

**UFRRJ**  
**INSTITUTO DE CIÊNCIAS EXATAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM**  
**MODELAGEM MATEMÁTICA E COMPUTACIONAL**

**DISSERTAÇÃO**

**Um estudo sobre a previsão do consumo de energia  
elétrica considerando as medidas de eficiência  
energética: aplicado à Universidade Federal Rural  
do Rio de Janeiro**

**Douglas Bortolassi Filgueiras**

**2022**



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO  
INSTITUTO DE CIÊNCIAS EXATAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM  
MATEMÁTICA E COMPUTACIONAL**

**UM ESTUDO SOBRE A PREVISÃO DO CONSUMO DE ENERGIA  
ELÉTRICA CONSIDERANDO AS MEDIDAS DE EFICIÊNCIA  
ENERGÉTICA: APLICADO À UNIVERSIDADE FEDERAL RURAL DO  
RIO DE JANEIRO**

**DOUGLAS BORTOLASSI FILGUEIRAS**

*Sob orientação de*  
**Felipe Leite Coelho da Silva**

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre** no Programa de Pós-Graduação em Modelagem Matemática e Computacional, Área de Concentração em Modelagem Matemática e Computacional.

Seropédica, RJ, Brasil  
Maio de 2022

Universidade Federal Rural do Rio de Janeiro  
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada  
com os dados fornecidos pelo(a) autor(a)

F475e Filgueiras, Douglas Bortolassi, 28/08/1992-  
Um estudo sobre a previsão do consumo de energia elétrica considerando as medidas de eficiência energética: aplicado à Universidade Federal Rural do Rio de Janeiro / Douglas Bortolassi Filgueiras. - Seropédica, 2022.  
61 f.

Orientador: Felipe Leite Coelho da Silva.  
Dissertação(Mestrado). -- Universidade Federal Rural do Rio de Janeiro, Programa de Pós-Graduação em Modelagem Matemática e Computacional, 2022.

1. Consumo de Energia Elétrica. 2. Séries Temporais. 3. Análise de Cluster. 4. Previsão. 5. Medidas de Eficiência Energética. I. da Silva, Felipe Leite Coelho , 28/05/1981-, orient. II Universidade Federal Rural do Rio de Janeiro. Programa de Pós Graduação em Modelagem Matemática e Computacional III. Título.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO  
INSTITUTO DE CIÊNCIAS EXATAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL**

**DOUGLAS BORTOLASSI FILGUEIRAS**

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre** no Programa de Pós-Graduação em Modelagem Matemática e Computacional, Área de Concentração em Modelagem Matemática e Computacional.

DISSERTAÇÃO APROVADA EM 30/05/2022.

---

Felipe Leite Coelho da Silva D. Sc. UFRRJ  
(Presidente)

---

Carlos Andrés Reyna Vera-Tudela D. Sc. UFRRJ

---

Fernando Luiz Cyrino Oliveira D. Sc. PUC-Rio

## Agradecimentos

A vida é repleta de escolhas e caminhos, agradeço a Deus por ter me possibilitado trilhar minhas escolhas e ter colocado pessoas na minha vida que ensinaram como seguir em frente e superar as adversidades. Portanto, a estas pessoas gostaria de expressar minha gratidão.

Ao Prof. Felipe Leite Coelho da Silva, o meu agradecimento pela disponibilidade e dedicação ao longo da construção do presente trabalho, estando sempre disposto e pronto a ajudar, retirando dúvidas e os obstáculos encontrados.

À minha família, aos meus pais Luis Vanderlei e Terezinha e minha irmã Mariana, que sempre dispuseram apoio em todos os momentos da minha vida, sempre respaldando com amor e carinho.

Aos professores do Programa de Pós-Graduação em Modelagem Matemática e Computacional da UFRRJ (PPGMMC-UFRRJ) pelo incentivo e conhecimento disponibilizado.

Aos colegas de labor e amigos da COPEA/UFRRJ pelo incentivo.

A todos os amigos que estiveram comigo neste período, pelo interesse em ouvir sobre este trabalho e motivação oferecida.

A todos que fizeram parte da minha formação profissional e meu crescimento pessoal, o meu mais sincero muito obrigado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

*"Uma mente que se abre a uma nova  
ideia jamais volta ao seu tamanho  
original." Albert Einstein*

BORTOLASSI FILGUEIRAS, Douglas. **UM ESTUDO SOBRE A PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA CONSIDERANDO AS MEDIDAS DE EFICIÊNCIA ENERGÉTICA: APLICADO À UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO**. 2022. 62f. Dissertação (Mestrado em Modelagem Matemática e Computacional). Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2022.

A previsão do consumo de energia elétrica é muito importante para o planejamento energético de um país, empresa ou região. O interesse pelas projeções do consumo de energia elétrica está relacionado, em geral, ao impacto financeiro que a distribuição de energia pode gerar, podendo causar imensos prejuízos. Neste trabalho, propomos uma metodologia utilizando a abordagem bottom-up por meio de modelos de séries temporais e análise de cluster para obter a previsão do consumo de energia elétrica. As medidas de eficiência energética foram inseridas na metodologia para avaliar a economia de energia elétrica. Em particular, esta metodologia foi aplicada aos dados de consumo de energia elétrica da Universidade Federal Rural do Rio de Janeiro (UFRRJ). Os resultados mostram que a metodologia apresentou um erro percentual absoluto médio de aproximadamente 1%. Além disso, essa metodologia demonstrou um grande potencial para avaliar a implementação de medidas de eficiência energética.

**Palavras-chave:** Consumo de Energia Elétrica, Séries Temporais, Análise de Cluster, Previsão, Medidas de Eficiência Energética.

## ABSTRACT

BORTOLASSI FILGUEIRAS, Douglas. **Titulo em Ingles** . 2022. 62p. Dissertation (Master in Mathematical and Computational Modeling). Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2022.

Electricity consumption forecasting is very important for the energy planning of a country, company or region. The interest in electricity consumption projections is related, in general, to the financial impact that energy distribution can generate, which can cause immense losses. In this work, we propose a methodology using the bottom-up approach through time series models and cluster analysis to obtain the prediction of electricity consumption. Energy efficiency measures were inserted into the methodology to assess electricity savings. In particular, this methodology was applied to electricity consumption data from the Federal Rural University of Rio de Janeiro (UFRRJ). The results show that the methodology presented an average absolute percentage error of approximately 1%. In addition, this methodology has shown great potential for evaluating the implementation of energy efficiency measures.

**Keywords:** Electricity Consumption, Time Series, Cluster Analysis, Forecast, Energy Efficiency Measures.



<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
1.1	Motivação	2
1.2	Objetivo	2
1.3	Organização da Dissertação	3
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>4</b>
<b>3</b>	<b>REVISÃO TEÓRICA</b>	<b>7</b>
<b>3.1</b>	<b>Análise de Cluster</b>	<b>7</b>
3.1.1	Medidas de similaridade ou dissimilaridade	7
3.1.1.1	Distância Euclidiana	8
3.1.1.2	Distância de Manhattan	8
3.1.1.3	Distância de Minkowski	8
3.1.2	Método de $k$ -médias ( $k$ -means)	9
<b>3.2</b>	<b>Séries Temporais</b>	<b>10</b>
<b>3.3</b>	<b>Modelos de Holt-Winters</b>	<b>11</b>
<b>3.4</b>	<b>Modelos de Box e Jenkins</b>	<b>13</b>
3.4.1	Função de Autocorrelação (FAC)	13
3.4.2	Função de Autocorrelação Parcial (FACP)	14
3.4.3	Modelo autorregressivo (AR(p))	14
3.4.3.1	Comportamento teórico da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP), modelo autorregressivo	14
3.4.4	Modelo de médias móveis (MA(q))	16
3.4.4.1	Comportamento teórico da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP), modelo de médias móveis	16
3.4.5	Modelo autorregressivo e de médias móveis (ARMA(p,q))	17
3.4.6	Modelo ARIMA não sazonal	17
3.4.7	Modelo ARIMA sazonal (SARIMA)	17
3.4.8	Estimação de parâmetros	18
<b>3.5</b>	<b>Modelos de Regressão Dinâmica</b>	<b>18</b>
<b>3.6</b>	<b>Modelos Lineares Dinâmicos</b>	<b>20</b>
3.6.1	Passeio aleatório mais ruído	21
3.6.2	Modelo de crescimento linear	22

3.6.3	Modelo de crescimento linear com sazonalidade .....	23
3.6.4	Filtro de Kalman para os modelos lineares dinâmicos .....	23
<b>3.7</b>	<b>Redes Neurais Artificiais</b> .....	24
3.7.1	Arquitetura das redes Neurais .....	25
3.7.2	Modelo para uma rede <i>feedforward</i> .....	25
3.7.3	Modelo NNAR .....	26
<b>3.8</b>	<b>Métricas para avaliar os modelos de séries temporais</b> .....	27
3.8.1	Erro médio absoluto percentual (MAPE) .....	27
3.8.2	Erro quadrático médio (MSE) .....	27
<b>3.9</b>	<b>Modelo bottom-up</b> .....	27
<b>3.10</b>	<b>Medidas de Eficiência Energética (MEE)</b> .....	27
<b>4</b>	<b>MATERIAIS E MÉTODOS</b> .....	29
<b>4.1</b>	<b>Materiais</b> .....	29
<b>4.2</b>	<b>Metodologia</b> .....	29
4.2.1	Modelo proposto para implementação de MEE .....	30
<b>5</b>	<b>RESULTADOS</b> .....	33
<b>5.1</b>	<b>Análise Descritiva dos Setores</b> .....	33
<b>5.2</b>	<b>Análise de Cluster</b> .....	34
<b>5.3</b>	<b>Cenário 1</b> .....	34
5.3.1	Previsões e ajustes para o Cluster 1 .....	35
5.3.2	Previsões e ajustes para o Cluster 2 .....	36
5.3.3	Previsões e ajustes para o Cluster 3 .....	37
5.3.4	Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, modelo linear dinâmico, modelo de redes neurais artificiais e do modelo de regressão dinâmica.....	38
5.3.5	Previsão do consumo de energia elétrica total da UFRRJ utilizando a abordagem bottom-up. ....	39
<b>5.4</b>	<b>Cenário 2</b> .....	40
5.4.1	Previsões e ajustes para o Cluster 1 .....	40
5.4.2	Previsões e ajustes para o Cluster 2 .....	41
5.4.3	Previsões e ajustes para o Cluster 3 .....	42
5.4.4	Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, modelo linear dinâmico, modelo de redes neurais artificiais e do modelo de regressão dinâmica.....	44
5.4.5	Previsão do consumo de energia elétrica total da UFRRJ utilizando a abordagem bottom-up. ....	44
<b>5.5</b>	<b>Cenário 3</b> .....	45
5.5.1	Previsões e ajustes para o Cluster 1 .....	45
5.5.2	Previsões e ajustes para o Cluster 2 .....	46
5.5.3	Previsões e ajustes para o Cluster 3 .....	47
5.5.4	Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, modelo linear dinâmico, modelo de redes neurais artificiais e do modelo de regressão dinâmica.....	49
5.5.5	Previsão do consumo de energia elétrica total da UFRRJ utilizando a abordagem bottom-up. ....	49

<b>5.6</b>	<b>Implementação de medidas de eficiência energética durante o horizonte de previsão do cenário 3 .....</b>	<b>50</b>
<b>6</b>	<b>CONCLUSÃO .....</b>	<b>55</b>
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>57</b>
<b>8</b>	<b>ANEXO .....</b>	<b>62</b>

## Lista de Figuras

Figura 3.1 – Série temporal do consumo de energia elétrica da Prefeitura Universitária da UFRRJ.	11
Figura 3.2 – Função de autocorrelação.	15
Figura 3.3 – Função de autocorrelação parcial.	16
Figura 3.4 – Diagrama de uma rede neural.	25
Figura 4.1 – Série temporal do consumo de energia elétrica do setor 17 da UFRRJ.	30
Figura 4.2 – Fluxograma com as etapas da metodologia.	31
Figura 5.1 – Clusters obtidos pelo método k-médias.	34
Figura 5.2 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018 para séries temporais do Cluster 1.	35
Figura 5.3 – Previsões dos modelos para o consumo de energia elétrica para o período de janeiro de 2019 a dezembro de 2019 para séries temporais do Cluster 1.	36
Figura 5.4 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018 para séries temporais do Cluster 2.	36
Figura 5.5 – Previsões dos modelos para o consumo de energia elétrica para o período de janeiro de 2019 a dezembro de 2019 para séries temporais do Cluster 2.	37
Figura 5.6 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018 para séries temporais do Cluster 3.	37
Figura 5.7 – Previsões dos modelos para o consumo de energia elétrica para o período de janeiro de 2019 a dezembro de 2019 para séries temporais do Cluster 3.	38
Figura 5.8 – Previsão do consumo de energia elétrica no período de janeiro de 2019 a dezembro de 2019 para a UFRRJ.	40
Figura 5.9 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020 para séries temporais do Cluster 1.	41
Figura 5.10 – Previsões dos modelos para o consumo de energia elétrica para o período de abril de 2020 a maio de 2021 para séries temporais do Cluster 1.	41
Figura 5.11 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020 para séries temporais do Cluster 2.	42
Figura 5.12 – Previsões dos modelos para o consumo de energia elétrica para o período de abril de 2020 a maio de 2021 para séries temporais do Cluster 2.	42
Figura 5.13 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020 para séries temporais do Cluster 3.	43
Figura 5.14 – Previsões dos modelos para o consumo de energia elétrica para o período de abril de 2020 a maio de 2021 para séries temporais do Cluster 3.	43

Figura 5.15 – Previsão do consumo de energia elétrica no período de abril de 2020 a maio de 2021 para a UFRRJ. ....	45
Figura 5.16 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021 para séries temporais do Cluster 1. ....	46
Figura 5.17 – Previsões dos modelos para o consumo de energia elétrica para o período de junho de 2021 a maio de 2026 para séries temporais do Cluster 1. ....	46
Figura 5.18 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021 para séries temporais do Cluster 2. ....	47
Figura 5.19 – Previsões dos modelos para o consumo de energia elétrica para o período de junho de 2021 a maio de 2026 para séries temporais do Cluster 2. ....	47
Figura 5.20 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021 para séries temporais do Cluster 3. ....	48
Figura 5.21 – Previsões dos modelos para o consumo de energia elétrica para o período de junho de 2021 a maio de 2026 para séries temporais do Cluster 3. ....	48
Figura 5.22 – Previsão do consumo de energia elétrica no período de junho de 2021 a maio de 2026 para a UFRRJ. ....	50
Figura 5.23 – Curvas de difusão das medidas de eficiência energética no período de junho de 2021 a maio de 2026 na UFRRJ. ....	51
Figura 5.24 – Previsão do consumo de energia elétrica no período de junho de 2021 a maio de 2026 para os diferentes cenários de difusão de medidas de eficiência energética na UFRRJ. ....	51
Figura 5.25 – Economia anual de energia elétrica em kWh para os diferentes cenários de difusão de medidas de eficiência energética para o horizonte de previsão de junho de 2021 a maio de 2026. ....	52

## Lista de Tabelas

Tabela 5.1 – Análise Descritiva dos setores. ....	33
Tabela 5.2 – Análise Descritiva dos Clusters. ....	34
Tabela 5.3 – Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, linear dinâmico, de redes neurais artificiais e o de regressão dinâmica, através da métrica MAPE para o Cenário 1. ....	39
Tabela 5.4 – Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, linear dinâmico, de redes neurais artificiais e o de regressão dinâmica, através da métrica MAPE para o Cenário 2. ....	44
Tabela 5.5 – Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, linear dinâmico, de redes neurais artificiais e o de regressão dinâmica, através da métrica MAPE para o Cenário 3. ....	49
Tabela 5.6 – Economia anual de energia elétrica em kWh para os diferentes cenários de difusão de medidas de eficiência energética no horizonte de previsão de junho de 2021 a maio de 2026. ....	52
5.7 Previsão do consumo de energia elétrica no período de junho de 2021 a maio de 2026 para os diferentes cenários de difusão de medidas de eficiência energética na UFRRJ. ....	52

A projeção do consumo de energia elétrica é importante para o planejamento de empresas e países desenvolvidos ou em desenvolvimento [1]. No Brasil, esse estudo é elaborado no Plano Decenal de Energia (PDE) realizado pela Empresa de Pesquisa Energética (EPE) [2]. Esse estudo propicia o planejamento da expansão dos sistemas energéticos que é importante para o desenvolvimento do país.

A partir das projeções do consumo de energia elétrica de um país, região, indústria ou até mesmo de um setor da economia é possível implementar medidas que podem suavizar o aumento ou reduzir o nível do consumo de energia elétrica. Uma das ações adotadas são denominadas medidas de eficiência energética (MEEs) [3].

Neste trabalho, será realizado a previsão do consumo de energia elétrica da Universidade Federal Rural do Rio de Janeiro (UFRRJ), considerando cenários de implementação de medidas de eficiência energética. Os dados de consumo de energia elétrica da UFRRJ são mensais e estão organizados em uma estrutura hierárquica, melhor dizendo, as séries temporais de consumo de energia elétrica da UFRRJ são hierárquicas. A estrutura hierárquica é composta pela agregação das séries temporais do consumo de energia elétrica dos setores da UFRRJ (níveis mais baixos) para formar a série temporal do consumo da UFRRJ (nível mais alto da hierarquia). Logo é necessário utilizar uma abordagem para séries temporais hierárquicas.

As abordagens bottom-up e top-down são estratégias de processamento de informação e ordenação do conhecimento, usadas em várias áreas de conhecimento [4]. No estudo de séries temporais, essas abordagens são utilizadas quando uma variável é organizada em uma estrutura hierárquica [5, 6]. Na abordagem top-down as previsões são geradas para série temporal completamente agregada, isto é, as variáveis de interesse são tratadas como um todo, fornecendo resultados no nível mais alto da estrutura hierárquica. Já na abordagem bottom-up, a estrutura hierárquica é caracterizada pelo agrupamento de séries temporais de níveis mais baixos para formar séries em outro nível da hierarquia, ou seja, as variáveis de interesse são desagregadas em componentes, essas componentes recebem um tratamento, e os resultados gerados pelas componentes são acumulados para compor as variáveis de interesse.

Quando faz-se a comparação entre esses dois tipos de abordagens, nota-se que a abordagem top-down utiliza uma quantidade menor de dados para chegar aos resultados. Portanto, na abordagem top-down há maior dificuldade na interpretação da influência de alguns agentes externos nos resultados obtidos, isto é, por exemplo avaliar o efeito de medidas de eficiência [3]. No entanto, na abordagem bottom-up tem-se uma melhor interpretação da influência de fatores externos nos resultados, mas em contrapartida, é necessário uma quantidade maior de dados. Portanto, a quantidade de informações necessárias para uma previsão hierárquica usando

abordagem bottom-up aumenta o nível de detalhamento da estrutura hierárquica e, consequentemente, aumenta o grau de complexidade da modelagem [7, 8].

Recentemente, a abordagem bottom-up vem sendo utilizada para avaliar os efeitos de medidas de eficiência no consumo de energia elétrica em diferentes setores do sistema energético, como será apresentado no Capítulo 2. Dessa forma, será utilizada no presente trabalho uma abordagem bottom-up para obter a previsão do consumo de energia elétrica, devido à estrutura hierárquica dos dados e a implementação de medidas de eficiência energética. Assim, a fim de prever o consumo de energia elétrica das séries temporais dos níveis mais baixos, propõe-se diferentes modelos de previsão, tais como: os modelos de Box e Jenkins, o método de Holt-Winters, os modelos lineares dinâmicos, os modelos de redes neurais artificiais e os modelos de regressão dinâmica. Além disso, será introduzido o algoritmo não supervisionado de K-médias para classificar os setores.

A metodologia utilizada neste trabalho diferencia-se das metodologias encontradas na literatura por utilizar a análise de cluster (método de K-médias) nas séries temporais dos níveis mais baixos, conseguindo realizar uma modelagem que se adéque aos grupos de séries temporais dos setores que compõem o consumo da UFRRJ. Todavia, essa metodologia pode ser empregada nas diferentes áreas do conhecimento, dada a ampla presença das séries temporais hierárquicas em diversificados seguimentos.

## 1.1 Motivação

O planejamento do consumo de eletricidade é importante para o desenvolvimento de um setor. Em particular, a Universidade Federal Rural do Rio de Janeiro não possui um estudo de planejamento do consumo de energia elétrica. As medidas que vem sendo implementadas com o intuito de minimizar o consumo de energia, como a implementação de medidas de eficiência energética, vêm sendo realizadas de maneiras isoladas e sem a possibilidade de avaliar o impacto futuro de tais medidas.

A importância de fazer a previsão do consumo de energia elétrica é notória, tendo em vista a magnitude dos gastos mensais em energia elétrica da Universidade Federal Rural do Rio de Janeiro, que é em torno dos R\$ 700.000,00, no período em estudo de janeiro de 2014 a maio de 2021. Neste contexto, há necessidade de realizar um estudo de previsão do consumo de energia elétrica considerando medidas de eficiência energética aplicada à UFRRJ. O presente estudo propiciará, com os seus resultados, cenários de economia de energia, devido às medidas de eficiência energética, o que dará embasamento para os gestores gerenciar o consumo de energia elétrica na Universidade. Ademais, pode ser utilizado para a captação de recursos destinados à implementação de medidas de eficiência energética para a instituição.

## 1.2 Objetivo

O objetivo deste trabalho é realizar a previsão do consumo de energia elétrica da UFRRJ e avaliar os efeitos de medidas de eficiência energética. Como objetivos específicos, pretende-se:

- Realizar o mapeamento dos setores da universidade;
- Analisar o consumo de energia elétrica dos setores;
- Classificar os setores em relação ao consumo de energia elétrica através da clusterização;



- Ajustar os modelos de séries temporais aos dados de consumo de energia elétrica;
- Obter a previsão do consumo de energia elétrica;
- Utilizar métricas para avaliar o desempenho dos modelos de previsão;
- Propor uma abordagem bottom-up para o setor analisado;
- Analisar e introduzir as medidas de eficiência energética.

### **1.3 Organização da Dissertação**

No Capítulo 2, são apresentadas algumas bibliografias relacionadas ao trabalho. Neste capítulo, foi realizada uma revisão de literatura sobre as abordagens utilizadas para a previsão do consumo de energia através das metodologias bottom-up e top-down. No Capítulo 3 são apresentadas as revisões de alguns assuntos já difundidos na literatura e utilizados na construção da metodologia do presente trabalho. O Capítulo 4 elucida a metodologia utilizada no presente trabalho, mostrando como os assuntos revisados no Capítulo 3 são utilizados para a construção desta metodologia. No Capítulo 5, são apresentados os resultados da aplicação da metodologia proposta. E, finalmente, no Capítulo 6 são apresentadas as conclusões deste trabalho, e apontando os possíveis trabalhos futuros.

## Revisão Bibliográfica

Neste Capítulo, será realizada uma revisão de alguns estudos envolvendo as abordagens bottom-up e top-down usadas para previsões do consumo de energia. Será destacado especialmente os efeitos das medidas de eficiência energética dos estudos utilizando a metodologia bottom-up [3].

A abordagem bottom-up vem sendo usada na análise de sistemas de energia desde 1980 [9, 10]. Herbst *et al.* [8] avaliaram a demanda e o fornecimento de energia futura utilizando modelos de energia bottom-up e top-down. Nesse estudo ressaltaram que um modelo energético bottom-up possui maior detalhamento tecnológico dos dados quando comparado a um modelo top-down. Também foi ressaltado que a abordagem híbrida desses modelos vem sendo empregada para realizar o planejamento de sistemas energéticos.

Os trabalhos de Kinney [12], Collins [13], Dunn *et al.* [14], Zellner e Tobias [5], e Dangerfield e Morris [15] mostraram que ao utilizar a abordagem bottom-up as previsões são mais precisas do que as previsões obtidas utilizando a abordagem top-down. Essa característica é devido principalmente ao detalhamento da estrutura dos dados utilizados na abordagem bottom-up.

Com o intuito de melhorar o desempenho da previsão das séries temporais hierárquicas, Fliedner [11] apresentou um detalhamento sobre as informações necessárias para realizar esse tipo de previsão, utilizando a abordagem bottom-up.

Jacobsen [16], Koopmans e Velde [17], Swan e Ugursal [18], Hall e Buckley [19], modelaram o consumo de energia realizando previsões de longo prazo em diferentes países utilizando as abordagens bottom-up, top-down e uma combinação das duas metodologias. Nesses trabalhos os autores afirmam que os modelos bottom-up são mais adequados para agregar novas tecnologias ao modelo. Um modelo bottom-up foi utilizado para obter o consumo de energia do setor residencial na Argélia até 2040 [20], obtendo resultados satisfatório para o efeito de medidas de eficiência.

Um modelo determinístico de programação linear inteira mista bottom-up foi apresentado por Georgiou [21], para o planejamento de energia a longo prazo na Grécia. Na China o planejamento de rede elétrica inter-regional foi modelado até 2030 usando abordagem bottom-up por Yi *et al.* [22]. Lee e Huh [23] usaram o modelo bottom-up para previsão de longo prazo do fornecimento de energia nova e renovável na Coreia do Sul. Berntsen e Trutnevte [24] apresentaram um modelo de energia bottom-up denominado de Modelo de Geração Alternativas (MGA) para analisar os cenários suíços de suprimento de eletricidade de longo prazo. MGA é uma técnica que pode gerar diferentes cenários de energia [25, 26, 27, 28, 29].

Outras obras sobre a previsão de consumo de energia elétrica de longo prazo que foram

desenvolvidos usando diferentes metodologias podem ser vistas em Koltsaklis *et al.* [30] e Kaboli *et al.* [31]. O primeiro estudo utilizou o método de Monte-Carlo e o segundo propôs uma abordagem ACS (busca cooperativa artificial).

Desde a crise do petróleo na década de 70, pode-se notar uma mudança na concepção sobre eficiência energética, uma vez que se intensificaram os estudos almejando a melhoria da eficiência energética no curto e médio prazo, como foi destacado por Beer *et al.* [32]. Esse estudo foi aplicado a indústria de papel e celulose com base nos países da Organização de Cooperação e Desenvolvimento Econômico (OECD), esses autores descreveram um método para detectar e caracterizar as tecnologias que podem melhorar a eficiência energética de longo prazo. Em torno de 570 medidas de eficiência energética aplicadas na indústria, praticadas pelos governos da Agência Internacional de Energia (IEA), China, Índia, México, Rússia e África do Sul, foram apresentadas em uma revisão desenvolvida por Tanaka [34]. Worrell e Price [33] realizaram um estudo sobre a existência de potenciais para melhoria da eficiência energética na indústria.

Fleiter *et al.* [35] apresentaram uma revisão dos modelos bottom-up de demanda de energia para o setor industrial, a fim de mostrar como esses modelos consideram as barreiras à implementação de novas medidas de eficiência energética.

Giraldo e Hyman [36], Giraldo e Hyman [37] em suas obras aplicadas à indústria de papel e celulose dos Estados Unidos da América (EUA), apresentaram um detalhamento dos modelos de energia de uso final. No estudo de Farla *et al.* [38] foi apresentado um método que compara o desenvolvimento da eficiência energética na indústria de transformação. Esse método fundamenta-se na utilização de dados de produção física como uma medida do crescimento da atividade para a indústria de transformação. Aplicou-se essa metodologia à indústria de papel e celulose de oito países da OECD. Dessa forma, com a aplicação da metodologia, foi possível acompanhar a evolução da eficiência energética, individualmente, para o consumo de combustível e eletricidade.

Kramer *et al.* [39] realizaram um estudo sobre as medidas de eficiência energética que mostra as tendências, estrutura e características de consumo de energia da indústria de papel e celulose dos EUA, juntamente com descrições das principais tecnologias de processo utilizadas na indústria. Essas medidas de eficiência energética podem ser aplicadas na indústria de papel e celulose de todo mundo.

Fracaro *et al.* [40] apresentaram um estudo sobre o desenvolvimento do consumo de energia da indústria brasileira de celulose e papel por meio de uma análise de decomposição de energia e uma abordagem de índice de eficiência energética entre 1979 e 2009. Camiato *et al.* [41] avaliaram a eficiência energética dos setores industriais brasileiros (papel e celulose, alimentos e bebidas, química, mineração, não metálicos e metalúrgicos, têxtil) de 1996 à 2009 usando um método de programação matemática denominado Data Envelopment Analysis (DEA). Esses estudos não avaliaram o impacto das medidas de eficiência energética de longo prazo. Silva *et al.* [42] mostraram um estudo preliminar para o consumo de energia elétrica do setor brasileiro de celulose e papel até 2035 usando um modelo bottom-up. Este trabalho não apresentou os detalhes do modelo utilizado. Silva *et al.* [3, 43] realizaram também a previsão do consumo de energia elétrica do setor brasileiro de celulose e papel utilizando a metodologia bottom-up e avaliaram o efeito das medidas de eficiência energética. A previsão do consumo de energia elétrica para o setor industrial brasileiro foi realizada por Silva *et al.* [44] através dos modelos Holt–Winters, SARIMA, Modelo Linear Dinâmico e TBATS (Trigonometric Box–Cox transform, ARMA errors, Trend e Seasonal components).

No Brasil, a EPE [2] é responsável pelo planejamento energético e vem desenvolvendo a previsão de demanda de eletricidade de médio e longo prazo, utilizando ambas as abordagens

bottom-up e top-down. A agência governamental não mostra o nível de detalhes da estrutura hierárquica usada em seu modelo bottom-up. Pao e Tsai [45] mostraram a importância da modelagem e previsão do consumo de energia no Brasil. Eles adotaram uma abordagem de top-down para prever o consumo anual de energia para o período de 2008 à 2013, usando os dados históricos de 1980 à 2007.

A abordagem bottom-up também vem sendo utilizada em estudos no setor residencial. Maçaira *et al.* [46] realizaram a previsão do consumo de energia elétrica para o setor residencial brasileiro por região utilizando a abordagem bottom-up e considerando as medidas de eficiência energética. A projeção do consumo de energia elétrica para o setor residencial no Brasil utilizando o método de Pegels Exponential Smoothing e considerando o aumento da eficiência energética foi realizado por Maçaira *et al.* [47]. Cabral *et al.* [48] realizaram a comparação entre três modelos utilizados para realizar a previsão do consumo de energia elétrica no setor residencial brasileiro. Irffi *et al.* [49] realizaram a previsão da demanda por energia elétrica nas classes residencial, comercial e industrial na região Nordeste do Brasil usando OLS dinâmico e mudança de regime.

A necessidade de se implementar medidas de eficiência energética foi destacada por Rodrigues [50] que salientou a importância de implementar essas medidas para minimizar o crescente consumo de energia elétrica que traz impactos econômicos e ambientais.

Neste Capítulo foi apresentada uma perspectiva ampla de diferentes trabalhos usando as abordagens bottom-up e top-down na previsão do consumo de energia. Embora a abordagem bottom-up seja intensiva em dados, exigindo uma grande quantidade de dados com boa qualidade o que aumenta a sua complexidade, ela foi escolhida porque permite uma melhor interpretação da influência de alguns fatores nos resultados. Além disso, nesta revisão de literatura não foi encontrada nenhuma metodologia que se propõe neste trabalho. O presente trabalho traz como novidade em sua metodologia o uso da análise de cluster concomitantemente com o uso de modelos de séries temporais. Isto possibilita realizar a previsão de uma quantidade reduzida de séries temporais agregadas, trazendo um ganho de tempo na aplicação dos modelos de séries temporais, pois a análise e previsão é aplicada em uma quantidade reduzida de séries temporais. As etapas desta metodologia serão apresentadas nos próximos capítulos.

Ademais, não foi encontrado na literatura nenhum trabalho de previsão do consumo de energia elétrica para um subsetor específico do setor público.

Neste capítulo serão apresentadas as revisões de alguns assuntos já difundidos na literatura e utilizados neste trabalho. Estes assuntos foram empregados na construção da metodologia desenvolvida no presente trabalho, sendo os percussores dos resultados que vão ser apresentados nos próximos capítulos.

### 3.1 Análise de Cluster

A análise de cluster nos permite encontrar em um certo conjunto de dados com uma estrutura de agrupamento os elementos da amostra, ou da população, que possuem características (valores das variáveis) semelhantes. Assim, permite avaliar a dimensionalidade, identificar valores discrepantes (outliers) e sugerir hipóteses acerca da estrutura de relações [51]. Para realizar a análise de agrupamento deve-se considerar três características importantes, sendo elas:

- A maneira de medir a similaridade (ou dissimilaridade) dos dados;
- O processo de formar os agrupamentos;
- O meio de decidir quantos grupos formar.

#### 3.1.1 Medidas de similaridade ou dissimilaridade

A análise de agrupamento (análise de cluster) busca identificar e agrupar as observações em grupos. Muitas técnicas usam um índice de similaridade ou proximidade entre cada par de observações. Uma medida de proximidade é a distância entre duas observações. Nesse caso, como a distância aumenta à medida que duas unidades se distanciam, a distância é na verdade uma medida de dissimilaridade. Diversos algoritmos de clusterização são aplicados em dados organizados em uma matriz de dissimilaridade, onde o elemento da coluna  $j$  e linha  $i$  da matriz é o número  $d(i; j)$ , que representa a distância entre os objetos  $i$  e  $j$ . Considere a matriz de dissimilaridade  $n \times n$  abaixo, com  $n$  objetos [51].

$$\begin{bmatrix} d(1,1) & d(1,2) & \cdots & d(1,n) \\ d(2,1) & d(2,2) & \cdots & d(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ d(n,1) & d(n,2) & \cdots & d(n,n) \end{bmatrix} \quad (3.1)$$

onde  $d(1,1) = d(2,2) = \cdots = d(n,n) = 0$ .

A função  $d$  é uma distância quando as seguintes condições necessárias e suficientes são satisfeitas, para quaisquer objetos  $i, j, k$ :

- I.  $d(i, j) \geq 0$ ;
- II.  $d(i, i) = 0$ ;
- III.  $d(i, j) = d(j, i)$  (simetria);
- IV.  $d(i, j) \leq d(i, k) + d(k, j)$  (desigualdade triangular).

A matriz de dissimilaridade pode ser obtida usando a distância Euclidiana, a distância Manhattan e a distância de Minkowski.

### 3.1.1.1 Distância Euclidiana

A distância Euclidiana é a distância geométrica no espaço multidimensional [51]. Para duas observações  $p$ -dimensionais,  $x^T = (x_1, x_2, \dots, x_p)$  e  $y^T = (y_1, y_2, \dots, y_p)$ , a distância Euclidiana é dada por:

$$\begin{aligned} d(x, y) &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} \\ &= \sqrt{(x - y)^T (x - y)} \end{aligned} \quad (3.2)$$

### 3.1.1.2 Distância de Manhattan

A distância de Manhattan (“City Block” ou “Geometria do Táxi”) é a soma das diferenças absolutas de suas coordenadas [51]. Para duas observações  $p$ -dimensionais,  $x^T = (x_1, x_2, \dots, x_p)$  e  $y^T = (y_1, y_2, \dots, y_p)$ , a distância de Manhattan é dada por:

$$d(x, y) = \sum_{i=1}^p |x_i - y_i| \quad (3.3)$$

### 3.1.1.3 Distância de Minkowski

A distância de Minkowski pode ser considerada como uma generalização de ambas as distâncias, Euclidiana e Manhattan [51]. Para duas observações  $p$ -dimensionais,  $x^T = (x_1, x_2, \dots, x_p)$  e  $y^T = (y_1, y_2, \dots, y_p)$ , a distância Minkowski é dada por:

$$d(x, y) = \left[ \sum_{i=1}^p |x_i - y_i|^m \right]^{1/m} \quad (3.4)$$

Na equação (3.4) para  $m = 1$  temos a distância de Manhattan e para  $m = 2$  a medida Euclidiana.

Embasado no conceito de distância tem-se o método K-médias que possui o objetivo de encontrar a similaridade entre os dados e agrupá-los.

### 3.1.2 Método de $k$ -médias ( $k$ -means)

O método de  $k$ -médias é uma heurística de agrupamento não hierárquico. Esse método é um dos algoritmos de aprendizado de máquina não supervisionado atualmente usado para particionar um determinado conjunto de dados em  $K$  grupos (ou seja,  $K$  clusters), em que  $K$  representa o número de grupos pré-estabelecidos pelo analista [51].

Este método classifica os objetos em vários grupos (ou seja, clusters), de modo que os objetos dentro do mesmo cluster sejam tão semelhantes quanto possível (ou seja, alta similaridade intraclasse), enquanto os objetos de diferentes clusters são tão diferentes quanto possível (ou seja, baixa inter-similaridade de classe). No agrupamento  $k$ -médias, cada cluster é representado por seu centro (ou seja, centróide) que corresponde à média dos pontos atribuídos ao cluster.

O método de  $k$ -médias busca distribuir os  $n$  indivíduos em um conjunto de dados multivariados em  $K$  grupos ou clusters,  $(G_1; G_2; \dots; G_K)$ , em que  $G_k$  denota o conjunto de  $n_k$  indivíduos no  $k$ -ésimo grupo, e  $K$  o número de cluster deve ser especificado previamente (pode ser obtido algum critério numérico). O algoritmo mais usado de agrupamento de  $k$ -médias foi desenvolvido por Hartigan e Wong [52] que procura encontrar a partição dos  $n$  indivíduos em  $K$  grupos que minimiza a soma dos quadrados dentro do grupo sobre todas as variáveis, podendo ser escrita como a função objetivo abaixo:

$$\text{minimize}_{G_1, G_2, \dots, G_k} \left\{ \sum_{k=1}^K \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \right\}, \quad (3.5)$$

em que  $|C_k|$  é o número de observações do  $k$ -ésimo cluster e  $x_{ij}$  é a variável  $j$  da observação  $i$ . Em outros termos, a variação dentro do cluster para o  $k$ -ésimo cluster é a soma de todas as distâncias Euclidianas ao quadrado entre as observações no  $k$ -ésimo cluster, dividido pelo número total de observações no  $k$ -ésimo cluster.

Para resolver a função objetivo (3.5) é necessário utilizar um algoritmo que apresente uma metodologia para alocar as observações em  $K$  clusters. Essa tarefa é difícil de resolver com precisão, dado que existem  $k^n$  maneiras de particionar  $n$  observações em  $k$  clusters. Na literatura temos alguns algoritmos já difundidos, sendo eles: Hartigan and Wong [52], MacQueen [53], Lloyd [54] e Forgy [55]. O algoritmo Hartigan-Wong geralmente apresenta bons resultados quando comparado a qualquer um deles [51], as etapas deste método (algoritmo) são estabelecidas por:

1. Número de clusters ( $K$ ) a serem criados, especificado pelo analista;
2. Próximo passo, deve-se escolher aleatoriamente  $k$ -centróides  $c_1, c_2, \dots, c_k$ , para iniciar o processo de partição;
3. Em seguida, cada elemento do conjunto de dados é comparado com cada centróide inicial, através de uma medida de distância, usualmente a distância Euclidiana. Assim, o elemento é agrupado ao grupo cuja distância é menor.
4. Após aplicar o passo 2, recalcula-se os valores dos centróides para cada novo grupo obtido. Repete-se o passo 2 para os centróides dos novos grupos.
5. Repete-se os passos 3 e 4 até que todos elementos estejam alocados em seus grupos, isto é, quando nenhuma realocação for necessária.

Este algoritmo é rápido, comumente convergindo em poucas iterações para uma configuração estável, na qual nenhuma observação está designada para um cluster cujo centro não lhe seja o mais próximo [56]. Um eventual problema é que esta condição enfatiza a questão da homogeneidade e ignora a importante questão da boa separação dos clusters. Dessa forma, pode-se ter uma separação errônea dos conjuntos no caso de uma inicialização ruim dos centros, inicialização esta que é feita de forma arbitrária (aleatória) no início da execução. A qualidade dos resultados também podem ser afetada pela escolha do número de conjuntos feita pelo usuário. Um número muito pequeno de conjuntos pode causar a junção de dois clusters naturais, enquanto que um número grande demais pode fazer com que um cluster natural seja quebrado artificialmente em dois [56].

### 3.2 Séries Temporais

Uma série temporal é uma sequência de observações feitas sequencialmente ao longo do tempo. A frequência de cada observação/medição de uma série temporal pode ser em minutos, diária, semanal, mensal, anual etc, podendo variar dependendo do fenômeno observado. As séries temporais estão abundantemente presentes em diversas áreas, como na economia, epidemiologia, medicina (eletrocardiograma), engenharia, ciências naturais (especialmente geofísica e meteorologia) e ciências sociais [57]. Uma característica importante de uma série temporal é que, normalmente, as observações adjacentes são dependentes [57]. O fato das observações serem registradas em instantes distintos, sucessivos e serem dependentes entre si faz com que a análise de séries temporais se distingue das demais análises estatísticas. Conforme Morettin e Tolo [58], os modelos utilizados para descrever séries temporais são processos estocásticos, o que significa que são processos controlados por leis probabilísticas.

Ao trabalhar com séries temporais tem-se o interesse por analisar e modelar a dependência entre as observações. Os principais objetivos desse interesse são [58]:

- investigar o mecanismo de geração da série temporal;
- obter previsões de valores futuros;
- descrever o comportamento da série temporal;
- controle estatístico de processos, ou seja, uma medida de qualidade de um processo.

Considere uma série temporal constituída pelas seguintes observações  $\{Z_t, t = 1, 2, \dots, N\}$ . Uma série temporal pode ser decomposta em três componentes não-observáveis, como na equação abaixo:

$$Z_t = T_t + S_t + a_t \quad (3.6)$$

onde  $T_t$  é a componente de tendência,  $S_t$  é a componente de sazonalidade e  $a_t$  é uma componente aleatória ou ruído, de média zero e variância constante  $\sigma^2$ . A componente de tendência pode ser interpretada como um aumento ou diminuição gradual das observações ao longo de um período. Por outro lado, a componente sazonal mostra as flutuações ocorridas em sub-períodos da série, a exemplo de uma série com um período anual seus possíveis sub-períodos podem ser medidos mensalmente, trimestralmente ou diariamente [58].

A Figura 3.1 traz um exemplo de uma série temporal com frequência mensal. Nesta Figura tem-se a série temporal do consumo de energia elétrica em MWh, de outubro de 2017 a outubro de 2020, da Prefeitura Universitária da UFRRJ.



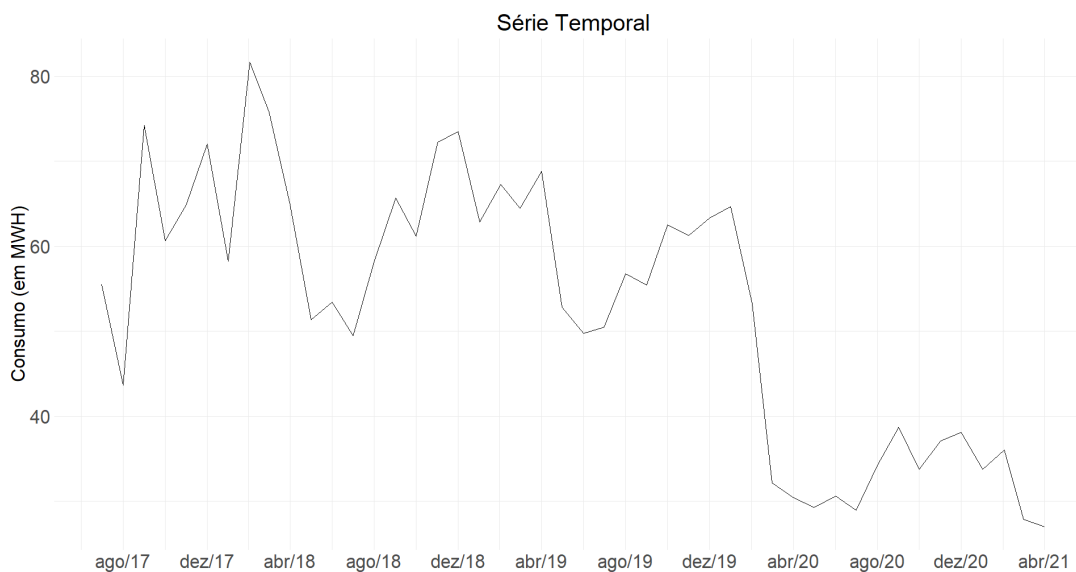


Figura 3.1 – Série temporal do consumo de energia elétrica da Prefeitura Universitária da UFRRJ.

Fonte: autor

### 3.3 Modelos de Holt-Winters

O método sazonal de Holt-Winters [61] é baseado em uma equação de previsão e três equações de suavização – uma para o nível, uma para a tendência, e uma para a sazonalidade. Existem duas variações do método de Holt-Winters, o método de Holt-Winters aditivo e o método de Holt-Winters multiplicativo. O método aditivo é utilizado quando as variações sazonais são aproximadamente constantes ao longo da série. E, o método multiplicativo é utilizado quando as variações sazonais estão mudando proporcionalmente ao nível da série.

A série temporal descrita pelo modelo de Holt-Winters aditivo por ser expressa por

$$Z_t = a_1 + b_1 t + S_t + \epsilon_t \quad (3.7)$$

onde  $\epsilon_t$  é o ruído e,  $a_1$ ,  $b_1$  e  $S_t$  são, respectivamente, os parâmetros de nível, tendência e sazonalidade.

A previsão dos valores futuros da série no método aditivo é dado por

$$\hat{Z}_t(h) = L_t + hT_t + S_{t+h-s}, \quad h = 1, 2, \dots, s \quad (3.8)$$

onde  $h$  é o horizonte de previsão e  $s$  é uma estação completa de sazonalidade o chamado período sazonal.

A componente de nível  $L_t$  é dada por

$$L_t = \alpha(Z_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad 0 < \alpha < 1 \quad (3.9)$$

onde  $\alpha$  é a constante de suavização da componente de nível.

A componente de tendência é dada por

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad 0 < \beta < 1 \quad (3.10)$$

onde  $\beta$  é a constante de suavização da componente de nível.

A componente de sazonalidade é dada por

$$S_t = \gamma(Z_t - L_t) + (1 - \gamma)S_{t-s}, \quad 0 < \gamma < 1 \quad (3.11)$$

onde  $\gamma$  é a constante de suavização da componente de sazonalidade.

Por outro lado, a série temporal descrita pelo modelo de Holt-Winters multiplicativo pode ser expressa por

$$Z_t = (a_1 + b_1 t)S_t + \epsilon_t \quad (3.12)$$

onde  $\epsilon_t$  é o ruído e,  $a_1$ ,  $b_1$  e  $S_t$  são, respectivamente, os parâmetros de nível, tendência e sazonalidade.

A previsão dos valores futuros da série no método multiplicativo é dado por

$$\hat{Z}_t(h) = (L_t + hT_t)S_{t-s+h}, \quad h = 1, 2, \dots, s \quad (3.13)$$

onde  $h$  é o horizonte de previsão e  $s$  é uma estação completa de sazonalidade.

A componente de nível  $L_t$  é dada por

$$L_t = \alpha \left( \frac{Z_t}{S_{t-s}} \right) + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad 0 < \alpha < 1 \quad (3.14)$$

onde  $\alpha$  é a constante de suavização da componente de nível.

A componente de tendência é dada por

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad 0 < \beta < 1 \quad (3.15)$$

onde  $\beta$  é a constante de suavização da componente de nível.

A componente de sazonalidade é dada por

$$S_t = \gamma \frac{Z_t}{L_t} + (1 - \gamma)S_{t-s}, \quad 0 < \gamma < 1 \quad (3.16)$$

onde  $\gamma$  é a constante de suavização da componente de sazonalidade.

Os métodos de Holt-Winters aditivos e multiplicativos permitem utilizar o amortecimento na tendência. Desta forma, são conhecidos como método de Holt-Winters com damped [62]. Um método que frequentemente apresenta previsões precisas e robustas para dados sazonais é o método Holt-Winters com uma tendência amortecida e sazonalidade multiplicativa.

A previsão dos valores futuros da série no método de Holt-Winters com damped é dado por

$$\hat{Z}_t(h) = (L_t + (\phi + \phi^2 + \dots + \phi^h)T_t)S_{t-s+h}, \quad h = 1, 2, \dots, s \quad (3.17)$$

onde  $h$  é o horizonte de previsão,  $s$  é uma estação completa de sazonalidade e  $\phi$  é a taxa de crescimento local.

A componente de nível  $L_t$  é dada por

$$L_t = \alpha \left( \frac{Z_t}{S_{t-s}} \right) + (1 - \alpha)(L_{t-1} + \phi T_{t-1}), \quad 0 < \alpha < 1 \quad (3.18)$$

onde  $\alpha$  é a constante de suavização da componente de nível.

A componente de tendência é dada por

$$T_t = \beta^* (L_t - L_{t-1}) + (1 - \beta^*)\phi T_{t-1}, \quad 0 < \beta < 1 \quad (3.19)$$

onde  $\beta$  é a constante de suavização da componente de nível.

A componente de sazonalidade é dada por

$$S_t = \gamma \frac{Z_t}{L_{t-1} + \phi T_{t-1}} + (1 - \gamma)S_{t-s}, \quad 0 < \gamma < 1 \quad (3.20)$$

onde  $\gamma$  é a constante de suavização da componente de sazonalidade.

A determinação dos parâmetros pode ser obtida minimizando a soma dos erros quadráticos de previsão para o valor do vetor  $(\alpha, \beta, \gamma)$ , isto é, obter o valor do vetor  $(\alpha, \beta, \gamma)$  que torne a soma dos erros quadráticos de previsão mínima. Maiores detalhes da metodologia de estimar os parâmetros são apontados por Granger e Newbold [61].

### 3.4 Modelos de Box e Jenkins

A metodologia de Box e Jenkins [59] é amplamente utilizada na análise de modelos paramétricos de séries temporais. Essa metodologia consiste em ajustar modelos auto-regressivos integrados de médias móveis, ARIMA(p, d, q), a um conjunto de dados, onde p é a ordem do modelo autorregressivo AR(p), q a ordem do modelo de médias móveis MA(q) e d o número de diferenças necessárias para tornar a série temporal estacionária. Para identificar os parâmetros q (referente à média-móvel), p (referente à parte autorregressiva) e d (referente ao número de diferenças) analisa-se o comportamento da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP). As seções seguintes trazem as definições de FAC e FACP, respectivamente.

#### 3.4.1 Função de Autocorrelação (FAC)

Seja um processo estacionário  $Z(t)$  com média  $\mu$  e variância  $\sigma^2$ , então a função de autocorrelação [57] pode ser expressa por

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{\gamma_j}{\sigma^2} \quad (3.21)$$

onde  $\gamma_j$  é a função de autocovariância.

Considerando as observações  $Z_1, Z_2, \dots, Z_N$ , a função de autocorrelação (FAC)  $\rho_j$  é estimado por

$$\rho_j \cong r_j, \quad (3.22)$$

$$r_j = \frac{c_j}{c_0}, j = 0, 1, \dots, N-1, \quad (3.23)$$

onde  $c_j$  é a estimativa da função de autocovariância  $\gamma_j$  definido por

$$c_j = E[Z_t - \bar{Z}][Z_{t+j} - \bar{Z}] = Cov[Z_t, Z_{t+j}] = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})], j = 0, 1, \dots, N-1 \quad (3.24)$$

sendo  $\bar{Z}$  a média amostral, dada por

$$\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t \quad (3.25)$$

### 3.4.2 Função de Autocorrelação Parcial (FACP)

A correlação entre  $Z_t$  e  $Z_{t-k}$ , removendo o efeito das variáveis  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$  para séries estacionárias, é chamada de autocorrelação parcial até a posição  $k$  e será denotada por  $\phi_{kk}$  [57] se  $Z_t$  é uma série normalmente distribuídas, ou seja,

$$\phi_{kk} = Cor(Z_t, Z_{t-k} | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}) \quad (3.26)$$

em que  $\phi_{kk}$  é o coeficiente de correlação da distribuição de  $Z_t, Z_{t-k}$  condicional a  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$ .

Seja um processo estacionário com FAC  $\rho_k$ , um método para obter a FACP é para um dado  $k$  mostrar que  $\phi_{kk}$  satisfaz as equações de Yule-Walker [60] dada por

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk}\rho_{j-p}, j = 1, 2, \dots, k. \quad (3.27)$$

Sendo assim, para uma série temporal observada, a função de autocorrelação parcial será estimada por

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j} \quad (3.28)$$

em que  $\phi_{kj} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j}$ , para  $j = 1, 2, \dots, k-1$ .

### 3.4.3 Modelo autorregressivo (AR(p))

O modelo autorregressivo [58] de ordem  $p$  (ou simplesmente  $AR(p)$ ) é definido por

$$Z_t = c + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (3.29)$$

onde  $c, \phi_1, \phi_2, \dots, \phi_p$  são os parâmetros,  $a_t$  é o ruído branco e o operador autorregressivo de ordem  $p$  é

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3.30)$$

onde  $B$  é o fator de atraso definido por

$$BZ_t = Z_{t-1}, \quad B^m Z_t = Z_{t-m} \quad (3.31)$$

#### 3.4.3.1 Comportamento teórico da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP), modelo autorregressivo

Quando o modelo autorregressivo é de ordem 1 (AR(1)) e  $0 < \phi < 1$ , a função de autocorrelação (FAC) apresenta decaimento exponencial (ou senoidal) e a função de autocorrelação parcial (FACP) é zero para  $k > 1$ . Já para o modelo autorregressivo de ordem 1 (AR(1)) e  $-1 < \phi < 0$ , a função de autocorrelação (FAC) apresenta decaimento para zero e a função de autocorrelação parcial (FACP) é zero para  $k > 1$ . Para o modelo autorregressivo de ordem  $p$  (AR(p)) a função de autocorrelação (FAC) apresenta decaimento para zero (exponenciais amortecidas e/ou senoide amortecida) e a função de autocorrelação parcial (FACP) é zero para  $k > p$ .

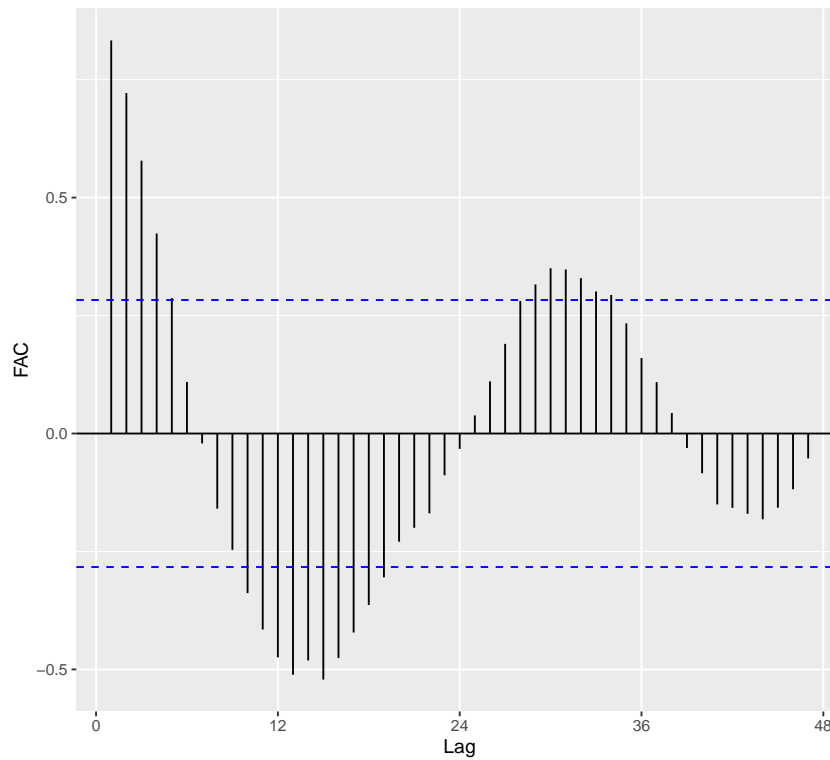


Figura 3.2 – Função de autocorrelação.  
 Fonte: autor

As Figuras 3.2 e 3.3 ilustram o comportamento da FAC e FACP, respectivamente, para um modelo AR(1) com  $0 < \phi < 1$ . Na Figura 3.2 a FAC apresenta decaimento senoidal. Todavia, na Figura 3.3 a FACP tende a zero para  $k > 1$ , sendo seu lag mais significativo em  $k = 1$ .

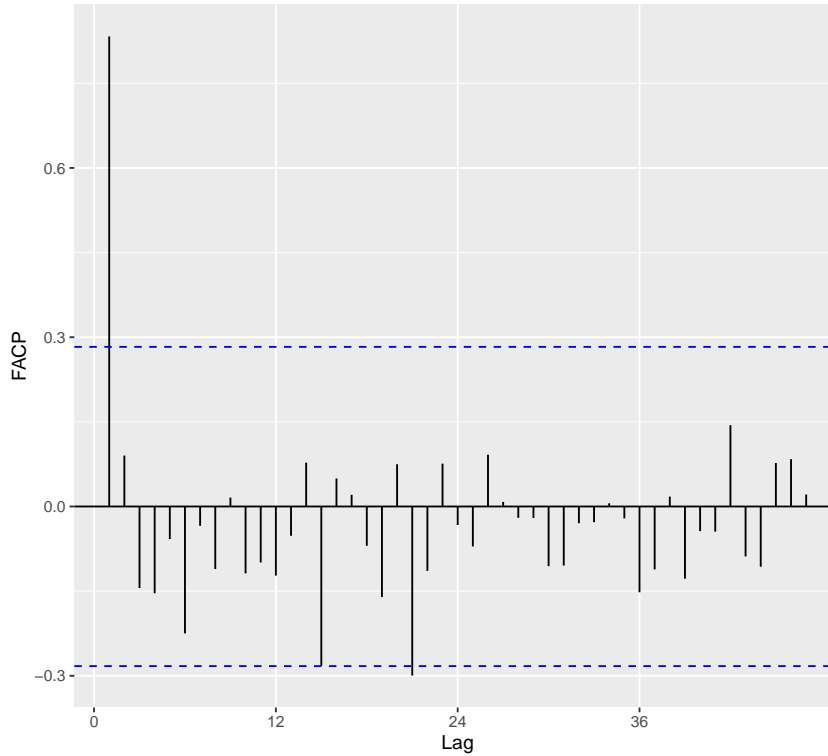


Figura 3.3 – Função de autocorrelação parcial.

Fonte: autor

### 3.4.4 Modelo de médias móveis (MA(q))

O modelo de médias móveis [58] de ordem  $q$  ( $MA(q)$ ) é definido por

$$Z_t = c + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (3.32)$$

onde  $a_t$  é o ruído branco e o operador de médias móveis de ordem  $q$  é dado por

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (3.33)$$

Fazendo  $\tilde{Z}_t = Z_t - c$ , tem-se

$$\tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t \quad (3.34)$$

#### 3.4.4.1 Comportamento teórico da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP), modelo de médias móveis

Quando o modelo de médias móveis é de ordem 1 ( $MA(1)$ ), a função de autocorrelação (FAC) é zero para  $k > 1$  e a função de autocorrelação parcial (FACP) apresenta decaimento oscilatório. Já para o modelo de médias móveis de ordem  $q$  ( $MA(q)$ ) a função de autocorrelação (FAC) anulam-se bruscamente e a função de autocorrelação parcial (FACP) não se anulam bruscamente (exponenciais amortecidas ou senoide).

### 3.4.5 Modelo autorregressivo e de médias móveis (ARMA(p,q))

O modelo autorregressivo e de médias móveis é a combinação dos modelos *AR* e *MA*. Representado por  $ARMA(p,q)$  [58] um processo autorregressivo e de médias móveis de ordem  $(p,q)$  pode ser escrito como

$$\phi(B)Z_t = \theta(B)a_t \quad (3.35)$$

onde  $\phi(B)$  é o operador autorregressivo (3.30) e  $\theta(B)$  é o operador de médias móveis (3.33), logo podemos reescrever a equação acima da seguinte forma

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.36)$$

### 3.4.6 Modelo ARIMA não sazonal

Quando a série temporal não é estacionária é necessário utilizar alguma transformação para torná-la estacionária. Dada uma série temporal  $Z_t$  não estacionária, a transformação de diferenças sucessivas pode ser aplicada com intuito de obter a série temporal estacionária transformada  $W_t$

$$W_t = \Delta^d Z_t \quad (3.37)$$

onde  $\Delta^d$  é o operador de diferenças sucessivas.

Assim pode-se escrever  $W_t$  por um modelo  $ARMA(p,q)$ , isto é,

$$\phi(B)W_t = \theta(B)a_t \quad (3.38)$$

onde  $\phi(B)$  é o operador autorregressivo (3.30),  $\theta(B)$  é o operador de médias móveis (3.33) e  $a_t$  é um ruído branco.

O processo autorregressivo integrado de médias móveis [58] denotado  $ARIMA(p; d; q)$  é dado por

$$W_t = c + \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} + a_t \quad (3.39)$$

Podendo ser reescrita por

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Z_t = c + (1 + \theta_1 B + \dots + \theta_q B^q) a_t \quad (3.40)$$

em que  $p$  e  $q$  são as ordens de  $\phi(B)$  e  $\theta(B)$ , respectivamente.

Ou ainda

$$\phi(B)\Delta^d Z_t = c + \theta(B)a_t \quad (3.41)$$

### 3.4.7 Modelo ARIMA sazonal (SARIMA)

Para dados que possuem sazonalidade os modelos  $ARIMA$  também são capazes de modelá-los. Para isso o modelo  $ARIMA$  sazonal (SARIMA) [58] é utilizado para absorver a componente de sazonalidade. A estrutura do modelo  $ARIMA$  sazonal de ordem  $(p,d,q) \times (P,D,Q)_s$  é dada por

$$\phi(B)\Phi(B^s)\Delta^d \Delta_s^D Z_t = \theta(B)\Theta(B^s)a_t \quad (3.42)$$

em que  $a_t$  é o ruído branco;  $\phi(B)$  é o operador autorregressivo de ordem  $p$ ;  $\theta(B)$  é o operador de médias móveis de ordem  $q$ ;  $\Phi(B^s)$  é o operador autorregressivo sazonal de ordem  $P$ ;  $\Theta(B^s)$  é o operador de médias móveis sazonal de ordem  $Q$ ;  $\Delta^d$  é o operador da diferença simples e;  $\Delta_s^D$  é o operador da diferença sazonal;  $s$  é o número de observações por ano (período).

O operador autorregressivo sazonal é dado por

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps} \quad (3.43)$$

O operador de médias móveis sazonal de ordem  $Q$  é

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs} \quad (3.44)$$

### 3.4.8 Estimação de parâmetros

A estimação dos parâmetros dos modelos pode ser realizada por um algoritmo de estimação que busca minimizar a diferença entre o valor real dos dados e o valor estimado pelo modelo. Assim, esta etapa pode ser considerada como a solução de um problema de otimização, ou seja, os parâmetros são determinados pela minimização de uma função objetivo.

Diversos métodos de estimação de parâmetros são encontrados na literatura, como o método de Mínimos Quadrados e suas extensões, Filtro de Kalman e o estimador de Máxima Verossimilhança.

O método de mínimos quadrados é um dos métodos mais comum e empregados em processos de estimação de parâmetros [63]. Este método pode ser empregado na estimação dos parâmetros  $(\phi_1, \phi_2, \dots, \phi_p)$  de um modelo  $AR(p)$ . Por outro lado, a estimação dos parâmetros  $(\theta_1, \theta_2, \dots, \theta_p)$  em modelos  $MA(p)$  não pode ser realizada pelo método de mínimos quadrados, uma vez que os erros  $(a_t)$  dos modelos  $MA(p)$  não se caracterizam como ruído branco, ou seja, os erros são funções não lineares. Assim, métodos computacionais iterativos precisam ser utilizados [57].

Nos modelos ARMA, ARIMA e SARIMA os problemas de estimação dos parâmetros são semelhantes aqueles apontados para os modelos MA no sentido de que um procedimento iterativo precisa ser utilizado. Neste trabalho não serão apresentados os detalhes das rotinas de estimação dos modelos abordados, tais detalhes podem ser encontrados em Box & Jenkins [59].

## 3.5 Modelos de Regressão Dinâmica

Os modelos de regressão dinâmica [64] combinam as características de variáveis explicativas com a dinâmica de séries temporais. Deve-se atentar que vocábulo “regressão dinâmica” não estabelece que os parâmetros do modelo evoluem com o tempo. Todavia, a palavra “dinâmica” caracteriza um modelo de regressão agregado pela estrutura de dependência de uma série temporal [65].

A aplicação dos modelos de regressão dinâmica fica condicionada a existência da estrutura de dependência entre a variável de interesse e variáveis causais e, concomitantemente a estrutura de correlação da série dependente (série a ser explicada) sugerir que não se pode supor a independência dos erros [65].

Os modelos de regressão dinâmica podem ser descritos por

$$\phi(B)Y_t = \beta x_t + \epsilon_t \quad (3.45)$$



onde  $Y_t$  é a variável dependente (endógena) no instante  $t$ ;  $\beta$  é o vetor de coeficientes das variáveis causais;  $x_t$  é o vetor de variáveis causais (exógenas) no instante  $t$ ;  $\epsilon_t$  é o ruído aleatório associado ao modelo, onde supõem-se que os  $\epsilon_t$  são independentes e identicamente distribuídos com densidade  $N(0, \sigma^2)$ ;  $\phi(B)$  é o polinômio autorregressivo de ordem  $p$ , dado por

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3.46)$$

sendo  $B$  o operador de atraso.

Assim, a estrutura do modelo de regressão dinâmica possibilita assumir como elementos  $x_t$  variáveis causais e juntamente suas defasagens. O polinômio  $\phi(B)$  presente no modelo proporciona uma grande flexibilidade desta classe de modelos, porém, dificulta a procura por um modelo adequado [65].

O modelo de regressão dinâmica (3.45), pode ser considerado como um caso particular do que é conhecido na literatura como modelo de regressão de Cochrane e Orcutt generalizado [65] que é dado por

$$\phi(B)y_t = \beta x_t + w_t \quad (3.47)$$

$$R(B)w_t = \epsilon_t \quad (3.48)$$

onde  $R(B)$  é um polinômio auto-regressivo.

A equação (3.47) pode ser reescrita isolando-se o erro  $w_t$ , da seguinte forma

$$w_t = \phi(B)y_t - \beta x_t \quad (3.49)$$

Substituindo a expressão (3.49) na equação (3.48), tem-se que

$$R(B) \cdot [\phi(B)y_t - \beta x_t] = \epsilon_t \quad (3.50)$$

Pode-se notar pela equação (3.50) que o modelo de regressão generalizado de Cochrane & Orcutt introduz defasagens tanto na variável dependente ( $Y_t$ ) quanto nas causais. Também é estabelecido por esta equação que a introdução do polinômio autorregressivo  $R(B)$  não afeta a relação de causalidade entre  $y_t$  e  $x_t$ . A equação (3.50) pode ser escrita em termos de novas variáveis  $Y_t$  e  $x_t$  dadas por

- $Y_t^* = R(B)y_t$
- $X_t^* = R(B)x_t$

onde  $R(B)$  é denominado de “fator comum” e representa a estrutura de correlação presente no erro  $w_t$ .

Com as novas variáveis obtêm-se a seguinte equação

$$\phi(B)Y_t^* = \beta X_t^* + \epsilon_t \quad (3.51)$$

Assim, o modelo de regressão generalizado de Cochrane & Orcutt se reduz ao modelo de regressão dinâmica usual ao se considerar as novas variáveis  $Y_t^*$  e  $X_t^*$ .

Para construir um modelo de regressão dinâmica é habitual utilizar uma estratégia bottom-up, isto é, inicia-se com um modelo simples e complementa-o com o intuito de o aperfeiçoar e inclui-se novas variáveis até obter um modelo propício. A obtenção de um modelo de regressão dinâmica é geralmente uma tarefa difícil, uma vez que além de escolher as variáveis que vão ser incluídas no modelo também é necessário escolher os lags (defasagens) destas variáveis.

Para determinar o modelo adequado, é preciso atentar-se a estrutura “lógica” do modelo e não somente na significância dos parâmetros. Em suma, para determinar o modelo de regressão a ser utilizado é indispensável verificar se os coeficientes estimados são coerentes e não somente obter um ajuste de parâmetros adequado [65].

As previsões originárias de um modelo de regressão dinâmica dependem dos valores passados da série e também dos valores previstos para as variáveis causais. Dessa forma, para obter as previsões da série  $Y_t$  em  $T + 1$ ,  $T + 2$ ,  $T + 3$ ,  $\dots$ , é necessário que seja provido ao modelo os valores futuros do vetor de variáveis causais  $x_t$ . Desta forma, se as previsões para as variáveis exógenas fornecidas ao modelo não forem satisfatórias, conseqüentemente o modelo de regressão dinâmica irá gerar previsões que também serão inadequadas.

### 3.6 Modelos Lineares Dinâmicos

Os modelos de espaço de estado [66], também conhecidos como modelos dinâmicos, consideram uma série temporal como a saída de um sistema dinâmico perturbado por distúrbios aleatórios. Eles permitem uma interpretação natural de uma série temporal como a combinação de vários componentes, como tendência, componentes sazonais ou regressivos. Ao mesmo tempo, eles têm uma estrutura probabilística elegante e poderosa, oferecendo uma estrutura flexível para uma ampla gama de aplicações. Os cálculos podem ser implementados por algoritmos recursivos. Os problemas de estimação e previsão são resolvidos pelo cálculo recursivo da distribuição condicional das quantidades de interesse, dadas as informações disponíveis. Nesse sentido, eles são tratados naturalmente dentro de uma estrutura bayesiana.

A classe dos modelos dinâmico é bastante abrangente, sendo os modelos lineares dinâmicos (MLD) um caso especial do modelo geral de espaço de estados, ou seja, uma subclasse destes modelos, sendo linear e gaussiano. Sendo assim, é suposto a normalidade para a variável resposta e para a evolução dos parâmetros dinâmicos através do tempo.

No modelo de regressão usual tem-se uma variável resposta  $y$  que é explicada por um conjunto de variáveis explicativas  $x_1, x_2, \dots, x_p$  através da relação

$$y = \theta_0 + \theta_1 x_1 + \dots + \theta_p x_p + \epsilon \quad (3.52)$$

onde  $\theta_1, \dots, \theta_p$  são os coeficientes de regressão e,  $\epsilon$  é uma perturbação aleatória assumindo-se que possui distribuição normal  $N(0, \sigma^2)$ .

Ainda pode-se escrever a equação (3.52) de forma mais compacta como

$$y = \mathbf{x}'\theta + \epsilon \quad (3.53)$$

onde

$$\mathbf{x}' = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_p \end{pmatrix} \quad e \quad \theta = \begin{pmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_p \end{pmatrix} \quad (3.54)$$

As variáveis explicativas ou regressoras possuem uma ampla série de origens. Desta forma, pode-se utilizar qualquer variável quantificável. Por sua vez, os coeficientes de regressão  $\theta_1, \dots, \theta_p$  informam sobre a influência que os regressores exercem sobre a resposta  $y$ .

Na abordagem de modelos dinâmicos os parâmetros mudam com o passar do tempo. Assim, o modelo de regressão usual (3.53) é estendido para

$$y_t = \mathbf{x}_t' \theta_t + \epsilon_t, \quad t = 1, \dots, n \quad (3.55)$$

em que a única mudança em relação ao modelo de regressão foi a indexação de  $\theta$ .

Assim, se o sistema é estático, como em regressão, tem-se

$$\theta_t = \theta_{t-1} = \theta \quad (3.56)$$

Em modelos dinâmicos, será admitido a forma mais geral, dada por

$$\theta_t = G_{t-1} \theta_{t-1} + w_t \quad (3.57)$$

em que  $G_t$  contém valores conhecidos e  $w_t$  é uma perturbação aleatória. Desta forma, tem-se que:

- A equação (3.57) é conhecida como equação do sistema;
- A matriz da evolução  $G_t$  controla a parte determinística da evolução do sistema e estabelece a propagação do sistema ao longo do tempo;
- A perturbação  $w_t$  é responsável pela introdução de incertezas devidas à passagem do tempo e a perda de informação;
- Se  $G_t = I$  e  $w_t = 0$ , o modelo se reduz ao caso estático.

O modelo linear dinâmico pode então ser definido por

$$\text{Equação das observações: } y_t = F_t' \theta_t + v_t, \quad v_t \sim N(0, V_t) \quad (3.58)$$

$$\text{Equação do sistema: } \theta_t = G_{t-1} \theta_{t-1} + w_t, \quad w_t \sim N(0, W_t) \quad (3.59)$$

$$\text{Informação inicial: } (\theta_0 | D_0) \sim N(m_0, C_0) \quad (3.60)$$

em que  $y_t$  é o vetor ( $n \times 1$ ) de observações e  $\theta_t$  é um vetor  $p$ -dimensional denominado parâmetro de estado do modelo dinâmico.  $m_0$  e  $C_0$  são os momentos da distribuição a priori inicial e são supostos conhecidos;  $F_t$  é uma matriz de regressão ( $n \times p$ ), cujos elementos são valores conhecidos;  $G_t$  é uma matriz ( $p \times p$ ) conhecida que descreve a evolução temporal dos parâmetros de estado;  $V_t$  é uma matriz de covariância ( $n \times n$ ) conhecida associada ao erro observacional  $v_t$ ;  $W_t$  é uma matriz de covariância ( $p \times p$ ) conhecida associada ao erro de evolução dos parâmetros de estado  $w_t$ .

O modelo linear dinâmico definido pela estrutura apresentada pelas equações em (3.58), (3.59) e (3.60) pode caracterizar uma grande variedade de modelos que são casos especiais dos modelos lineares dinâmicos. A escolha de  $F_t$  e  $G_t$  define a subclasse do modelo linear dinâmico escolhido. Nas seções a seguir serão apresentados o passeio aleatório mais ruído e modelo de crescimento linear, que são dois casos especiais dos modelos lineares dinâmicos.

### 3.6.1 Passeio aleatório mais ruído

O modelo mais simples para uma série temporal univariada ( $y_t : t = 1, 2, \dots$ ) é o chamado modelo de passeio aleatório mais ruído, definido por

$$y_t = \mu_t + v_t, \quad v_t \sim N(0, V) \quad (3.61)$$

$$\mu_t = \mu_{t-1} + w_t, \quad w_t \sim N(0, W) \quad (3.62)$$

Pode-se perceber que o modelo representado pelas equações (3.61) e (3.62) é um MLD com:

- $n = p = 1$ ;
- $\theta_t = \mu_t, \quad t \geq 1$ ;
- $F_t = G_t = 1, \quad t \geq 1$ ;
- $V_t = V, \quad t \geq 1$ ;
- $W_t = W, \quad t \geq 1$ ;
- Sequências de erros ( $v_t$ ) e ( $w_t$ ) são internamente e mutuamente independentes.

Este modelo também é conhecido como modelo de nível local, sendo apropriado para séries temporais que não possuem variação de tendência ou sazonalidade. O modelo considera que as observações provêm de um ruído mais um nível  $\mu_t$ , assumindo que o nível está sujeito a perturbações aleatórias ao longo do tempo. Outro aspecto a ser observado é que se  $W$  for aproximadamente zero, tem-se aproximadamente um modelo normal, com média e variância constantes.

### 3.6.2 Modelo de crescimento linear

O modelo de crescimento linear, ou modelo de tendência linear local, tem a mesma equação de observação que o modelo de nível local, mas inclui uma inclinação variável no tempo na dinâmica para  $\mu_t$ .

$$y_t = \mu_t + v_t, \quad v_t \sim N(0, \sigma_{obs}^2) \quad (3.63)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + w_{t,1}, \quad w_{t,1} \sim N(0, \sigma_\mu^2) \quad (3.64)$$

$$\beta_t = \beta_{t-1} + w_{t,2}, \quad w_{t,2} \sim N(0, \sigma_\beta^2) \quad (3.65)$$

O modelo representado pelas equações (3.63), (3.64) e (3.65) é um MLD especificado pelo vetor de estado  $\theta_t$  e por  $\{G_t, F_t, V_t, W_t\}$ . Assim:

- $n = 1$ ;
- $p = 2$ ;
- $F_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad t \geq 1$ ;
- $\theta_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}, \quad t \geq 1$ ;
- $G_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad t \geq 1$ ;
- $V_t = V, \quad t \geq 1$ ;

- $w_t = \begin{bmatrix} w_{t,1} \\ w_{t,2} \end{bmatrix};$
- $W_t = \begin{bmatrix} \sigma_\mu^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix}, \quad t \geq 1;$
- Parâmetros para as variações dos erros:  $e = \begin{bmatrix} \sigma_{obs}^2 & \sigma_\mu^2 & \sigma_\beta^2 \end{bmatrix}^T.$

Quando  $\sigma_\beta^2$  é próximo de zero, tem-se aproximadamente um modelo com incremento constante ao longo do tempo. Neste exemplo, as matrizes  $F_t$  e  $G_t$ , e as matrizes de covariâncias  $V_t$  e  $W_t$  são constantes ao longo do tempo. Nestes casos, o modelo é dito invariante no tempo.

### 3.6.3 Modelo de crescimento linear com sazonalidade

O modelo linear dinâmico de crescimento linear e sazonalidade é definido por

$$Y_t = \mu_t + \gamma_t + \epsilon_t, \quad \epsilon_t \sim N(0, V_\epsilon) \quad (3.66)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \delta_{\mu t}, \quad \delta_{\mu t} \sim N(0, V_\mu) \quad (3.67)$$

$$\beta_t = \beta_{t-1} + \delta_{\beta t}, \quad \delta_{\beta t} \sim N(0, V_\beta) \quad (3.68)$$

$$\gamma_t = - \sum_{i=1}^{11} \gamma_{t-1} + \delta_{\gamma t}, \quad \delta_{\gamma t} \sim N(0, V_\gamma) \quad (3.69)$$

em que  $\mu_t$  é o nível em  $t$ ,  $\beta_t$  é a inclinação em  $t$ ,  $\gamma_t$  é o fator sazonal,  $V_\epsilon$  é a variância observacional,  $V_\mu$ ,  $V_\beta$ ,  $V_\gamma$ , são as variâncias dos estados do sistema, e os erros  $\epsilon_t$ ,  $\delta_{\mu t}$ ,  $\delta_{\beta t}$  e  $\delta_{\gamma t}$  são internamente e mutuamente independentes.

### 3.6.4 Filtro de Kalman para os modelos lineares dinâmicos

O Filtro de Kalman [66] refere-se a um conjunto de procedimentos recursivos para estimação em modelos dinâmicos. Considerando o MLD definido em (3.58), (3.59) e (3.60) com  $\{G_t, F_t, V_t, W_t\}$  conhecidos para  $t = 1, 2, \dots$ . O filtro de Kalman é dividido em três etapas dependentes: evolução, previsão e atualização. Sendo  $D_t$  a informação total obtida até o tempo  $t$ ,  $(\theta_t|D_t)$  a distribuição a posteriori e  $(\theta_t|D_{t-1})$  a distribuição a priori, o filtro de Kalman pode ser ilustrado por

$$\begin{array}{ccccccc} \cdots & (\theta_{t-1}|D_{t-1}) & \xrightarrow{\text{Evolução}} & (\theta_t|D_{t-1}) & \xrightarrow{\text{Atualização}} & (\theta_t|D_t) & \cdots \\ & & & \Downarrow \text{Previsão} & & & \\ & & & (Y_t|D_{t-1}) & & & \end{array} \quad (3.70)$$

Conforme ilustrado por (3.70), a evolução é a passagem de  $(\theta_{t-1}|D_{t-1})$  para  $(\theta_t|D_{t-1})$ , a previsão é a obtenção de  $(Y_t|D_{t-1})$  e a atualização é a passagem de  $(\theta_t|D_{t-1})$  para  $(\theta_t|D_t)$ .

Pelo filtro de Kalman, considerando o MDL especificado por (3.58), (3.59) e (3.60), e a partir da distribuição

$$(\theta_{t-1}|D_{t-1}) \sim N(m_{t-1}, c_{t-1}) \quad (3.71)$$

temos que:

I. A distribuição preditiva um passo a frente de  $\theta_t$  dado  $D_{t-1}$  é Normal, com parâmetros:

$$a_t = E[\theta_t|D_{t-1}] = G_t m_{t-1}, \quad (3.72)$$

$$R_t = Var[\theta_t|D_{t-1}] = G_t C_{t-1} G_t' + W_t \quad (3.73)$$

II. A distribuição preditiva um passo à frente de  $Y_t$  dado  $D_{t-1}$  é Normal, com parâmetros:

$$f_t = E[Y_t|D_{t-1}] = F_t a_t, \quad (3.74)$$

$$Q_t = Var[Y_t|D_{t-1}] = F_t R_{t-1} F_t' + V_t \quad (3.75)$$

III. Distribuição de filtragem de  $\theta_t$  dado  $D_t$  é Normal, com parâmetros:

$$m_t = E[\theta_t|D_t] = a_t R_t F_t' Q_t^{-1} e_t, \quad (3.76)$$

$$C_t = Var[\theta_t|D_t] = R_t - R_t F_t' Q_t^{-1} + F_t R_t \quad (3.77)$$

em que  $e_t = Y_t - f_t$  é o erro de previsão.

As equações de atualização são dadas por:  $(\theta_t) \sim N(m_t, c_t)$  em que

$$m_t = a_t + A_t(y_t - f_t), \quad (3.78)$$

$$C_t = R_t + A_t Q_t A_t', \quad (3.79)$$

$$A_t = R_t + F_t' Q_t^{-1} \quad (3.80)$$

E, as equações de previsão  $k$  passos a frente podem ser escritas por

$$(\theta_{t+k}|D_t) \sim N(a_t(k), R_t(k)), \quad (3.81)$$

$$(Y_{t+k}|D_t) \sim N(f_t(k), Q_t(k)), \quad (3.82)$$

### 3.7 Redes Neurais Artificiais

As redes neurais artificiais foram propostas pela primeira vez por Warren McCullough e Walter Pitts [67]. Desde então, o uso de redes neurais nas mais diversas áreas do conhecimento vem sendo cada vez mais recorrente [65]. O bom desempenho da metodologia tem sido observado nos mais diversos campos, inclusive nas previsões utilizando séries temporais [65].

Segundo Zanini [65] a metodologia foi motivada pela tentativa de modelar a rede de neurônios humanos visando compreender o funcionamento do cérebro, porém, apenas no início do desenvolvimento de redes neurais a motivação fisiológica desempenhou papel relevante. Nas pesquisas atuais, neste campo do conhecimento, não há relação com as peculiaridades do funcionamento do cérebro como um todo ou do neurônio mais especificamente.

Uma Rede Neural Artificial (ANN - Artificial Neural Network) modela a relação entre um conjunto de sinais de entrada e um sinal de saída usando um modelo derivado de nossa compreensão de como um cérebro biológico responde a estímulos de entradas sensoriais. Assim como o cérebro usa uma rede de células interconectadas chamadas neurônios para criar um processador paralelo massivo, uma ANN é construída a partir de uma rede de neurônios ou nós artificiais para resolver problemas de aprendizagem.

### 3.7.1 Arquitetura das redes Neurais

A “arquitetura” de uma rede neural artificial refere-se a maneira com que as unidades (ou neurônios) e suas conexões estão arranjadas [65]. Sucintamente, uma rede neural artificial pode ser dividida em três partes, chamadas de: camada de entrada, camada oculta e camada de saída.

A camada de entrada é responsável por receber a informação, sendo esta um dado vindo do meio externo. Por sua vez, as camadas ocultas são dotadas de neurônios que extraem as características associadas ao processo ou sistema a ser inferido, sendo praticamente toda a responsabilidade de processamento interno da rede realizado nesta camada. E, por fim, a camada de saída que é também constituída de neurônios e, é responsável pela produção e apresentação dos resultados finais da rede os quais vieram dos processamentos dos neurônios das camadas anteriores [68].

As principais arquiteturas de redes neurais artificiais podem ser divididas em: redes *feedforward* de camadas simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas [68].

### 3.7.2 Modelo para uma rede *feedforward*

Dentre as arquiteturas existentes a mais comum é a denominada *feedforward*, Figura 3.4. Esta Figura traz uma rede *feedforward* com três entradas, uma camada oculta (quatro neurônios) e uma camada de saída (dois neurônios). Neste arranjo o sinal é propagado para frente, ou seja, o fluxo de dados será da entrada para a saída, sem a realimentação dos neurônios.

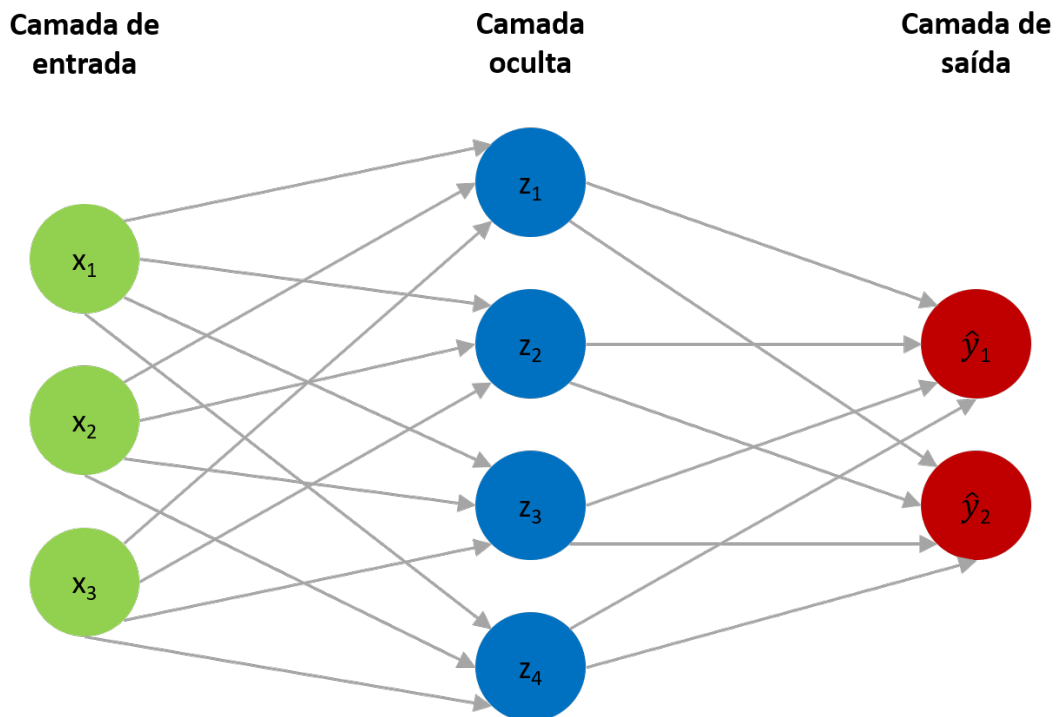


Figura 3.4 – Diagrama de uma rede neural.

Fonte: autor

De acordo com Everitt et al. [51], o modelo especificado na Figura 3.4 pode ser escrito por

$$z_j = f(\alpha_{j0} + \alpha'_j \mathbf{x}), \quad j = 1, 2, \dots, 4; \quad (3.83)$$

$$\hat{y}_k = g_k(\beta_{k0} + \beta'_k \mathbf{z}), \quad z = 1, 2. \quad (3.84)$$

onde  $f$  é conhecida como função de ativação,  $\beta_{k0}$  denota o intercepto do neurônio de saída,  $\alpha_{j0}$  representa o intercepto do  $j$ -ésimo neurônio oculto,  $\mathbf{x}' = (x_1, x_2, x_3)$  são as observações (covariáveis ou neurônios) da camada de entrada,  $\mathbf{z}' = (z_1, z_2, z_3, z_4)$  são os neurônios da camada oculta,  $\hat{y}'_k = (\hat{y}_1, \hat{y}_2)$  são os neurônios da camada de saída,  $\alpha'_j = (\alpha_{j1}, \alpha_{j2}, \alpha_{j3})$  são os pesos atribuídos à conexão entre a entrada e a camada oculta,  $\beta'_k = (\beta_{k1}, \beta_{k2}, \beta_{k3}, \beta_{k4})$  são os pesos entre a camada oculta e a camada de saída e  $g_k$  é uma função de integração incluída para permitir uma transformação final da saída.

Os parâmetros de uma rede neural são os pesos. Os pesos usados em um modelo de rede neural são estimados a partir dos dados do conjunto de treinamento por mínimos quadrados, por exemplo, para a rede descrita acima, minimizando

$$R(\alpha, \beta) = \sum_k (y_k - \hat{y}_k)^2 \quad (3.85)$$

um critério que não é linear nos parâmetros. Muitas das vezes não é fácil minimizar  $R$ , uma vez que ele pode ter mínimos locais e normalmente as redes neurais são super-parametrizadas, muitas das vezes com mais parâmetros do que observações. Na literatura de rede neural, este estágio de estimativa é frequentemente descrito como “treinamento” da rede.

A função de ativação pode ser escolhida dentre as seguintes funções:

- identidade:  $f(z) = z$ ;
- logística:  $f(z) = \frac{1}{1+e^{-z}}$ ;
- tangente hiperbólica:  $f(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$ ;

### 3.7.3 Modelo NNAR

Com os dados da série temporal, os valores defasados da série temporal podem ser usados como entradas para uma rede neural, assim como usamos os valores defasados em um modelo de autorregressão linear. Chamamos isso de rede neural autorregressiva ou modelo NNAR.

Neste estudo será considerado apenas a rede unidirecional (feedforward) com uma camada oculta e representada pela notação  $NNAR(p, k)$  para indicar que há  $p$  entradas defasadas e  $k$  nós na camada oculta. Por exemplo, um modelo  $NNAR(10, 6)$  é uma rede neural com as últimas dez observações ( $y_{t-1}, y_{t-2}, \dots, y_{t-10}$ ) usado como entrada para prever a saída  $y_t$  e com seis neurônios na camada oculta. Agora, com dados sazonais, é importante também adicionar os últimos valores observados da mesma estação como entradas. Por exemplo, um modelo  $NNAR(4, 1, 3)_{12}$  tem entradas ( $y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$ ) e  $y_{t-12}$  e três neurônios na camada oculta. De forma geral a notação  $NNAR(p, P, k)_m$ ,  $p$  número de defasagens não sazonais usadas como entradas ( $p = 4$ , são as 4 últimas observações),  $P$  representa o número de defasagens sazonais usadas como entradas (por exemplo, para  $P = 1$  e  $m = 12$ , é considerado o valor de 12 amostras/meses atrás), e  $k$  é o número de nós (neurônios) na camada oculta, [62].



### 3.8 Métricas para avaliar os modelos de séries temporais

Neste trabalho, para realizar a escolha do modelo de previsão mais adequado às séries temporais em análise, utilizou-se a métrica MAPE e a métrica MSE.

#### 3.8.1 Erro médio absoluto percentual (MAPE)

O erro médio absoluto percentual (MAPE) é definido por

$$MAPE = 100 \times \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3.86)$$

onde  $N$  é o número de pontos ajustados,  $y_i$  é o valor observado, e  $\hat{y}_i$  é o valor previsto também para o instante  $i$ .

#### 3.8.2 Erro quadrático médio (MSE)

O erro quadrático médio (MSE) é definido por

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3.87)$$

onde  $N$  é o número de pontos ajustados,  $y_i$  é o valor observado, e  $\hat{y}_i$  é o valor previsto também para o instante  $i$ .

### 3.9 Modelo bottom-up

As abordagens bottom-up e top-down são estratégias de processamento de informação e ordenação do conhecimento, usadas em várias áreas de conhecimento [47]. No estudo de séries temporais, essas abordagens são utilizadas quando uma variável é organizada em uma estrutura hierárquica [62]. Na abordagem top-down as previsões são geradas para série temporal completamente agregada, isto é, as variáveis de interesse são tratadas como um todo, fornecendo resultados no nível mais alto da estrutura hierárquica. Já na abordagem bottom-up, a estrutura hierárquica é caracterizada pelo agrupamento de séries temporais de níveis mais baixos para formar séries em outro nível da hierarquia, ou seja, as variáveis de interesse são desagregadas em componentes, essas componentes recebem um tratamento, e os resultados gerados pelas componentes são acumulados para compor as variáveis de interesse.

Neste trabalho foi proposto uma abordagem bottom-up, considerando dois níveis na estrutura hierárquica. Desta forma, o consumo de energia elétrica do setor, sendo o nível superior da estrutura hierárquica, é definido por

$$W_t = \sum_{i=1}^p W_{i,t} \quad (3.88)$$

onde  $W_t$  é o consumo de energia elétrica total do setor no tempo  $t$ ,  $W_{i,t}$  é o consumo de energia elétrica da  $i$ -ésima série temporal agregada por cluster no tempo  $t$ , e  $p$  é o número de cluster.

#### 3.10 Medidas de Eficiência Energética (MEE)

A eficiência energética busca realizar uma mesma atividade consumindo menos energia [69]. Ou seja, a eficiência energética é definida como a relação entre a quantidade de energia

empregada em uma atividade e aquela disponibilizada para sua realização. Neste contexto, as medidas de eficiência energética são utilizadas para reduzir o consumo de energia ao longo do tempo.

Para descrever a aceitação ao longo do tempo de uma nova tecnologia em substituição à tecnologias existentes comumente utiliza-se uma curva logística (ou curva de difusão em forma de S) [70], [71]. A curva de difusão resultante é um resumo matemático de como as medidas de eficiência energética vão ser implementadas ao longo do tempo, podendo ter a origem, por exemplo, em uma estratégia de implementação.

Segundo Mathews e Baroni [71], Jaffe e Stavins [72] e Da Silva [3] os valores de uma curva de difusão podem ser estimados por uma função logística. Assim, para descrever a penetração das medidas de eficiência energética através dos valores de uma curva de difusão definiu-se neste estudo uma função logística dada por

$$d_x(t) = \frac{D_{x,0}D_{x,f}e^{\lambda_x(t-t_{x,0})}}{D_{x,f} + D_{x,0}[e^{\lambda_x(t-t_{x,0})} - 1]} \quad (3.89)$$

onde  $x$  é o tipo de difusão,  $t$  é o ano de interesse,  $\lambda$  é o parâmetro de inclinação da curva logística,  $t_{x,0}$  é o ano inicial,  $t_{x,f}$  é ano final,  $D_{x,0}$  é o valor inicial da curva de difusão em  $t_{x,0}$  e  $D_{x,f}$  é o valor final da curva de difusão em  $t_{x,f}$ .

## Materiais e Métodos

### 4.1 Materiais

Os dados de consumo de energia elétrica da UFRRJ estão distribuídos por setores. Para o presente trabalho foram utilizados os dados de janeiro de 2014 a maio de 2021. Sendo que de janeiro de 2014 a setembro de 2017 foram obtidos a partir de informações da UFRRJ, e de outubro de 2017 a maio de 2021 foram disponibilizados pela concessionária de energia que possui a concessão da região da UFRRJ. A variável regressora utilizada neste trabalho denota o número de servidores trabalhando presencialmente na Universidade no período janeiro de 2014 a maio de 2021, informações estas extraídas da UFRRJ. Atualmente, o consumo de energia elétrica da Universidade é integrado por 107 setores com medidores individuais.

Nesse estudo, são considerados 20 setores devido à falta de dados para alguns medidores. Mas, os setores investigados são responsáveis por aproximadamente 70% do consumo de energia elétrica da UFRRJ.

O período analisado foi impactado pela pandemia de COVID-19, o que acarretou em uma redução do consumo de energia elétrica na UFRRJ a partir de março de 2020, como pode ser observado na Figura 4.1.

### 4.2 Metodologia

As etapas da metodologia utilizada são apresentadas pelo fluxograma da Figura 4.2.

Esta metodologia consiste de uma primeira etapa de mapeamento dos dados. Em seguida, a partir dos setores mapeados aplica-se análise cluster para agrupá-los. Na próxima etapa, utiliza os modelos de séries temporais para avaliar a capacidade preditiva em cada cluster obtido. Posteriormente, através de uma abordagem bottom-up é possível agrupar as previsões dos clusters. Finalmente, é realizada a projeção do consumo de energia elétrica da UFRRJ considerando a implementação de MEE. A descrição de cada etapa é apresentada no Capítulo 3.

Para facilitar a análise do desempenho dos modelos aplicados às séries temporais do consumo de energia elétrica, propõe-se a elaboração de uma análise de Cluster, utilizando a técnica k-médias [51], para os dados de consumo de energia elétrica dos 20 setores da UFRRJ, ocasionando em um ganho de tempo por evitar a análise em separado de cada série temporal de cada setor. Em seguida, foi utilizada os modelos de Box e Jenkins [59], o método de Holt-Winters [61], os modelos lineares dinâmicos [66], os modelos de redes neurais artificiais [62] e os modelos de regressão dinâmica [65] para realizar o ajuste e a previsão da série temporal

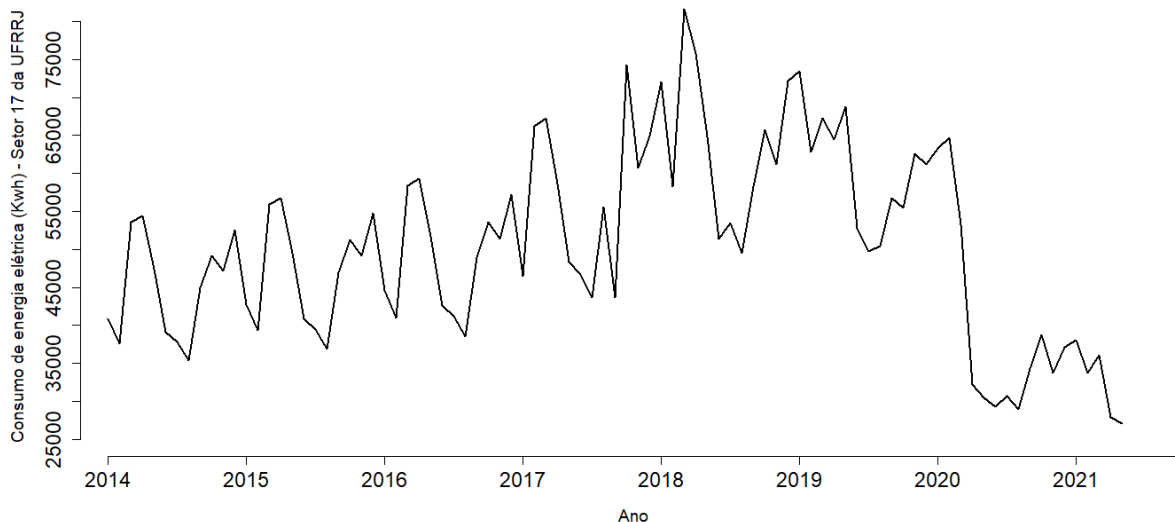


Figura 4.1 – Série temporal do consumo de energia elétrica do setor 17 da UFRRJ.  
Fonte: autor

resultante de cada Cluster. Consecutivamente, para avaliar o modelo de previsão obtido foi utilizado o erro médio absoluto percentual (MAPE) e o erro quadrático médio (MSE). Após classificar os modelos adotados que apresentaram o melhor desempenho preditivo através da métrica MAPE, utilizou-se a estratégia bottom-up para obter a previsão do consumo de energia da UFRRJ. Finalmente, de posse da previsão do consumo de energia elétrica da UFRRJ obtida pela estratégia bottom-up, realizou-se a projeção do consumo de energia elétrica da UFRRJ para diferentes cenários considerando o efeito das MEE.

Para análise e previsão das séries temporais do consumo de eletricidade foram considerados os seguintes cenários:

- Cenário 1: Ajuste dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018. Previsões dos modelos para o consumo de energia elétrica no período de janeiro de 2019 a dezembro de 2019.
- Cenário 2: Ajuste dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020. Previsões dos modelos para o consumo de energia elétrica no período de abril de 2020 a maio de 2021.
- Cenário 3: Ajuste dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021. Previsões dos modelos para o consumo de energia elétrica no período de junho de 2021 a maio de 2026.

As análises estatísticas foram realizadas através do software R 4.0.5 (The R Foundation for Statistical Computing, Vienna, Austria; <http://www.r-project.org>), conforme Anexo.

#### 4.2.1 Modelo proposto para implementação de MEE

Neste trabalho foram utilizadas 3 medidas de eficiência energética, sendo elas: troca de lâmpadas fluorescente por LED, substituição de aparelhos de ar condicionado de janela por

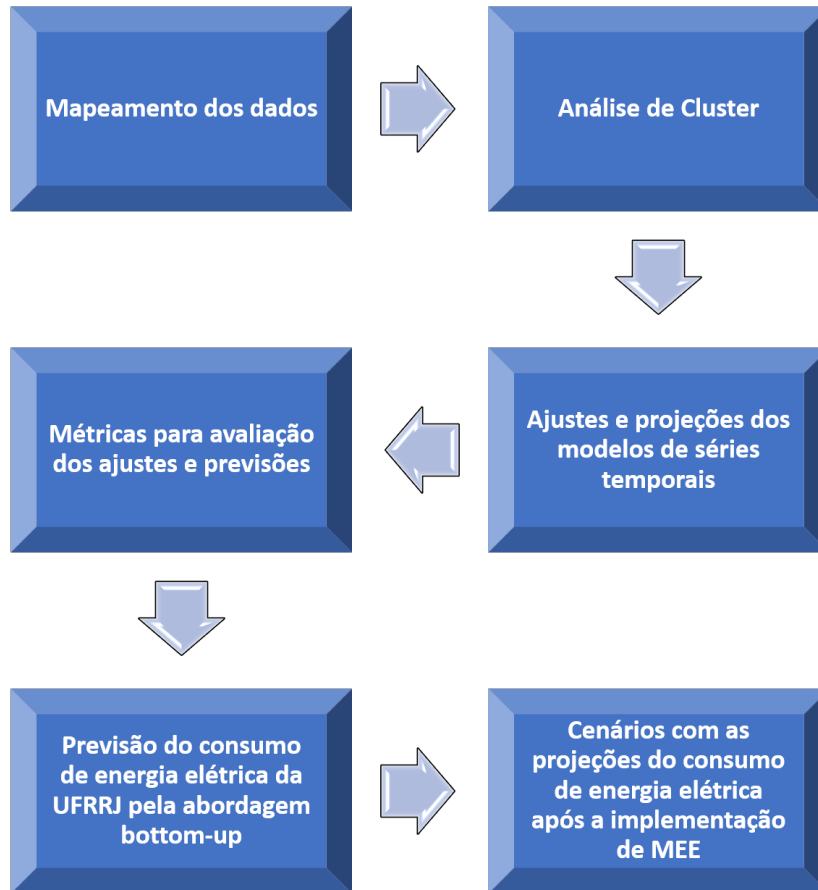


Figura 4.2 – Fluxograma com as etapas da metodologia.

Fonte: autor

aparelhos de ar condicionado Split e substituição de monitores de cinescópio por monitores de LED. O potencial de redução do consumo de energia elétrica das MEE foram retiradas no site do Immetro [73] e do Procel [74]. A representação percentual do consumo de energia elétrica dos equipamentos em relação ao consumo da energia elétrica total da Universidade considerou, conforme dados do Procel, que cerca de 70% do consumo de energia elétrica nos prédios públicos se deve ao uso dos sistemas de iluminação e climatização dessas edificações [75].

Assim o consumo de energia elétrica total da UFRRJ após a implementação das MEE, pode ser escrito por

$$W_t = \sum_{i=1}^n W_{i,t} \quad (4.1)$$

onde  $n$  é o número de clusters e,  $W_{i,t}$  é o consumo de energia elétrica da  $i$ -ésima série temporal agregada por cluster no tempo  $t$  após medidas de eficiência energética, dado por

$$W_{i,t} = \sum_{j=1}^J (C_{i,j,t} - E_{i,j,t}) \quad (4.2)$$

onde  $J$  é o número de grupos de equipamentos,  $C_{i,j,t}$  é o consumo de energia elétrica do grupo de equipamentos  $j$  do cluster  $i$  no tempo  $t$  e,  $E_{i,j,t}$  é a energia economizada pelo grupo de equipamentos  $j$  do cluster  $i$  no tempo  $t$ .

O consumo de energia elétrica do grupo de equipamentos  $j$  do cluster  $i$  no tempo  $t$  é definido por

$$C_{i,j,t} = \phi_j \times G_{i,t} \quad (4.3)$$

onde  $\phi_j$  é o percentual de consumo de energia elétrica do grupo de equipamentos  $j$  e  $G_{i,t}$  é o consumo de energia elétrica da  $i$ -ésima série temporal agregada por cluster no tempo  $t$  inicialmente calculada sem medidas de eficiência energética.

A energia economizada pelo grupo de equipamentos  $j$  do cluster  $i$  no tempo  $t$  é definida como

$$E_{i,j,t} = \sum_{k=1}^m (\alpha_{i,j,k,t} \times C_{i,j,t}) \quad (4.4)$$

onde  $m$  é o número de medidas de eficiência energética e,  $\alpha_{i,j,k,t}$  é o percentual de economia de energia da medida de eficiência energética  $k$  aplicada ao grupo de equipamentos  $j$  do cluster  $i$  no tempo  $t$ .

Neste trabalho foram considerados quatro cenários para descrever a difusão das medidas de eficiência energética, sendo nomeados por Cenário Congelado, Cenário Mínimo, Cenário Médio e Cenário Máximo. O Cenário Congelado é o cenário onde não há implementação de nenhuma medida de eficiência energética. O Cenário Mínimo possui curva de difusão atingindo 50% de implementação de medidas de eficiência energética. O Cenário Médio possui curva de difusão atingindo 75% de implementação de medidas de eficiência energética. E por fim, o Cenário Máximo possui curva de difusão atingindo 100% de implementação de medidas de eficiência energética.

### 5.1 Análise Descritiva dos Setores

Na tabela 5.1 é apresentado a análise descritiva dos vinte setores em estudo no período de janeiro de 2014 a março de 2020. Nessa tabela é apresentado a média, variância, desvio padrão e o coeficiente de variação para os dados de consumo de energia elétrica (kWh) que compõem a série temporal de cada setor.

Tabela 5.1 – Análise Descritiva dos setores.

Setor	Média	Variância	Desvio Padrão	Coeficiente de Variação
1	5705,85	5476580,20	2340,21	41,01
2	10864,83	16546352,74	4067,72	37,44
3	15400,98	83225706,11	9122,81	59,24
4	4557,19	1252915,89	1119,34	24,56
5	24776,33	57796621,86	7602,41	30,68
6	19490,60	45881611,65	6773,60	34,75
7	23500,08	54947419,48	7412,65	31,54
8	17446,26	27784234,54	5271,08	30,21
9	61198,87	357573621,70	18909,62	30,90
10	33709,27	71144744,07	8434,73	25,02
11	8832,28	126986850,60	11268,84	127,59
12	55667,07	268478834,90	16385,32	29,43
13	16829,18	8379268,79	2894,70	17,20
14	16290,29	13954746,56	3735,61	22,93
15	29697,97	201711601,40	14202,52	47,82
16	37111,33	30834809,89	5552,91	14,96
17	66682,61	114066751,20	10680,20	16,02
18	6519,23	9414217,67	3068,26	47,06
19	2514,87	145090,93	380,91	15,15
20	30290,33	190519961,60	13802,90	45,57

O setor com maior consumo de energia elétrica no período em estudo foi o setor 17, o que teve o menor consumo foi o setor 19. Em relação a variabilidade dos dados, o setor 11 apresentou a maior variabilidade e o setor com a menor variabilidade é o setor 16.

## 5.2 Análise de Cluster

A Figura 5.1 apresenta o resultado da análise de Cluster para os dados de consumo dos 20 setores em estudo da UFRRJ. Pode-se observar que foram obtidos 3 clusters.

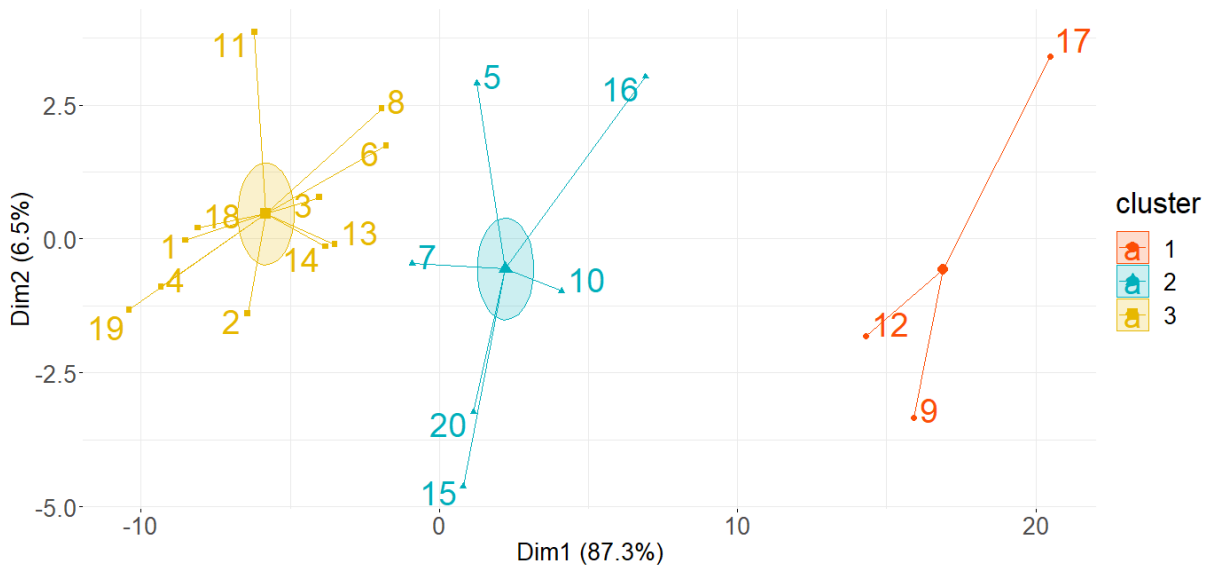


Figura 5.1 – Clusters obtidos pelo método k-médias.

A Tabela 5.2 traz a análise descritiva dos dados que compõem a série temporal resultante da integração dos setores que formam cada um dos três clusters encontrados na Figura 5.1. Na comparação entre a Tabela 5.1 e a Tabela 5.2, pode-se observar que o coeficiente de variação dos clusters tem resultados melhores que os encontrados para a maioria dos setores. Isso significa que os dados agrupados nos clusters tem baixa dispersão.

Tabela 5.2 – Análise Descritiva dos Clusters.

Cluster	Média	Variância	Desvio Padrão	Coefficiente de Variação
1	183548,55	1533350039,00	39158,01	21,33
2	179085,30	1646205911,00	40573,46	22,66
3	124451,57	1051969549,00	32434,08	26,06

Nas próximas seções serão apresentados três cenários onde são realizados os ajustes e as previsões para as séries temporais agregadas dos clusters obtidos. Estes cenários possibilitam analisar o desempenho dos modelos de Box e Jenkins, do método de Holt-Winters, dos modelos lineares dinâmicos, dos modelos de redes neurais artificiais e dos modelos de regressão dinâmica no período pré-pandemia, de pandemia e pós-pandemia de COVID-19. Sendo Cenário 1 o período de análise pré-pandemia, o Cenário 2 o período da pandemia e por fim o Cenário 3 sendo o período pós-pandemia.

## 5.3 Cenário 1

Nesta seção serão apresentados os ajustes dos modelos no período de janeiro de 2014 a dezembro de 2018, e as previsões serão realizadas para o período de janeiro de 2019 a de-



zembro de 2019. Neste período os dados não sofreram influência da pandemia de COVID-19 e, portanto, este Cenário é considerado pré-pandemia.

### 5.3.1 Previsões e ajustes para o Cluster 1

Nas Figuras 5.2 e 5.3 tem-se os ajustes e as previsões, respectivamente, dos modelos ARIMA  $(0,0,0)(0,1,1)_{12}$ , linear dinâmico, de redes neurais artificiais, de Holt-Winters aditivo e de regressão dinâmica para série temporal do consumo de energia elétrica proveniente do Cluster 1. O modelo de regressão dinâmica apresentou o melhor ajuste aos dados de consumo de energia elétrica provenientes do Cluster 1, enquanto a melhor previsão foi alcançada pelo modelo de redes neurais.

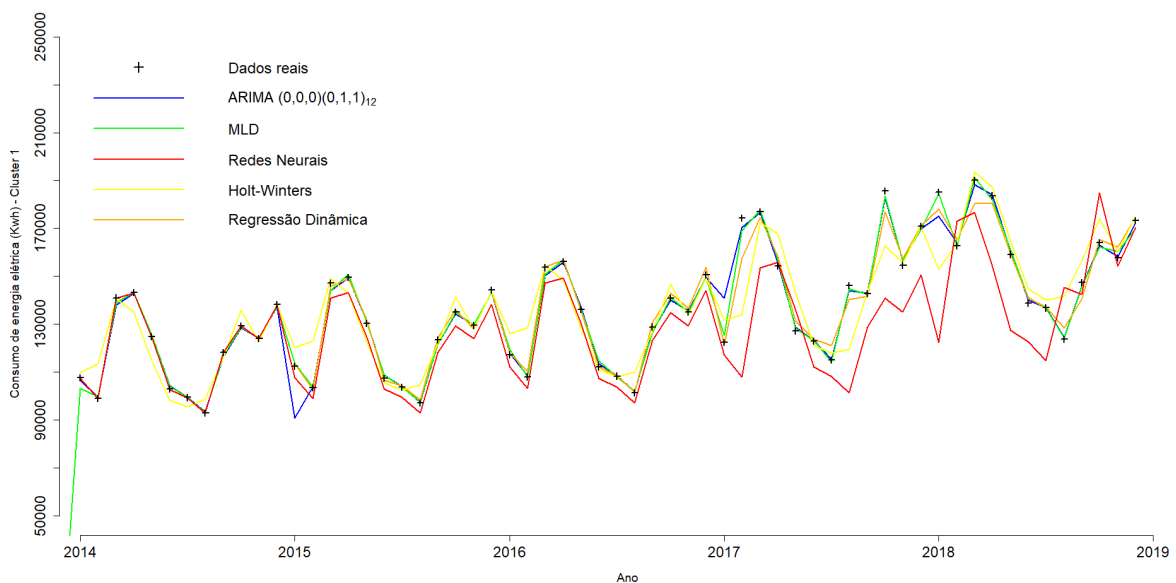


Figura 5.2 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018 para séries temporais do Cluster 1.

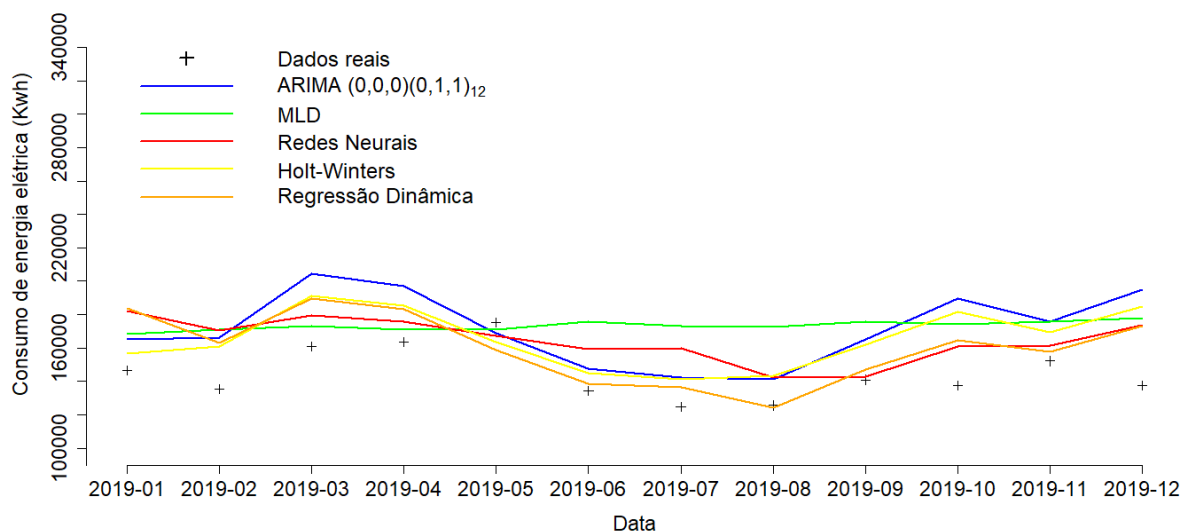


Figura 5.3 – Previsões dos modelos para o consumo de energia elétrica para o período de janeiro de 2019 a dezembro de 2019 para séries temporais do Cluster 1.

### 5.3.2 Previsões e ajustes para o Cluster 2

Nas Figuras 5.4 e 5.5 tem-se os ajustes e as previsões, respectivamente, dos modelos ARIMA  $(0, 1, 1)(1, 0, 0)_{12}$ , linear dinâmico, de redes neurais artificiais, de Holt-Winters aditivo e de regressão dinâmica para série temporal do consumo de energia elétrica proveniente do Cluster 2. O modelo de regressão dinâmica apresentou o melhor ajuste aos dados de consumo de energia elétrica provenientes do Cluster 2, enquanto a melhor previsão foi alcançada pelos modelos de Box e Jenkins.

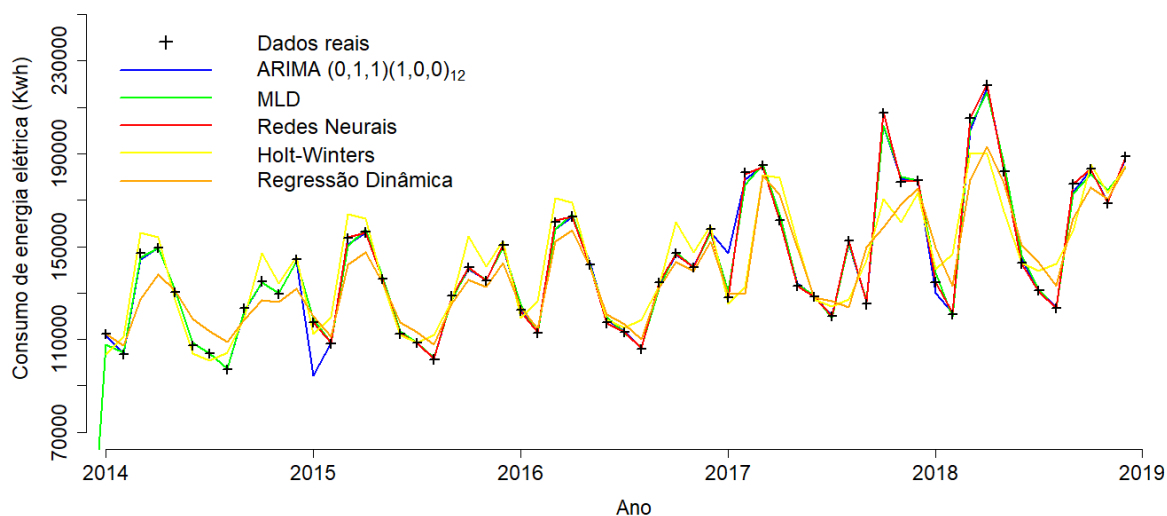


Figura 5.4 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018 para séries temporais do Cluster 2.

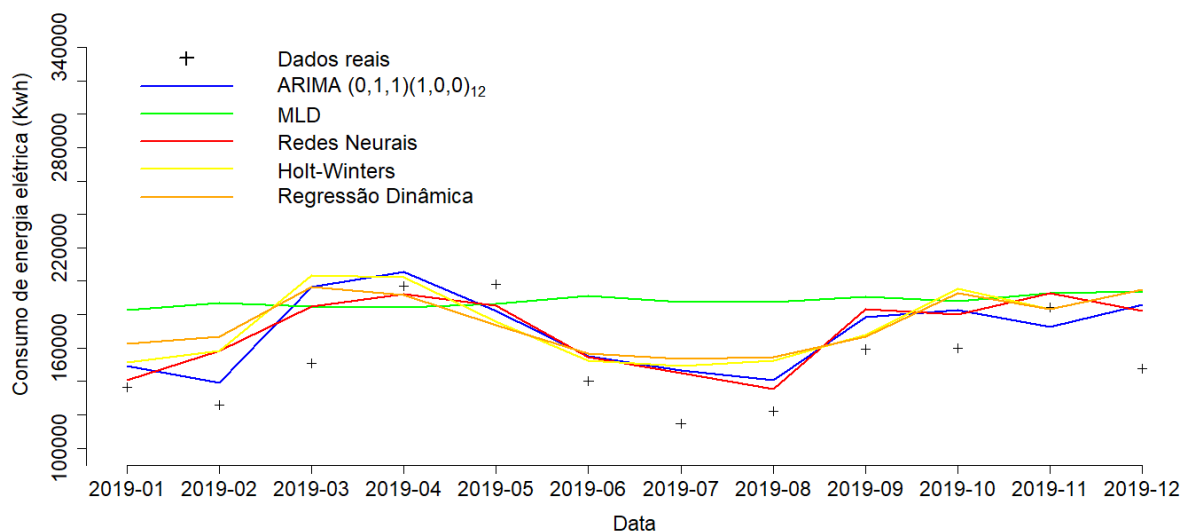


Figura 5.5 – Previsões dos modelos para o consumo de energia elétrica para o período de janeiro de 2019 a dezembro de 2019 para séries temporais do Cluster 2.

### 5.3.3 Previsões e ajustes para o Cluster 3

Nas Figuras 5.6 e 5.7 tem-se os ajustes e as previsões, respectivamente, dos modelos ARIMA  $(0, 1, 1)(1, 1, 0)_{12}$ , linear dinâmico, de redes neurais artificiais, de Holt-Winters aditivo e de regressão dinâmica para série temporal do consumo de energia elétrica proveniente do Cluster 3. O modelo de regressão dinâmica apresentou o melhor ajuste aos dados de consumo de energia elétrica provenientes do Cluster 3, enquanto a melhor previsão foi alcançada pelo modelo de redes neurais.

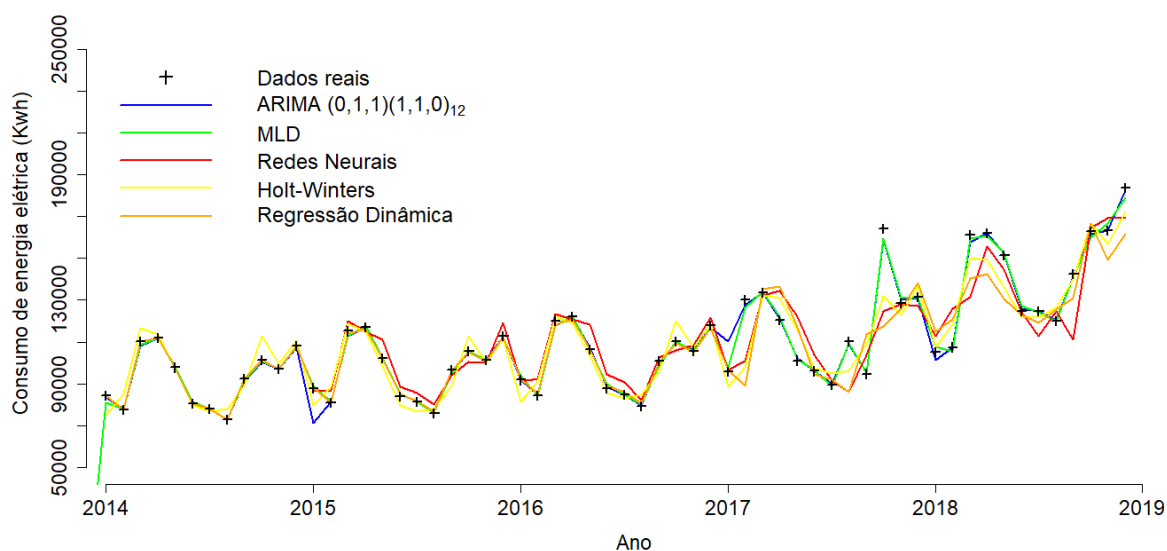


Figura 5.6 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a dezembro de 2018 para séries temporais do Cluster 3.

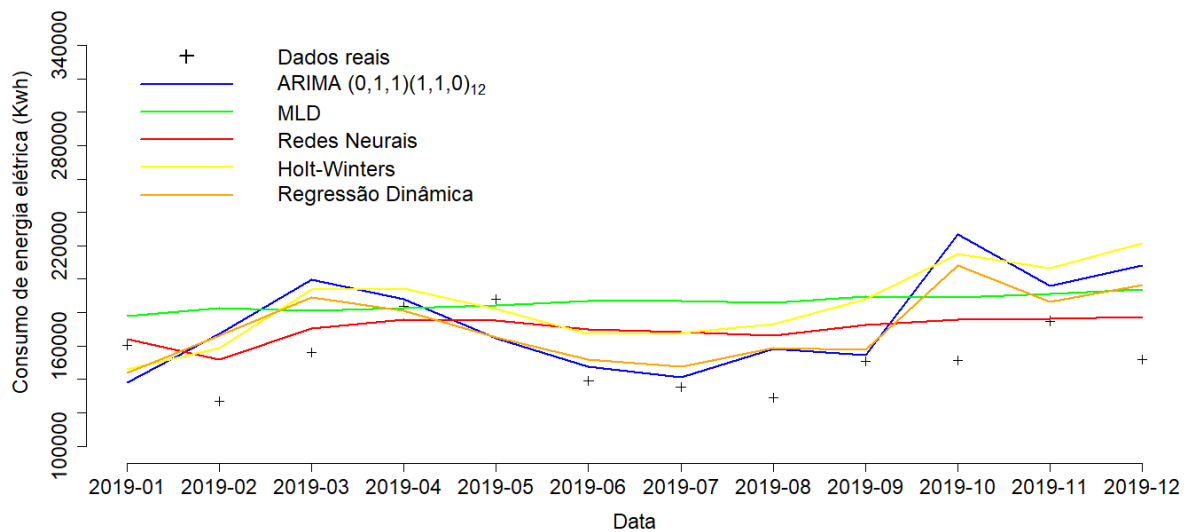


Figura 5.7 – Previsões dos modelos para o consumo de energia elétrica para o período de janeiro de 2019 a dezembro de 2019 para séries temporais do Cluster 3.

### 5.3.4 Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, modelo linear dinâmico, modelo de redes neurais artificiais e do modelo de regressão dinâmica.

A Tabela 5.3 apresenta a comparação entre os modelos utilizados neste trabalho para o Cenário 1. Nota-se que para o Cluster 1 o modelo de redes neurais artificiais tem o melhor desempenho, para o Cluster 2 o modelo de Box e Jenkins apresentou melhores resultados do que os demais, enquanto para o Cluster 3 a melhor previsão também é obtida pelo modelo de redes neurais artificiais.

Tabela 5.3 – Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, linear dinâmico, de redes neurais artificiais e o de regressão dinâmica, através da métrica MAPE para o Cenário 1.

Cluster	Modelo	MAPE (%)	MAPE (%)	MSE	MSE
		Ajuste	Previsão	Ajuste	Previsão
1	Box & Jenkins	0,45	1,47	0,0071	0,038
	Holt-Winters aditivo	0,43	1,30	0,0058	0,032
	Modelo linear dinâmico	0,19	1,70	0,0042	0,050
	Redes neurais	0,15	1,18	0,0060	0,026
	Regressão dinâmica	0,12	1,27	0,0012	0,032
2	Box & Jenkins	0,83	1,13	0,0184	0,020
	Holt-Winters aditivo	0,47	1,15	0,0071	0,033
	Modelo linear dinâmico	0,21	2,12	0,0043	0,083
	Redes neurais	0,14	1,15	0,0016	0,021
	Regressão dinâmica	0,11	2,15	0,0012	0,090
3	Box & Jenkins	0,39	1,39	0,0087	0,042
	Holt-Winters aditivo	0,45	1,88	0,0059	0,068
	Modelo linear dinâmico	0,21	1,79	0,0041	0,060
	Redes neurais	0,61	1,10	0,0101	0,021
	Regressão dinâmica	0,11	1,71	0,0012	0,032

### 5.3.5 Previsão do consumo de energia elétrica total da UFRRJ utilizando a abordagem bottom-up.

A Figura 5.8 apresenta os resultados da metodologia bottom-up aplicada aos clusters das séries temporais do consumo de energia elétrica da UFRRJ para o Cenário 1. O consumo de energia elétrica total previsto é o somatório da previsão do consumo do Cluster 1, Cluster 2 e Cluster 3 que foram obtidos pela aplicação do modelo de redes neurais artificiais ao Cluster 1, o modelo de Box e Jenkins ao Cluster 2 e o modelo de redes neurais artificiais ao Cluster 3, uma vez que estes modelos apresentaram os menores valores de MAPE para os respectivos clusters. O MAPE de previsão obtido pela abordagem bottom-up foi de aproximadamente 1%.

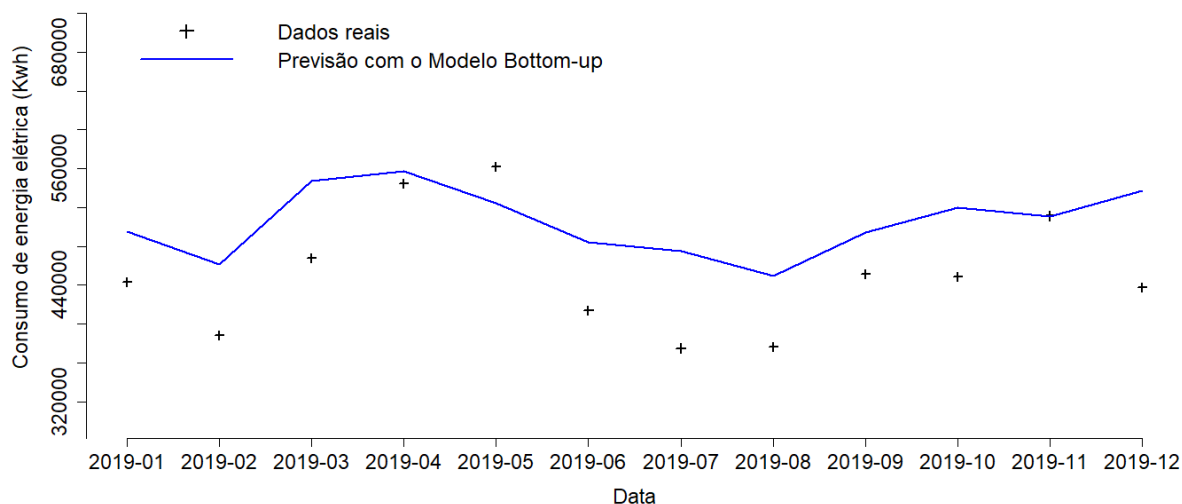


Figura 5.8 – Previsão do consumo de energia elétrica no período de janeiro de 2019 a dezembro de 2019 para a UFRRJ.

## 5.4 Cenário 2

Nesta seção serão apresentados os ajustes dos modelos no período de janeiro de 2014 a março de 2020, e as previsões serão realizadas para o período de abril de 2020 a maio de 2021. Neste período os dados de consumo de energia elétrica foram impactados pela pandemia de COVID-19 e, portanto, este Cenário é considerado durante a pandemia.

### 5.4.1 Previsões e ajustes para o Cluster 1

A Figura 5.9 mostra o ajuste do modelo ARIMA  $(0, 1, 1)(0, 1, 1)_{12}$ , modelo linear dinâmico, modelo de redes neurais artificiais, modelo de Holt-Winters aditivo e modelo de regressão dinâmica aos dados de consumo de energia elétrica para o Cluster 1. Nesta Figura, pode-se observar que os modelos conseguiram capturar a estrutura da série temporal (sazonalidade, tendência etc). Pela Tabela 5.4.4, o modelo de regressão dinâmica apresentou a melhor métrica de ajuste sendo um MAPE de 0,12%. A Figura 5.10 mostra a previsão dos modelos utilizados. O modelo de regressão dinâmica apresentou o melhor resultado para previsão do consumo de energia elétrica (Tabela 5.4.4).

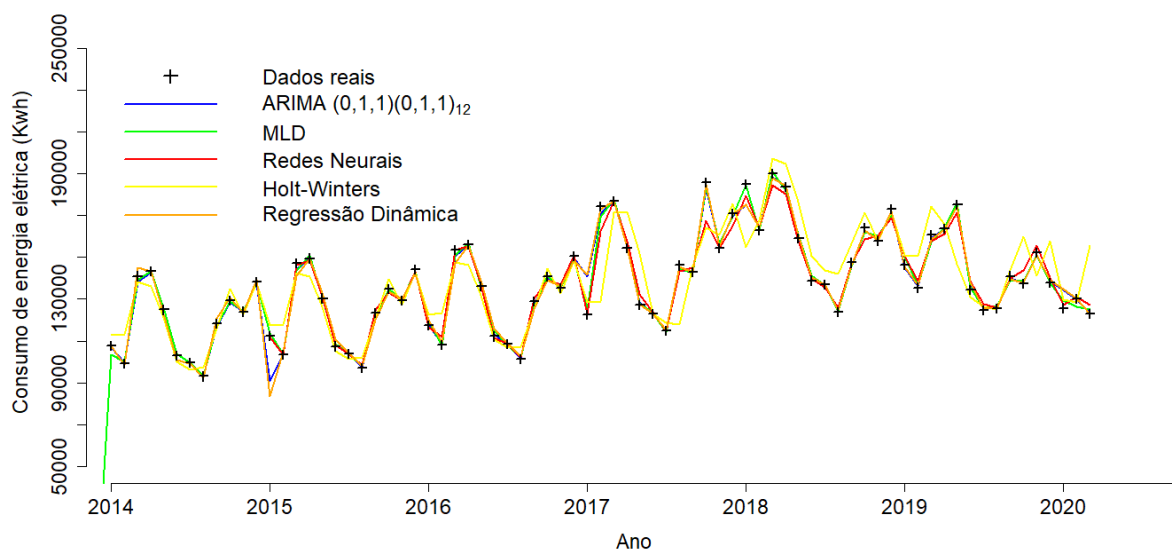


Figura 5.9 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020 para séries temporais do Cluster 1.

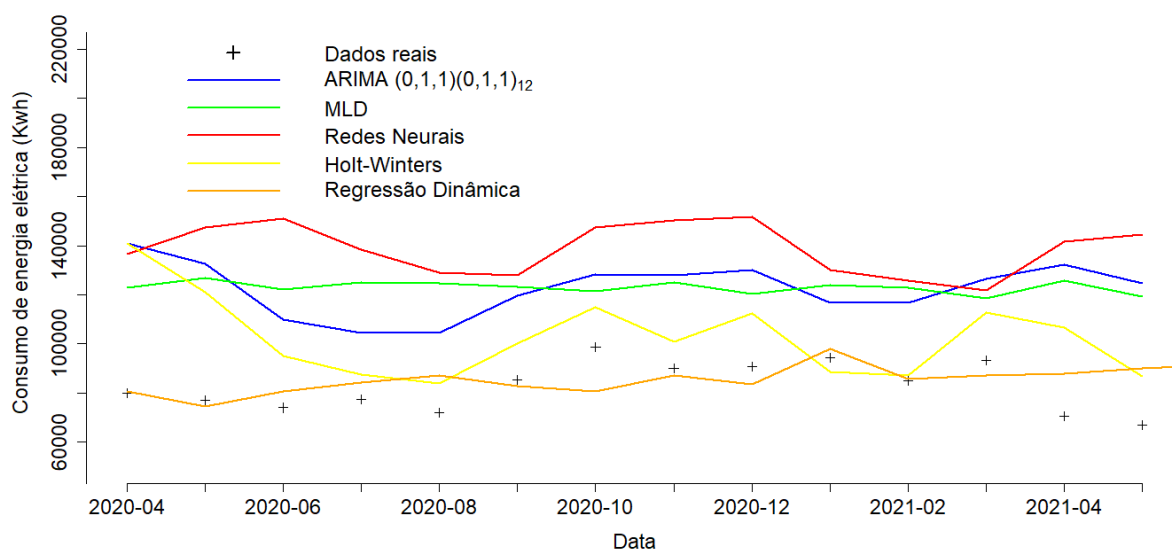


Figura 5.10 – Previsões dos modelos para o consumo de energia elétrica para o período de abril de 2020 a maio de 2021 para séries temporais do Cluster 1.

## 5.4.2 Previsões e ajustes para o Cluster 2

A Figura 5.11 mostra o ajuste do modelo  $ARIMA(0,1,1)(0,1,1)_{12}$ , modelo linear dinâmico, modelo de redes neurais artificiais, modelo de Holt-Winters aditivo e modelo de regressão dinâmica aos dados de consumo de energia elétrica para o Cluster 2. Nesta Figura, pode-se observar que os modelos conseguiram capturar a estrutura da série temporal (sazonalidade, tendência etc). Pela Tabela 5.4.4, o modelo de regressão dinâmica apresentou a melhor métrica de ajuste sendo um MAPE de 0,10%. A Figura 5.12 mostra a previsão dos modelos utilizados.

O modelo de regressão dinâmica apresentou o melhor resultado para previsão do consumo de energia elétrica (Tabela 5.4.4).

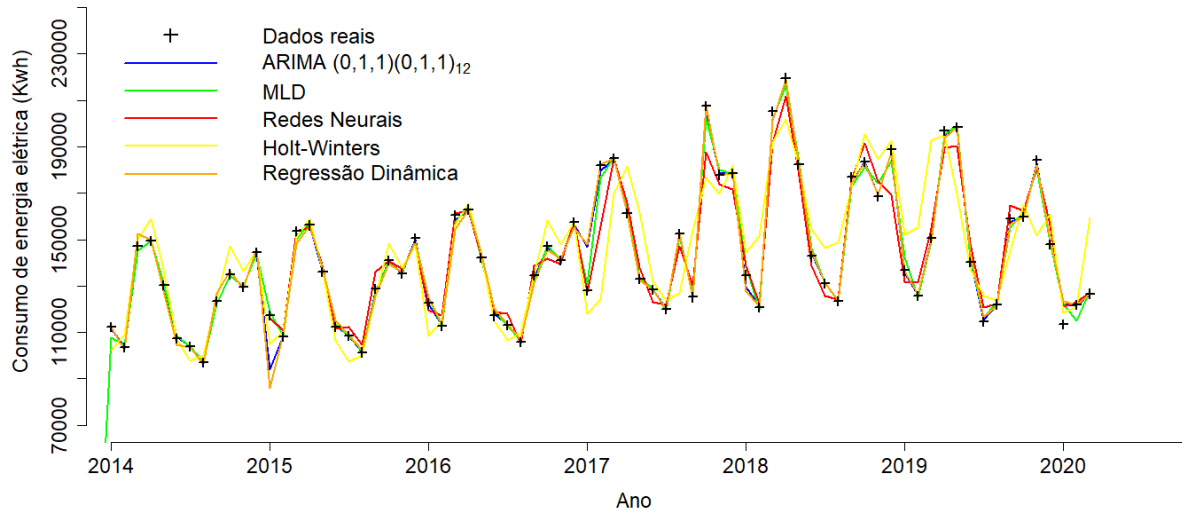


Figura 5.11 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020 para séries temporais do Cluster 2.

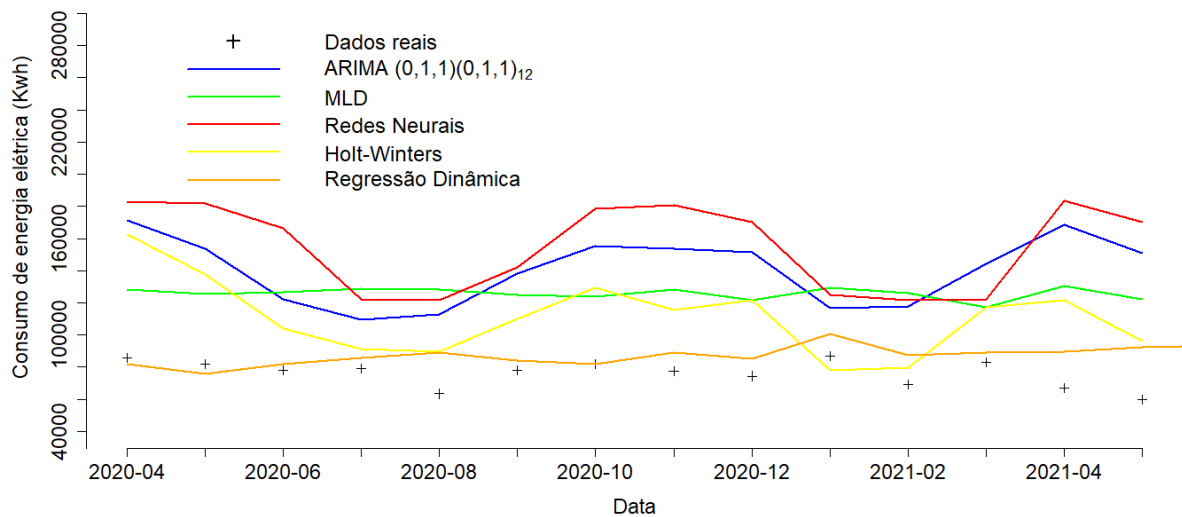


Figura 5.12 – Previsões dos modelos para o consumo de energia elétrica para o período de abril de 2020 a maio de 2021 para séries temporais do Cluster 2.

### 5.4.3 Previsões e ajustes para o Cluster 3

A Figura 5.13 mostra o ajuste do modelo ARIMA (0,0,2)(0,1,1)12, modelo linear dinâmico, modelo de redes neurais artificiais, modelo de Holt-Winters aditivo e modelo de regressão dinâmica aos dados de consumo de energia elétrica para o Cluster 3. Nesta Figura, pode-se



observar que os modelos conseguiram capturar a estrutura da série temporal (sazonalidade, tendência etc). Pela Tabela 5.4.4, o modelo de regressão dinâmica apresentou a melhor métrica de ajuste sendo um MAPE de 0,10%. A Figura 5.14 mostra a previsão dos modelos utilizados. O modelo de regressão dinâmica apresentou o melhor resultado para previsão do consumo de energia elétrica (Tabela 5.4.4).

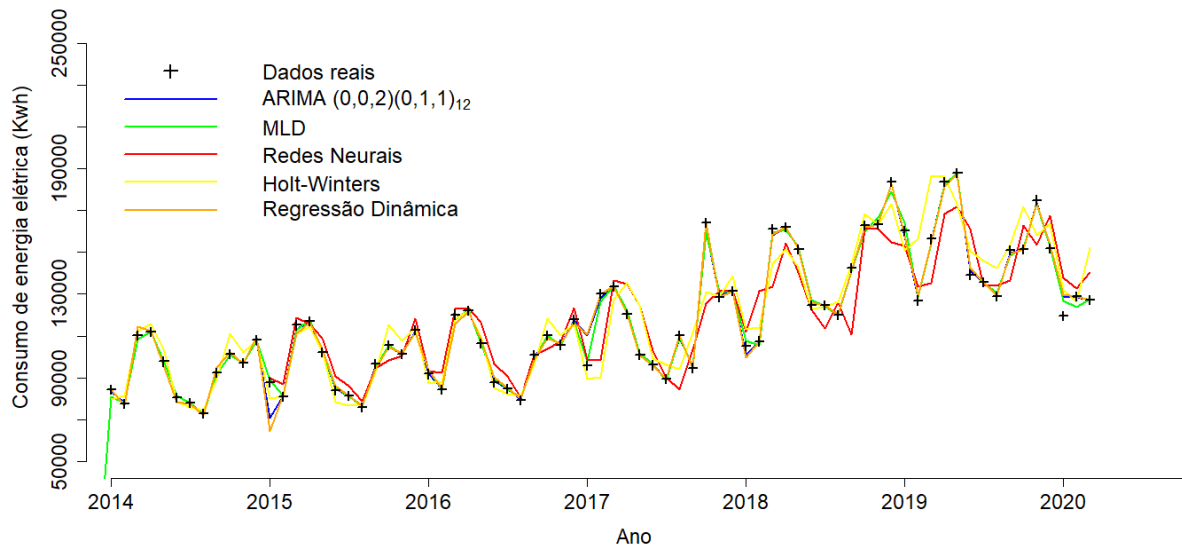


Figura 5.13 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a março de 2020 para séries temporais do Cluster 3.

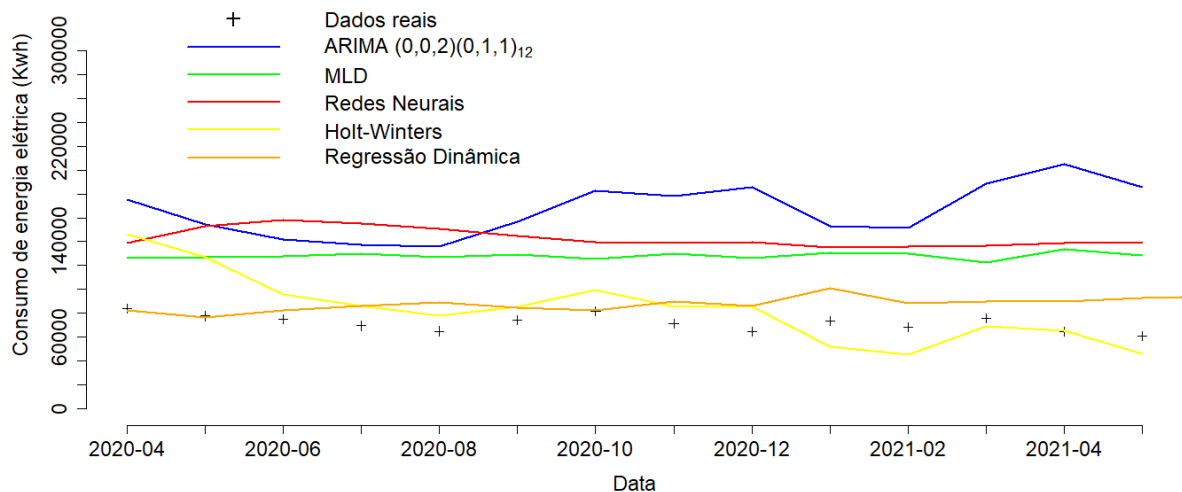


Figura 5.14 – Previsões dos modelos para o consumo de energia elétrica para o período de abril de 2020 a maio de 2021 para séries temporais do Cluster 3.

#### 5.4.4 Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, modelo linear dinâmico, modelo de redes neurais artificiais e do modelo de regressão dinâmica.

A Tabela 5.4.4 apresenta a comparação entre os modelos utilizados neste trabalho para o Cenário 2. Nota-se que tanto para o Cluster 1, Cluster 2 e Cluster 3 o modelo de regressão dinâmica tem o melhor desempenho quando comparado aos modelos de Box e Jenkins, Holt-winters aditivo, modelo linear dinâmico e o modelo de redes neurais artificiais. Os modelos utilizados neste estudo apresentaram um MAPE inferior a 1% para o ajuste e inferior a 8% para previsão. O modelo de regressão dinâmica apresentou as melhores métricas de MAPE e MSE (Tabela 5.4.4).

Tabela 5.4 – Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, linear dinâmico, de redes neurais artificiais e o de regressão dinâmica, através da métrica MAPE para o Cenário 2.

Cluster	Modelo	MAPE (%)	MAPE (%)	MSE	MSE
		Ajuste	Previsão	Ajuste	Previsão
1	Box & Jenkins	0,42	3,53	0,0093	0,176
	Holt-Winters aditivo	0,53	2,38	0,0077	0,094
	Modelo linear dinâmico	0,18	3,58	0,0034	0,177
	Redes neurais	0,13	4,65	0,0050	0,297
	Regressão dinâmica	0,11	0,88	0,0010	0,018
2	Box & Jenkins	0,56	5,41	0,0131	0,400
	Holt-Winters aditivo	0,56	3,79	0,0094	0,202
	Modelo linear dinâmico	0,21	4,50	0,0036	0,267
	Redes neurais	0,24	5,95	0,0017	0,534
	Regressão dinâmica	0,10	1,43	0,0010	0,041
3	Box & Jenkins	0,62	7,50	0,0100	0,728
	Holt-Winters aditivo	0,58	3,21	0,0087	0,148
	Modelo linear dinâmico	0,20	5,24	0,0033	0,353
	Redes neurais	0,66	6,22	0,0115	0,492
	Regressão dinâmica	0,10	1,91	0,0010	0,062

#### 5.4.5 Previsão do consumo de energia elétrica total da UFRRJ utilizando a abordagem bottom-up.

A Figura 5.15 apresenta os resultados da metodologia bottom-up aplicada aos clusters das séries temporais do consumo de energia elétrica da UFRRJ. O consumo de energia elétrica total previsto é o somatório da previsão do consumo do Cluster 1, Cluster 2 e Cluster 3 que foram obtidos pela aplicação do modelo de regressão dinâmica em cada cluster, uma vez que este modelo apresentou menor MAPE para os três clusters. O MAPE de previsão obtido pela abordagem bottom-up foi de 1,12%.

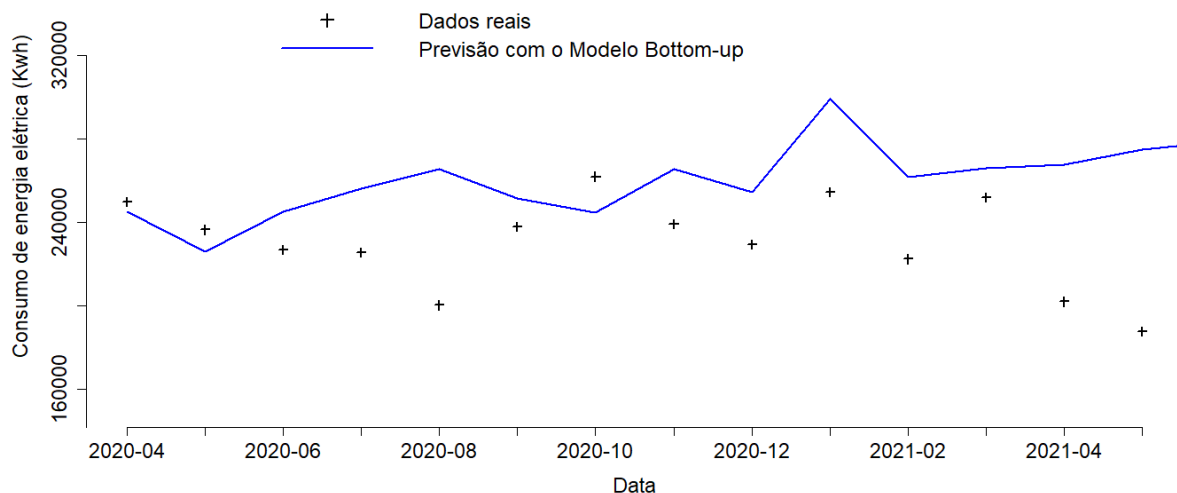


Figura 5.15 – Previsão do consumo de energia elétrica no período de abril de 2020 a maio de 2021 para a UFRRJ.

## 5.5 Cenário 3

Nesta seção serão apresentados os ajustes dos modelos no período de janeiro de 2014 a maio de 2021, e as previsões serão realizadas para o período de junho de 2021 a maio de 2026. Neste período os dados de consumo de energia elétrica foram influenciados pela pandemia de COVID-19 e a previsão foi realizada para o período em que as atividades de aulas presenciais na universidade foram retomadas, portanto devido ao horizonte de previsão este Cenário é considerado pós-pandemia. Ressalta-se que o fim da pandemia na Universidade é caracterizada pela volta às aulas presenciais e retomada dos servidores ao trabalho presencial.

### 5.5.1 Previsões e ajustes para o Cluster 1

Nas Figuras 5.16 e 5.17 tem-se os ajustes e as previsões, respectivamente, dos modelos ARIMA  $(0, 1, 1)(0, 1, 1)_{12}$ , linear dinâmico, de redes neurais artificiais, de Holt-Winters aditivo e de regressão dinâmica para série temporal do consumo de energia elétrica proveniente do Cluster 1. O melhor ajuste foi obtido pelo modelo de regressão dinâmica, e a previsão que consegue um comportamento conforme o esperado também foi alcançada pelo modelo de regressão dinâmica, pois espera-se que o consumo de energia elétrica na Universidade volte a crescer após o término da pandemia de COVID-19.

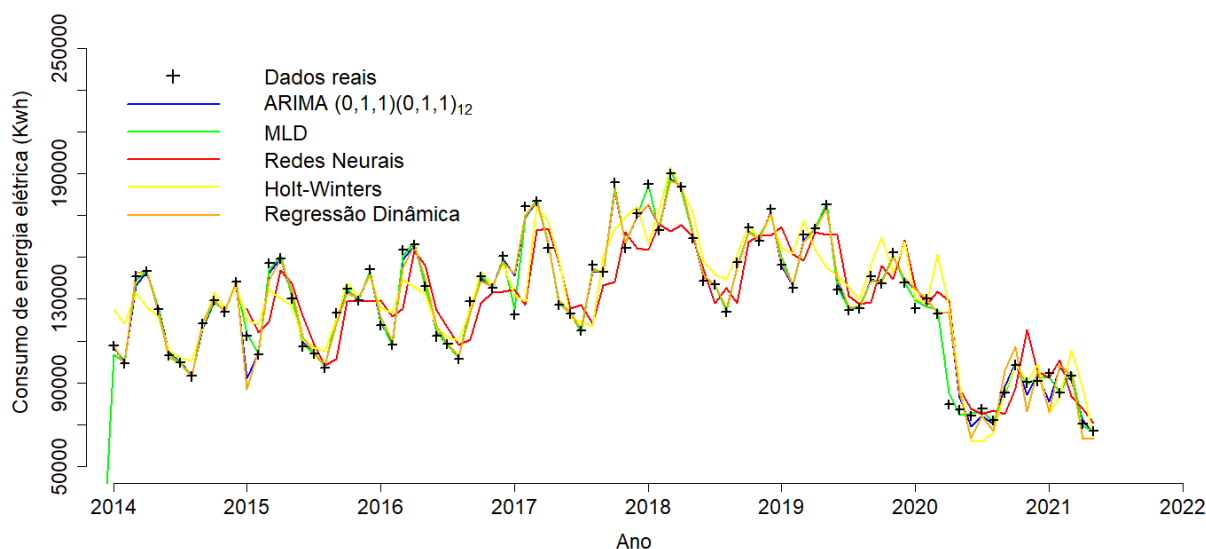


Figura 5.16 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021 para séries temporais do Cluster 1.

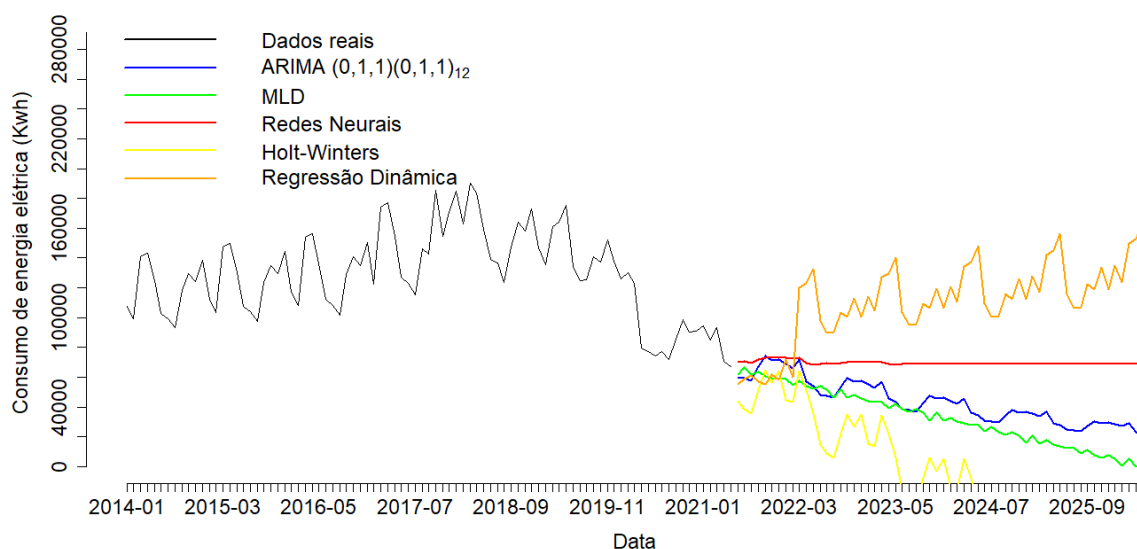


Figura 5.17 – Previsões dos modelos para o consumo de energia elétrica para o período de junho de 2021 a maio de 2026 para séries temporais do Cluster 1.

### 5.5.2 Previsões e ajustes para o Cluster 2

Nas Figuras 5.18 e 5.19 tem-se os ajustes e as previsões, respectivamente, dos modelos ARIMA  $(0, 1, 2)(1, 0, 2)_{12}$ , linear dinâmico, de redes neurais artificiais, de Holt-Winters aditivo e de regressão dinâmica para série temporal do consumo de energia elétrica proveniente do Cluster 2. O melhor ajuste foi obtido pelo modelo de regressão dinâmica, e a previsão que consegue um comportamento conforme o esperado também foi alcançada pelo modelo de regressão dinâmica, apresentando um aumento no consumo de energia elétrica após o término da pandemia

de COVID-19.

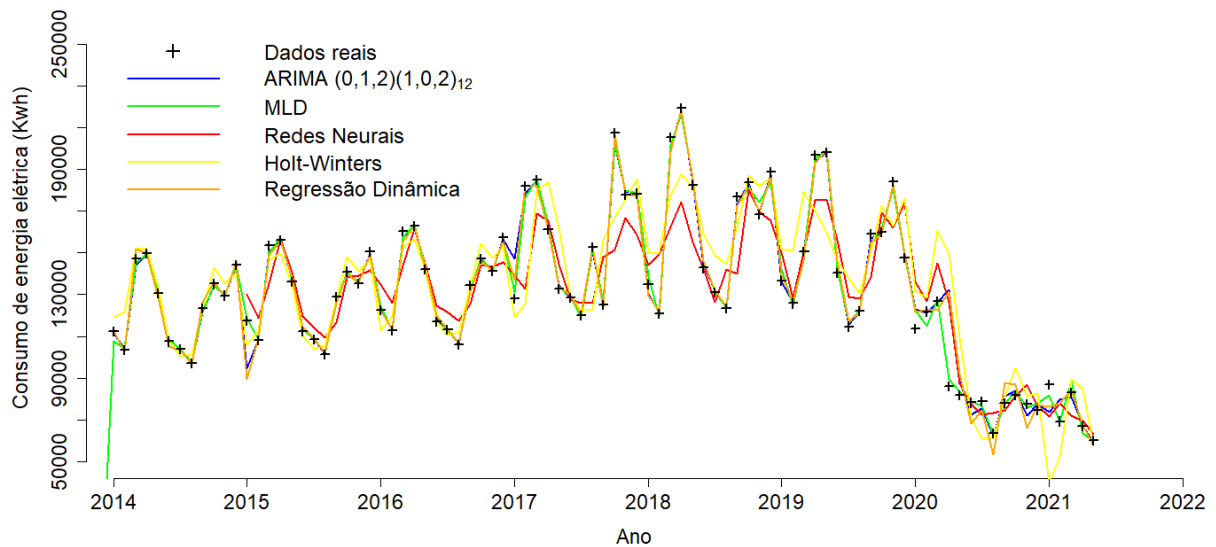


Figura 5.18 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021 para séries temporais do Cluster 2.

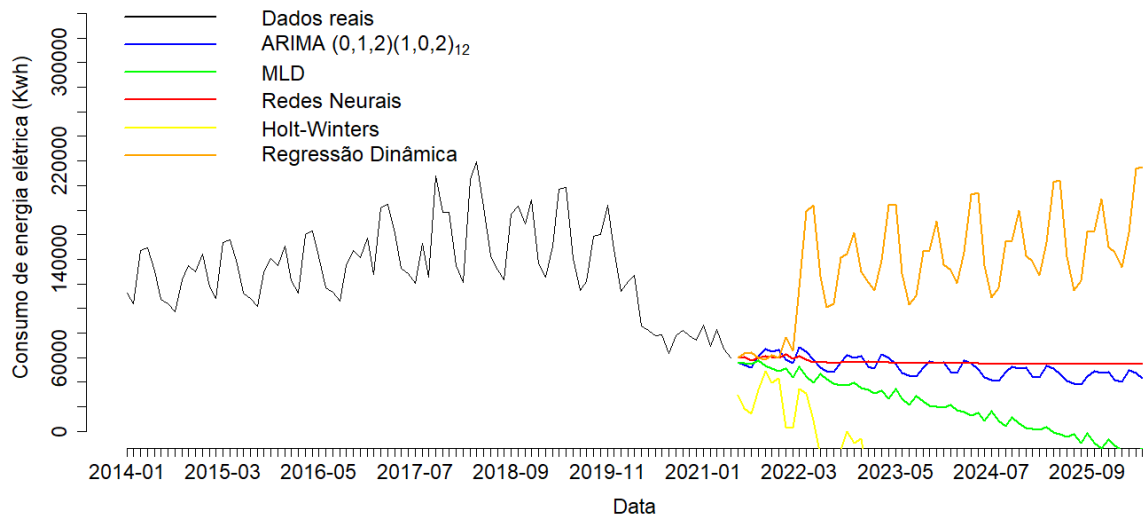


Figura 5.19 – Previsões dos modelos para o consumo de energia elétrica para o período de junho de 2021 a maio de 2026 para séries temporais do Cluster 2.

### 5.5.3 Previsões e ajustes para o Cluster 3

Nas Figuras 5.20 e 5.21 tem-se os ajustes e as previsões, respectivamente, dos modelos ARIMA (0, 1, 1)(0, 1, 1)<sub>12</sub>, linear dinâmico, de redes neurais artificiais, de Holt-Winters aditivo e de regressão dinâmica para série temporal do consumo de energia elétrica proveniente do Cluster 3. O melhor ajuste foi obtido pelo modelo de regressão dinâmica, e a previsão que consegue

um comportamento conforme o esperado também foi alcançada pelo modelo de regressão dinâmica, apresentando um aumento no consumo de energia elétrica após o término da pandemia de COVID-19.

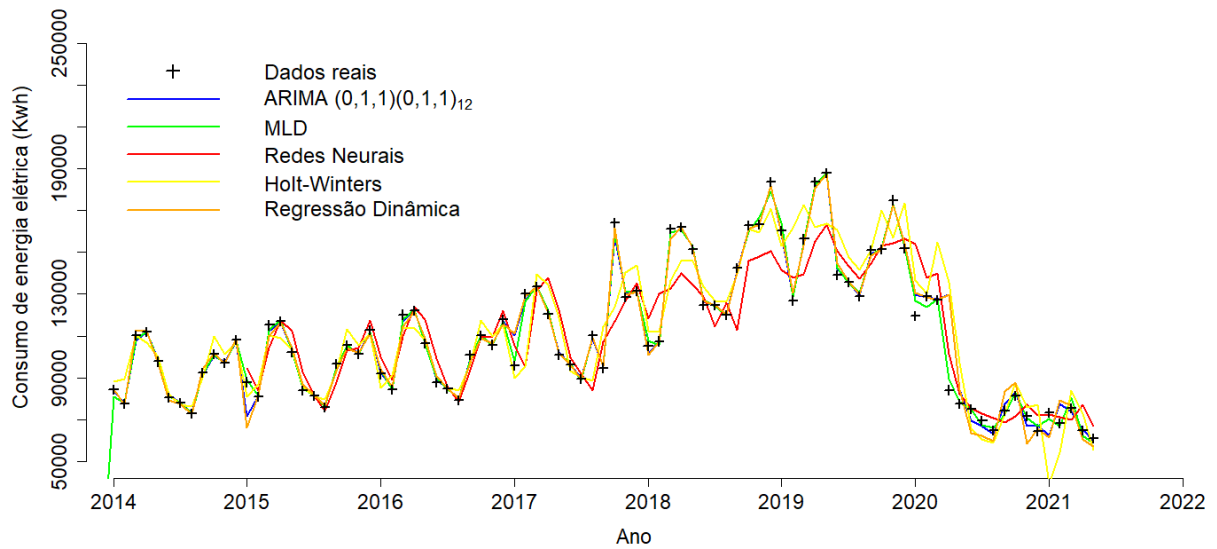


Figura 5.20 – Ajustes dos modelos aos dados de consumo de energia elétrica no período de janeiro de 2014 a maio de 2021 para séries temporais do Cluster 3.

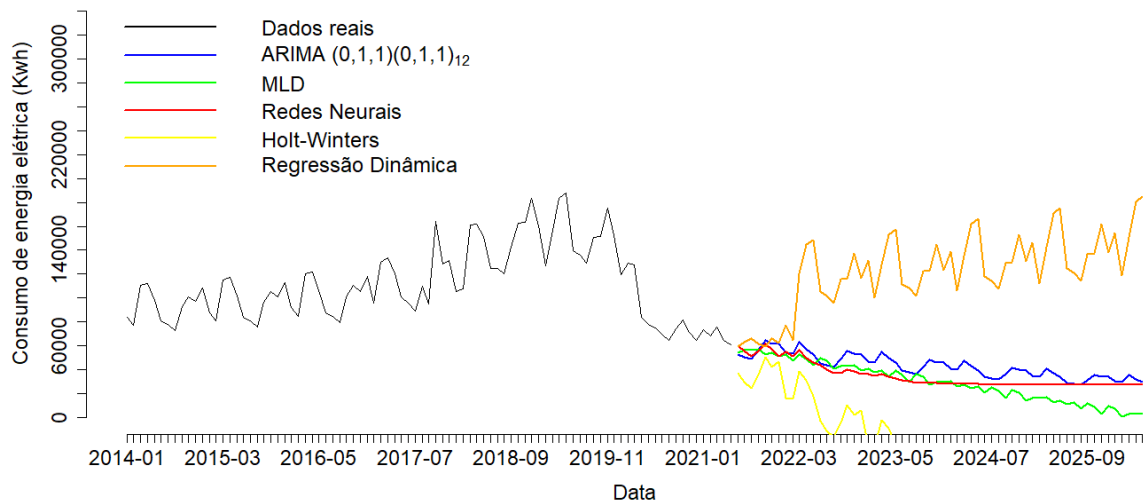


Figura 5.21 – Previsões dos modelos para o consumo de energia elétrica para o período de junho de 2021 a maio de 2026 para séries temporais do Cluster 3.

#### 5.5.4 Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, modelo linear dinâmico, modelo de redes neurais artificiais e do modelo de regressão dinâmica.

A Tabela 5.5 apresenta a comparação entre os modelos utilizados neste trabalho para o Cenário 3. Nota-se que tanto para o Cluster 1, Cluster 2 e Cluster 3 o modelo de regressão dinâmica tem o melhor ajuste quando comparado aos modelos de Box e Jenkins, Holt-winters aditivo, linear dinâmico e de redes neurais artificiais. Porém, como pode ser observado nas Figuras 5.17, 5.19 e 5.21, a previsão obtida pelo modelo de regressão dinâmica conseguiu recuperar o comportamento de consumo de energia elétrica da Universidade após considerar o fim da pandemia do COVID-19 em março de 2022. Assim, este modelo conseguiu restabelecer o consumo de energia elétrica após o fim da pandemia aos níveis de consumo do período antes da pandemia. O término da pandemia do COVID-19 é caracterizada pela retomada das aulas presenciais e retomada dos servidores em trabalho presencial, ambos em março de 2022.

Tabela 5.5 – Comparação entre os modelos de Box e Jenkins, Holt-Winters aditivo, linear dinâmico, de redes neurais artificiais e o de regressão dinâmica, através da métrica MAPE para o Cenário 3.

Cluster	Modelo	MAPE (%)	MSE
		Ajuste	Ajuste
1	Box & Jenkins	0,52	0,0124
	Holt-Winters aditivo	0,58	0,0120
	Modelo linear dinâmico	0,37	0,0069
	Redes neurais	0,84	0,0143
	Regressão dinâmica	0,23	0,0037
2	Box & Jenkins	0,91	0,0194
	Holt-Winters aditivo	0,74	0,0156
	Modelo linear dinâmico	0,51	0,0081
	Redes neurais	0,91	0,0149
	Regressão dinâmica	0,22	0,0038
3	Box & Jenkins	0,63	0,0153
	Holt-Winters aditivo	0,75	0,0142
	Modelo linear dinâmico	0,35	0,0078
	Redes neurais	0,79	0,0149
	Regressão dinâmica	0,22	0,0038

#### 5.5.5 Previsão do consumo de energia elétrica total da UFRRJ utilizando a abordagem bottom-up.

A Figura 5.22 apresenta os resultados da metodologia bottom-up aplicada aos clusters das séries temporais do consumo de energia elétrica da UFRRJ. O consumo de energia elétrica total previsto é o somatório da previsão do consumo do Cluster 1, Cluster 2 e Cluster 3 que foram obtidos pela aplicação do modelo de regressão dinâmica em cada cluster, uma vez que este modelo conseguiu descrever o comportamento da série temporal dos dados no período investigado. Pode-se observar nesta figura que a partir de março de 2022, período considerado como volta das atividades presenciais na Universidade, que o consumo de energia elétrica foi

recuperado através da metodologia bottom-up que utilizou o modelo de regressão dinâmica nas previsões dos clusters.

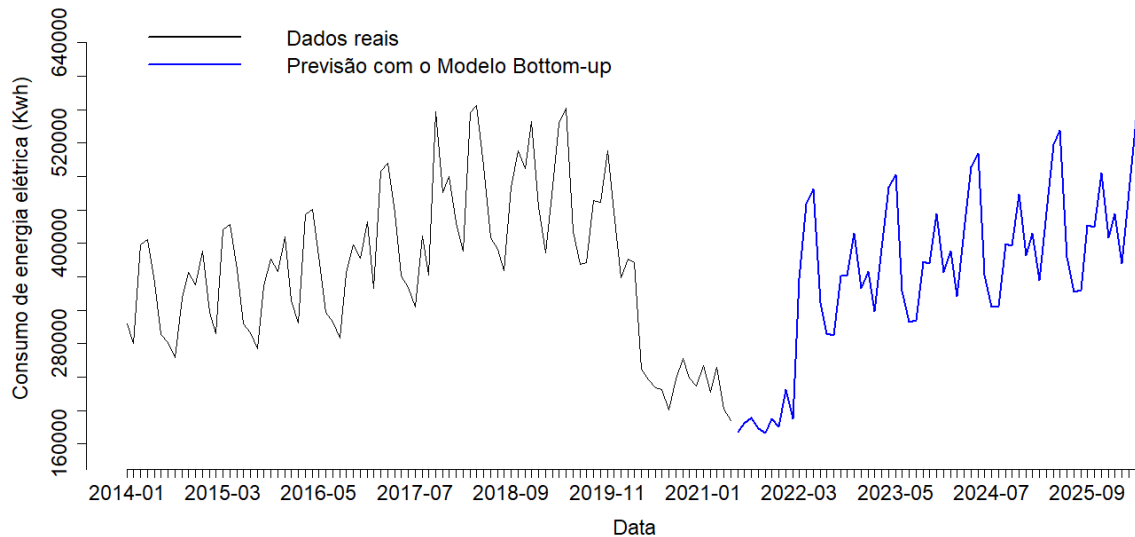


Figura 5.22 – Previsão do consumo de energia elétrica no período de junho de 2021 a maio de 2026 para a UFRRJ.

## 5.6 Implementação de medidas de eficiência energética durante o horizonte de previsão do cenário 3

A Figura 5.23 mostra as curvas de difusão para as medidas de eficiência energética. Estas curvas mostra como as MEE foram inseridas no horizonte de previsão.

A Figura 5.24 traz a previsão do consumo de energia elétrica da UFRRJ para quatro contextos de implementação e difusão de medidas de eficiência energética, sendo eles: Cenário Congelado, Cenário mínimo, Cenário Médio e Cenário Máximo. O Cenário Congelado é o cenário onde não há implementação de nenhuma medida de eficiência energética e, por outro lado, os Cenários mínimo, Cenário Médio e Cenário Máximo possuem a implementação de 50% , 75% e 100% de medidas de eficiência energética no consumo de energia elétrica da Universidade, respectivamente. Nesta figura pode-se observar uma significativa redução do consumo de energia elétrica com a implementação das medidas de eficiência energética. Os valores desta previsão são apresentados na Tabela 5.7.

A Figura 5.25 mostra a economia de energia elétrica anual nos diferentes cenários de difusão das medidas de eficiência energética. Para o horizonte de previsão, a economia gerada no Cenário Mínimo, Cenário Médio e Cenário Máximo foi de 2.758.680 kWh, 3.883.005 kWh e 5.373.199 kWh, respectivamente, conforme Tabela 5.6.

Fazendo uma breve análise financeira, considerando que o valor do kWh é aproximadamente R\$ 1,00, a economia financeira no decorrer dos cinco anos de implementação das medidas de eficiência energética seria de R\$ 2.758.680,00, R\$ 3.883.005,00 e R\$ 5.373.199,00, respectivamente para os Cenário Mínimo, Cenário Médio e Cenário Máximo. Além disso, considerando que os gastos mensais da Universidade com energia elétrica é de aproximadamente R\$ 700.000,00, a economia mensal após a implementação das medidas de eficiência energética para os Cenário Mínimo, Cenário Médio e Cenário Máximo seria de R\$129.500,00, R\$



194.250,00 e R\$ 259.000,00, na devida ordem.

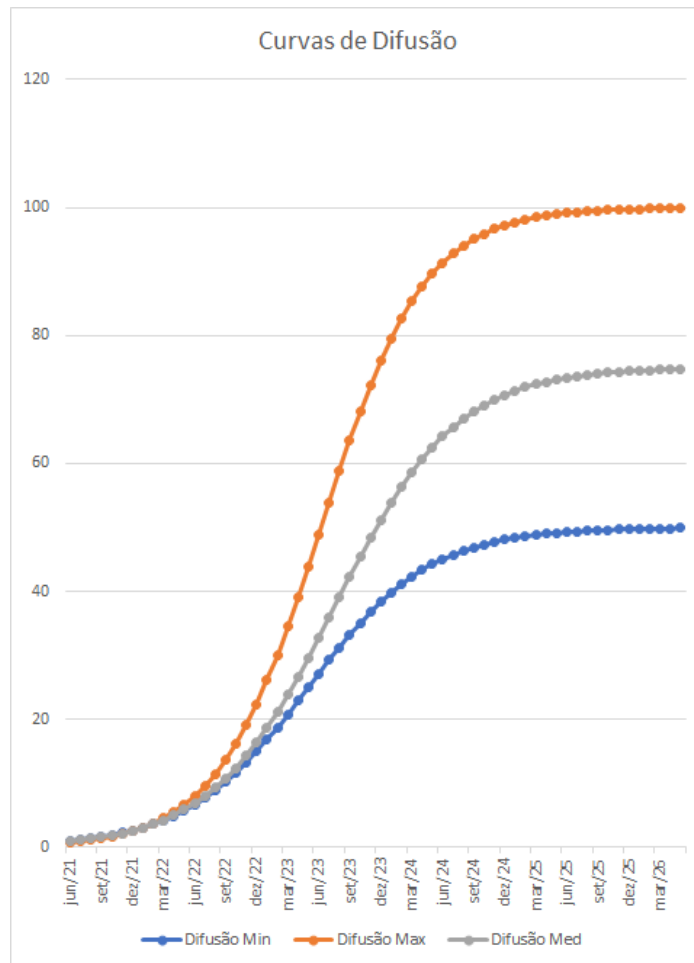


Figura 5.23 – Curvas de difusão das medidas de eficiência energética no período de junho de 2021 a maio de 2026 na UFRRJ.

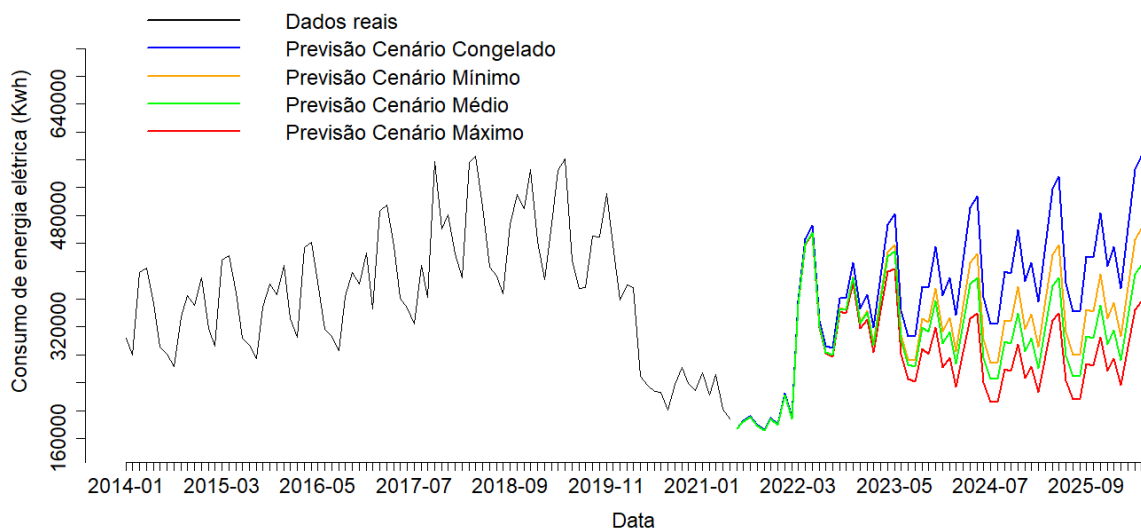


Figura 5.24 – Previsão do consumo de energia elétrica no período de junho de 2021 a maio de 2026 para os diferentes cenários de difusão de medidas de eficiência energética na UFRRJ.

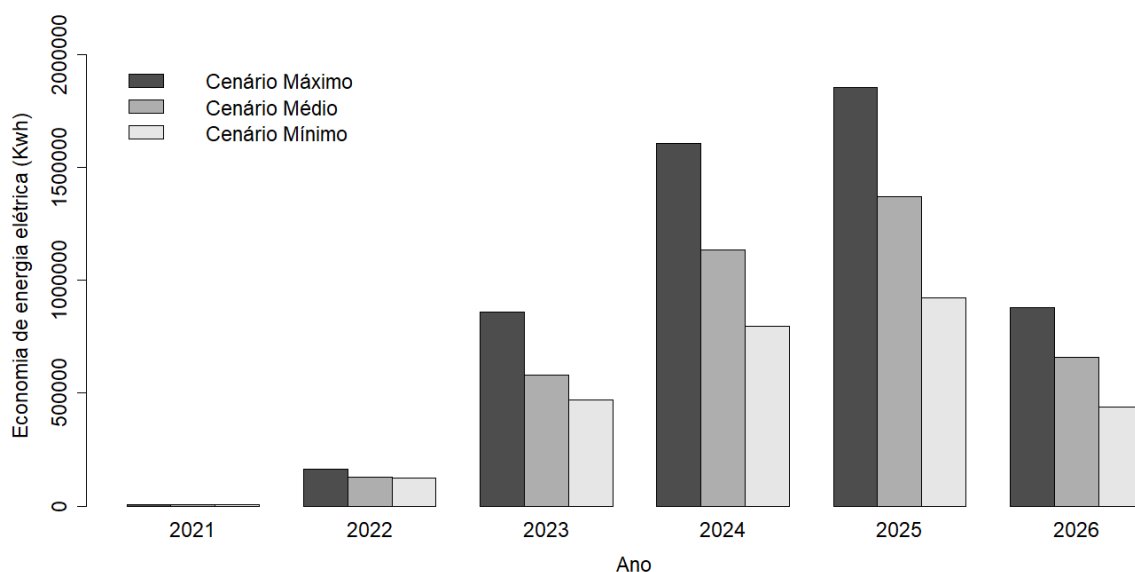


Figura 5.25 – Economia anual de energia elétrica em kWh para os diferentes cenários de difusão de medidas de eficiência energética para o horizonte de previsão de junho de 2021 a maio de 2026.

Tabela 5.6 – Economia anual de energia elétrica em kWh para os diferentes cenários de difusão de medidas de eficiência energética no horizonte de previsão de junho de 2021 a maio de 2026.

Ano	Economia de energia elétrica (kWh)		
	Cenário Mínimo	Cenário Médio	Cenário Máximo
2021	8.038,6	7.969,9	7.201,2
2022	123.762,4	130.137,8	162.017,4
2023	469.235,7	580.209,7	860.269,2
2024	795.675,4	1.133.800,9	1.608.017,7
2025	922.223,1	1.372.345,3	1.854.756,1
2026	439.744,8	658.541,4	880.937,4
<b>Total</b>	<b>2.758.680</b>	<b>3.883.005</b>	<b>5.373.199</b>

Tabela 5.7: Previsão do consumo de energia elétrica no período de junho de 2021 a maio de 2026 para os diferentes cenários de difusão de medidas de eficiência energética na UFRRJ.

Ano	Consumo de energia elétrica (kWh)			
	Cenário Congelado	Cenário Mínimo	Cenário Médio	Cenário Máximo
jun/21	174276	173641	173651	173771
jul/21	185131	184334	184344	184477
ago/21	191674	190701	190712	19085
set/21	178525	177455	177466	177589
out/21	172609	171391	171401	171507
nov/21	189811	188233	188243	188336

*Continua na próxima página*

Tabela 5.7 – Continuação da tabela

Ano	Consumo de energia elétrica (kWh)			
	Cenário Congelado	Cenário Mínimo	Cenário Médio	Cenário Máximo
dez/21	180569	178805	178811	178864
jan/22	224889	222309	222309	222309
fev/22	190578	188016	188007	187926
mar/22	35787	352244	352198	351838
abr/22	446738	438543	438433	437634
mai/22	465555	455614	455422	454105
jun/22	329357	321196	320982	319607
jul/22	292268	283892	283606	281885
ago/22	289654	280089	279676	277347
set/22	360659	346995	346265	342406
out/22	360721	345114	344099	339078
nov/22	411917	391671	390091	382777
dez/22	346534	327292	325513	317808
jan/23	366539	343686	341211	331182
fev/23	31859	296429	293646	283102
mar/23	393711	363364	358988	343492
abr/23	467073	427459	420963	399467
mai/23	481848	437202	428953	403452
jun/23	343261	308763	301646	281099
jul/23	306151	273016	265452	245055
ago/23	307856	272226	263303	240823
set/23	378235	331738	319071	289239
out/23	376216	327407	313059	28144
nov/23	435088	375876	357241	318752
dez/23	365156	313319	295979	26235
jan/24	390831	333254	312924	275811
fev/24	336539	285331	266366	233683
mar/24	414834	349911	324837	28392
abr/24	491555	412726	381148	33218
mai/24	508168	424942	390523	339627
jun/24	362651	302173	276461	24007
jul/24	323685	268861	244991	212534
ago/24	325477	269611	244783	212234
set/24	399375	330042	298682	258902
out/24	397254	327619	295644	256265
nov/24	459071	377934	340194	294929
dez/24	385641	317004	284725	246908
jan/25	412599	338725	303657	263418
fev/25	355594	291604	260986	226491
mar/25	437803	358683	320567	278311
abr/25	518361	424342	378789	328992
mai/25	535804	438323	390862	339613
jun/25	383011	313147	278992	242503

*Continua na próxima página*

Tabela 5.7 – *Continuação da tabela*

Ano	Consumo de energia elétrica (kWh)			
	<b>Cenário Congelado</b>	<b>Cenário Mínimo</b>	<b>Cenário Médio</b>	<b>Cenário Máximo</b>
jul/25	342097	279559	248879	216404
ago/25	343978	280979	249982	217434
set/25	421571	344241	306096	266321
out/25	419344	342321	304247	264783
nov/25	484251	395206	351111	305642
dez/25	407151	332212	295047	256893
jan/26	435456	355243	315413	274677
fev/26	375601	306367	271951	236869
mar/26	461921	376727	334341	291252
abr/26	546506	445664	395452	344533
mai/26	564821	460558	408608	356036

*Fim da tabela*

Neste trabalho, foi proposta uma metodologia para obter a previsão do consumo de energia elétrica e analisar os efeitos das medidas de eficiência energética no horizonte de previsão, aplicado ao setor público do sistema elétrico brasileiro, em particular para UFRRJ. Essa metodologia combinou os modelos de séries temporais, análise de cluster e abordagem bottom-up.

Com a aplicação da metodologia de clusterização foi possível analisar apenas três séries temporais agregadas e não mais as 20 séries dos setores, o que resultou em um ganho de tempo para análise e previsão das séries temporais do consumo de energia elétrica.

A metodologia foi testada para os dados da UFRRJ, considerando três cenários com horizontes de previsão diferentes. O Cenário 1 possui horizonte de previsão de 12 meses à frente, a partir dos dados de janeiro de 2014 até dezembro de 2018. O Cenário 2 possui horizonte de previsão de 14 meses à frente, a partir dos dados de janeiro de 2014 até março de 2020. E finalmente, o Cenário 3 possui horizonte de previsão de 60 meses (5 anos) à frente, a partir dos dados de janeiro de 2014 até maio de 2021.

O Cenário 1 não sofreu a interferência da pandemia nos dados. Os resultados da métrica MAPE da Tabela 5.3 mostra que para o cluster 1 o modelo de redes neurais artificiais tem o melhor desempenho, para o cluster 2 os modelos de Box e Jenkins sobressai aos outros modelos, enquanto para o cluster 3 a melhor previsão também é conseguida pelo modelo de redes neurais artificiais, conforme destacado na seção 5.3.4.

Por outro lado, o Cenário 2 realiza a previsão do consumo de energia elétrica na UFRRJ durante a pandemia de COVID-19. Neste Cenário, os resultados da métrica MAPE da Tabela 5.4.4 apontam que tanto para o cluster 1, cluster 2 e cluster 3 o modelo de regressão dinâmica tem o melhor desempenho quando comparado aos modelos de Box e Jenkins, Holt-winters aditivo, modelo linear dinâmico e o modelo de redes neurais artificiais, conforme seção 5.4.4. Este bom desempenho, do modelo de regressão dinâmica, na previsão durante a pandemia deve-se a utilização da variável regressora que traz as características do comportamento presencial na UFRRJ durante a pandemia.

No Cenário 3 a previsão do consumo de energia elétrica na UFRRJ é realizada durante uma parte da pandemia e, grande parte para o período pós-pandemia. Da mesma forma que ocorreu no Cenário 2, no presente Cenário o modelo que conseguiu realizar a melhor previsão para o consumo de energia elétrica de todos os Clusters é o modelo de regressão dinâmica, conforme pode ser observado nas Figuras 5.17, 5.19 e 5.21.

A previsão do consumo de energia elétrica total da UFRRJ obtido pela metodologia bottom-up, através dos valores previstos pelos modelos de séries temporais adotados para os clusters obtidos, mostrou-se satisfatória apresentando um MAPE de 1% e 1,12% para o Cenário 1 e para o Cenário 2, respectivamente.

A implementação de medidas de eficiência energética na UFRRJ mostrou-se com grande potencial para a redução nos gastos da Universidade com energia elétrica. Durante o período de implementação das MEE a economia seria de 2.758.680 kWh, 3.883.005 kWh e 5.373.199 kWh para o Cenário Mínimo, Cenário Médio e Cenário Máximo, de modo respectivo, conforme salientado na seção 5.6. A economia percentual nos gastos mensais com energia elétrica após a implementação das MEE seria de 18,5% no Cenário Mínimo, 27,75% no Cenário Médio e 37% no Cenário Máximo.

Uma questão interessante para pesquisas futuras seria realizar uma análise de investimentos para a implementação das medidas de eficiência energética neste seguimento do setor público. Também seria notável a inclusão de outras medidas de eficiência energética ao modelo.

## Referências bibliográficas

- 1 HAHN, H.; MEYER-NIEBERG, S.; PICKL, S. Electric load forecasting methods: tools for decision making. *Eur J Oper Res*, 2009;199(3):902e7. <https://doi.org/10.1016/j.ejor.2009.01.062>.
- 2 EMPRESA DE PESQUISA ENERGÉTICA (EPE). *ESTUDOS DA DEMANDA DE ENERGIA: Demanda de Energia 2050*. 2016. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-227/topico-458/DEA%2013-15%20Demanda%20de%20Energia%202050.pdf>. Acesso em: 16 fev. 2022.
- 3 SILVA, F. L. C. et al. A bottom-up methodology for long term electricity consumption forecasting of an industrial sector – Application to pulp and paper sector in Brazil, *Energy*, 144:1107–1118, 2018. <https://doi.org/10.1016/j.energy.2017.12.078>.
- 4 MAÇAIRA, Paula Medina. *Modelagem e previsão das séries de consumo de energia elétrica no Brasil com métodos de suavização exponencial de Pegels e abordagem bottom up por uso final*. Rio de Janeiro, 2015. Dissertação (Programa de Pós Graduação em Engenharia Elétrica) - Pontifícia Universidade Católica do Rio De Janeiro.
- 5 ZELLNER, A.; TOBIAS, J. A note on aggregation, disaggregation and forecasting performance. *J. Forecasting*, 19:457-69, 2000. [https://doi.org/10.1002/1099-131X\(200009\)19:5<457::AID-FOR761>3.0.CO;2-6](https://doi.org/10.1002/1099-131X(200009)19:5<457::AID-FOR761>3.0.CO;2-6)
- 6 HYNDMAN, R. et al. Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis*, 2011;55:2579e89. <https://doi.org/10.1016/j.csda.2011.03.006>
- 7 FLEITER, T. et al. Energy efficiency in the german pulp and paper industry – a model-based assessment of saving potentials. *Energy*, v. 40, p. 84–99, 2012.
- 8 HERBST, A. et al. Introduction to energy systems modelling. *Swiss Journal of Economics and statistic*, v. 148, p. 111–135, 2012.
- 9 CHATEAU, B.; LAPILLONNE, B. Long-term energy demand forecasting a new approach. *Energy Pol*, 1978;6:140e57. [https://doi.org/10.1016/0003-4916\(63\)90068-X](https://doi.org/10.1016/0003-4916(63)90068-X).
- 10 GEROSKI P. A. Models of technology diffusion. *Res Pol*, 2000;29:603e25. [https://doi.org/10.1016/S0048-7333\(99\)00092-X](https://doi.org/10.1016/S0048-7333(99)00092-X)
- 11 FLIEDNER, G. Hierarchical forecasting: issues and use guidelines. *Industrial Management & Data Systems*, v. 101, p. 5–12, 2001.
- 12 KINNEY, W. R. Predicting earnings: entity versus sub-entity data. *Journal of Accounting Research*, v. 9, p. 127–136, 1971.

- 13 COLLINS, D. W. Predicting earnings with sub-entity data: some further evidence. *Journal of Accounting Research*, v. 14, p. 163–177, 1976.
- 14 DUNN, D. M.; WILLIAMS, W. H.; DECHAINED, T. L. Aggregate versus sub aggregate models in local area forecasting. *Journal of the American Statistical Association*, v. 71, p. 68–71, 1976.
- 15 DANGERFIELD, B. J.; MORRIS, J. S. Top-down or bottom-up: aggregate versus disaggregate extrapolations. *International Journal of Forecasting*, v. 8, p.233–241, 1992.
- 16 JACOBSEN, H. K. Integrating the bottom-up and top-down approach to energyeconomy modelling: the case of Denmark. *Energy Econ.* 1998;20:443e61. [https://doi.org/10.1016/S0140-9883\(98\)00002-4](https://doi.org/10.1016/S0140-9883(98)00002-4).
- 17 KOOPMANS, C.C.; VELDE, D.W. Bridging the energy efficiency gap: using bottomup information in a top-down energy demand model. *Energy Econ.* 2001;23: 57e75. [https://doi.org/10.1016/S0140-9883\(00\)00054-2](https://doi.org/10.1016/S0140-9883(00)00054-2).
- 18 SWAN, L. G.; UGURSAL, V. I. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. *Renew Sustain Energy.* Rev 2009;13:1819e35. <https://doi.org/10.1016/j.rser.2008.09.033>.
- 19 HALL, L. M.; BUCKLEY, A. R. A review of energy systems models in the UK: prevalent usage and categorisation. *Appl Energy.* 2016;169(Supplement C):607e28. <https://doi.org/10.1016/j.apenergy.2016.02.044>.
- 20 GHEDAMSI, R. et al. Modeling and forecasting energy consumption for residential buildings in Algeria using bottom-up. *Energy Build.* 2016;121:309e17. <https://doi.org/10.1016/j.enbuild.2015.12.030>.
- 21 GEORGIU, P. N. A bottom-up optimization model for the long-term energy planning of the greek power supply sector integrating mainland and insular electric systems. *Computers & Operations Research*, v. 66, n. Supplement C, p. 292 – 312, 2016. ISSN 0305-0548.
- 22 YI, B.-W.; XU, J.-H.; FAN, Y. Inter-regional power grid planning up to 2030 in china considering renewable energy development and regional pollutant control: A multi-region bottom-up optimization model. *Applied Energy*, v. 184, n. Supplement C, p. 641–658, 2016. ISSN 0306-2619.
- 23 LEE, C.-Y.; HUH, S.-Y. Forecasting new and renewable energy supply through a bottom-up approach: The case of south korea. *Renewable and Sustainable Energy Reviews*, v. 69, n. Supplement C, p. 207 – 217, 2017. ISSN 1364-0321.
- 24 BERNTSEN, P. B.; TRUTNEVYTE, E. Ensuring diversity of national energy scenarios: Bottom-up energy system model with modeling to generate alternatives. *Energy*, v. 126, p. 886–898, 2017. ISSN 0360-5442.
- 25 BRILL, J. E. D.; CHANG, S.-Y.; HOPKINS, L. Modeling to generate alternatives: The hsj approach and an illustration using a problem in land use planning. *Management Science*, v. 28, p. 221–235, 1982.
- 26 DECAROLIS, J. F. Using modeling to generate alternatives (mga) to expand our thinking on energy futures. *Energy Economics*, v. 33, n. 2, p. 145 – 152, 2011. ISSN 0140-9883.
- 27 DECAROLIS, J. F. Using modeling to generate alternatives (mga) to expand our thinking on energy futures. *Energy Economics*, v. 33, n. 2, p. 145 – 152, 2011. ISSN 0140-9883.
- 28 TRUTNEVYTE, E. Expanse methodology for evaluating the economic potential of renewable energy from an energy mix perspective. *Applied Energy*, v. 111, n. Supplement C, p. 593 – 601, 2013. ISSN 0306-2619.
- 29 TRUTNEVYTE, E. Does cost optimization approximate the real-world energy transition. *Energy*, v. 106, n. Supplement C, p. 182 – 193, 2016. ISSN 0360-5442.



- 30 KOLTSAKLIS, N. E.; LIU, P.; GEORGIADIS, M. C. An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response. *Energy*, v. 82, n. Supplement C, p. 865 – 888, 2015. ISSN 0360-5442.
- 31 KABOLI, S. H. A.; SELVARAJ, J.; RAHIM, N. Long-term electric energy consumption forecasting via artificial cooperative search algorithm. *Energy*, v. 115, n. Part 1, p. 857 – 871, 2016. ISSN 0360-5442.
- 32 BEER, J.; WORRELL, E.; BLOKT, K. Long-term energy-efficiency improvements in the paper and board industry. *Energy*, v. 23, p. 21–42, 1997.
- 33 WORRELL, E.; PRICE, L. Policy scenarios for energy efficiency improvement in industry *Energy Policy*, v. 29, p. 1223–1241, 2001.
- 34 TANAKA, K. Review of policies and measures for energy efficiency in industry sector. *Energy Policy*, v. 39, p. 6532–6550, 2011.
- 35 FLEITER, T.; WORRELL, E.; EICHHAMMER, W. Barriers to energy efficiency in industrial bottom-up energy demand models - a review. *Renew Sustain Energy*, Rev 2011;15(6):3099e111. <https://doi.org/10.1016/j.rser.2011.03.025>.
- 36 GIRALDO, L.; HYMAN, B. Energy end-use models for pulp, paper, and paperboard mills. *Energy*, v. 20, p. 1005–1019, 1995.
- 37 GIRALDO, L.; HYMAN, B. An energy process-step model for manufacturing paper and paperboard. *Energy*, v. 21, p. 667–681, 1996.
- 38 FARLA, J.; BLOK, K.; SCHIPPER, L. Energy efficiency developments in the pulp and paper industry: A cross-country comparison using physical production data. *Energy Policy*, v. 97, p. 745–758, 1997.
- 39 KRAMER, K. J. et al. *Energy efficiency improvement and cost saving opportunities for the pulp and paper industry*. <https://www.energystar.gov/ia/business/industry/downloads/Pulpand-Paper-Energy-Guide.pdf>, 2009.
- 40 FRACARO, G. et al. Energy efficiency improvement and cost saving opportunities for the pulp and paper industry. *Energies*, 2012;5:3550e72. <https://doi.org/10.3390/en5093550>.
- 41 CAMIOTO, F. D. C.; MARIANO, E. B.; REBELATTO, D. A. D. N. efficiency in Brazil's industrial sectors in terms of energy and sustainable development. *Environ Sci Pol*, 2014;37:50e60. <https://doi.org/10.1016/j.envsci.2013.08.007>.
- 42 DA SILVA, F. L. C. et al. Forecast of long-term electricity consumption of the industrial sub-sector of pulp and paper in Brazil using a bottom-up approach. *Procedia Computer Science*, 2015;55:514e22. <https://doi.org/10.1016/j.procs.2015.07.028>.
- 43 DA SILVA, F. L. C. et al. A bottom-up bayesian extension for long term electricity consumption forecasting. *Energy*, 2019;198:210;167. <http://dx.doi.org/10.1016/j.energy.2018.10.201>.
- 44 DA SILVA, F. L. C. et al. Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector. *Energies*, 2022; 15, 588. <https://doi.org/10.3390/en15020588>.
- 45 PAO, H. T.; TSAI, C. M. Modeling and forecasting the CO2 emissions, energy consumption, and economic growth in Brazil. *Energy*, 2011;36:2450e8. <https://doi.org/10.1016/j.energy.2011.01.032>.
- 46 MAÇAIRA, P. M.; ELSLAND, R.; SOUZA, R. C.; OLIVEIRA, F. L. C.; FERNANDES, G Forecasting residential electricity consumption: a bottom-up approach for Brazil by region *Energy Efficiency* , 2020; 13,911–934. <https://doi.org/10.1007/s12053-020-09860-w>

- 47 MAÇAIRA, P. M.; SOUZA, R. C.; OLIVEIRA, F. L. C. Modelling and Forecasting the Residential Electricity Consumption in Brazil with Pegels Exponential Smoothing Techniques. *Procedia Computer Science*. 328-335 p, 2015
- 48 CABRAL, J. A.; CABRAL, M. V. F.; JÚNIOR, A. O. P. Elasticity estimation and forecasting: An analysis of residential electricity demand in Brazil. *Utilities Policy*. 101108 p, 2020
- 49 IRFFI, G. et al. Previsão da demanda por energia elétrica para classes de consumo na região Nordeste, usando OLS dinâmico e mudança de regime. *Econ. Apl.* 13, 2009. <https://doi.org/10.1590/S1413-80502009000100004>
- 50 RODRIGUES, A. *GreenSDN: energy efficiency in software defined networks*. São Paulo, 2016 Dissertação (Engenharia de Computação) - Universidade de São Paulo.
- 51 EVERITT, B. S. et al. *Cluster Analysis*, 5a. edição. Wiley, United Kingdom, 2011.
- 52 HARTIGAN, J. A.; WONG, M. A. Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, 1979, 28, 100-108. doi: 10.2307/2346830.
- 53 MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, eds L. M. Le Cam & J. Neyman, 1967, 1, pp. 281-297. Berkeley, CA: University of California Press.
- 54 LLOYD, S. P. Least squares quantization in PCM. Technical Note, Bell Laboratories. Published in 1982 in IEEE Transactions on Information Theory, 28, 128-137.
- 55 FORGY, E. W. Cluster analysis of multivariate data: efficiency vs interpretability of classifications. *Biometrics*, 1965, 21, 768-769.
- 56 LINDEN, R. Técnicas de Agrupamento. *Revista de Sistemas de Informação da FSMA*, n. 4 (2009) pp. 18-36, <http://www.fsma.edu.br/si/sistemas.html>
- 57 EHLERS, R.S. *Análise de séries temporais*, 5a. edição. 2009.
- 58 MORETTIN, P. A.; TOLOI, C. M. C. *Análise de séries temporais*, 2a. edição. E. Blücher, São Paulo, 2006.
- 59 BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Time Series Analysis: Forecasting and Control*, 4a. edição. Wiley, New Jersey, 2008.
- 60 SARNAGLIA, A. J. Q.; REISEN, V. A. Estimção Robusta em Processos Periódicos Auto-regressivos na Presença de Outliers Aditivos. In: 19º SINAPE, n. 19º. 2010, São Pedro, 2010.
- 61 GRANGER, C. W. J.; NEWBOLD, P. *Forecasting Economic Time Series*, Academic Press, New York, 1977.
- 62 HYNDMAN, R. ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 2. ed. Monash University, Australia
- 63 MATOS, D. B. *Técnicas de Estimção de Parâmetros Utilizadas para a Modelagem Matemática de Propulsores Eletromecânicos*. Ijuí, 2018 Dissertação (Modelagem Matemática) - Universidade Regional do Noroeste do Estado do Rio Grande do Sul.
- 64 PANKRATZ, A. *Forecasting with Dynamic Regression Models*. 1 ed. John Wiley & Sons, 1991.
- 65 ZANINI, A. *Redes Neurais e Regressão Dinâmica: um modelo híbrido para a previsão de curto prazo da demanda de gasolina automotiva no Brasil*. Rio de Janeiro, 2000 Dissertação (Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro.

- 66 WEST, M.; Harrison, J. *Bayesian Forecasting and Dynamic Models*. 2 ed. Springer, 1997.
- 67 MCCULLOCH, W. S; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. 115–133 p, 1943.
- 68 DA SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. *Redes Neurais Artificiais: para engenharia e ciências aplicadas*. 2 ed. São Paulo: Artliber, 2016.
- 69 EMPRESA DE PESQUISA ENERGÉTICA. *Eficiência Energética*. Disponível em: <https://www.epe.gov.br/pt/abcdenergia/eficiencia-energetica>. Acesso em: 3 fev. 2022
- 70 DA SILVA, F. L. C. *Extensão Estocástica Bayesiana da Abordagem Bottom-up Determinística para a Previsão de Longo Prazo do Consumo de Energia*. Rio de Janeiro, 2017 Tese (Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro.
- 71 MATHEWS, J. A.; BARONI, P. The industrial logistic surface: Displaying the impact of energy policy on uptake of new technologies, *Energy*, v. 57, p. 733–740, 2013. <http://dx.doi.org/10.1016/j.energy.2013.05.048>.
- 72 JAFFE, A. B.; STAVINS, R. N. The energy paradox and the diffusion of conservation technology, *Resource and Energy Economics*, v. 16, p. 91–122, 1994. [https://doi.org/10.1016/0928-7655\(94\)90001-9](https://doi.org/10.1016/0928-7655(94)90001-9)
- 73 INMETRO. *Informação ao Consumidor*. Disponível em: <http://www.inmetro.gov.br/consumidor/tabelas.asp>. Acesso em: 3 fev. 2022.
- 74 PROCEL. *PROCEL INFO*. Disponível em: <http://www.procel.gov.br/main.asp?View=B70B5A3C-19EF-499D-B7BC-D6FF3BABE5FA>. Acesso em: 3 fev. 2022.
- 75 PROCEL. *Edificações*. Disponível em: <http://www.procelinfo.com.br/main.asp?TeamID=%7B82BBD82C-FB89-48CA-98A9-620D5F9DBD04%7D>. Acesso em: 3 fev. 2022.

Neste trabalho as análises estatísticas foram realizadas através do software R 4.0.5 (The R Foundation for Statistical Computing, Vienna, Austria; <http://www.r-project.org>). As bibliotecas e funções utilizadas são:

1. Análise de Cluster

- Biblioteca: factoextra.
- Função: kmeans.

2. Modelos de Holt-Winters

- Biblioteca: forecast.
- Função: hw.

3. Modelos de Box e Jenkins

- Biblioteca: forecast.
- Função: auto.arima.

4. Modelos de Regressão Dinâmica

- Biblioteca: forecast.
- Função: forecast.

5. Modelos Lineares Dinâmicos

- Biblioteca: dlm.
- Função: dlm.

6. Redes Neurais Artificiais

- Biblioteca: forecast.
- Função: nnetar.