA, H.; NETO, A. A.; PERERA, L. C. J. Previsão de preços de commodities com modelos arima-garch e redes neurais com ondaletas: velhas tecnologias - novos resultados. **Revista de Administração**, v. 45, n. 2, p. 188–202, 2010. Citado na página 32.
- LONGO, M.; ZANINELLI, D.; SIANO, P.; PICCOLO, A. Evaluating innovative fcn networks for energy prices' forecasting. In: **2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)**. [S.l.: s.n.], 2016. p. 315–320. Citado na página 34.
- MELO, J.; ALVES, A.; BRIGATTO, G.; NEGRETE, L. Mercado de energia elétrica: Volatilidade do preço spot e risco da geração hidrelétrica com despacho centralizado. In: . [S.l.: s.n.], 2015. Citado na página 25.
- MINISTÉRIO DE MINAS E ENERGIA, EMPRESA DE PESQUISA ENERGÉTICA. **Plano Decenal de Expansão de Energia 2029**. Brasília, 2019. 391 p. Citado na página 25.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. 2º. ed. [S.l.]: Blucher, 2006. ISBN 978-85-212-0389-6. Citado nas páginas 40 e 41.
- NETO, A. A. **Mercado Financeiro**. 14º. ed. [S.l.]: Atlas, 2018. 268 p. ISBN 9788597017793. Citado na página 61.
- QUEIROZ, F. C. B. P.; ANDRADE, H. R. H. D. V. P.; QUEIROZ, J. V.; MACÊDO, D. M. D. PREVISÃO E SÉRIES TEMPORAIS PARA TOMADA DE DECISÃO EMPRESARIAL EM UMA INDÚSTRIA MOVELEIRA DA REGIÃO DE CRICIÚMA–SC. **Revista Catarinense da Ciência Contábil**, v. 11, p. 26 – 42, 07 2012. ISSN 1808-3781. Citado na página 31.
- REIS, A. J. R.; SILVA, A. P. A. d. Aplicação da transformada wavelet discreta na previsão de carga a curto prazo via redes neurais. **Sba: Controle Automação Sociedade Brasileira de Automatica**, sciELO, v. 15, p. 101 – 108, 03 2004. ISSN 0103-1759. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-17592004000100013&nrm=iso>. Citado na página 33.
- RODRIGUES, A. L. **Redes neurais artificiais aplicadas na previsão de preços do mercado spot de energia elétrica**. Dissertação (Mestrado) — Universidade de São Paulo, São Paulo, 2009. Citado na página 34.
- SAMOHYL, O. F. C. W. e Elisa Henning e Graciela Moro e R. Aplicação de um modelo sarima na previsão de vendas de motocicletas. **Exacta**, v. 11, n. 1, p. 77–88, 2013. ISSN 1983-9308. Disponível em: <<https://periodicos.uninove.br/exacta/article/view/3897>>. Citado na página 44.
- SHIRI, A.; AFSHAR, M.; RAHIMI-KIAN, A.; MAHAM, B. Electricity price forecasting using support vector machines by considering oil and natural gas price impacts. In: IEEE INTERNATIONAL CONFERENCE ON SMART ENERGY GRID ENGINEERING (SEGE), 2015. Oshawa, ON, Canada, 2015. Citado na página 33.
- SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais: para engenharia e ciências aplicadas**. 2º. ed. [S.l.]: Artliber, 2016. ISBN 8588098873. Citado na página 53.
- THIELE, C. C.; ADAMI, A. G. Previsão de séries temporais financeiras: modelo baseado em redes neurais artificiais. **Revista Brasileira de Computação Aplicada**, v. 8, p. 113 – 128, 07 2016. ISSN 2176-6649. Citado na página 32.

TOLMASQUIM, M. T. **Novo Modelo do Setor Elétrico Brasileiro**. 2^o. ed. [S.l.]: Synergia, 2015. ISBN 978-85-68483-06-0. Citado na página 24.

WONG, W. C.; CHEE, E.; LI, J.; WANG, X. Recurrent neural network-based model predictive control for continuous pharmaceutical manufacturing. **Mathematics**, v. 6, n. 11, 2018. ISSN 2227-7390. Disponível em: <<https://www.mdpi.com/2227-7390/6/11/242>>. Citado na página 54.

