

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO RIO GRANDE
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL

ANÁLISE DA INFLUÊNCIA DAS VARIÁVEIS METEOROLÓGICAS NA DEMANDA DE
ELETRICIDADE

por

Liara Kurtz da Veiga

Dissertação para obtenção do Título de
Mestre em Modelagem Computacional

Orientadora: Profa. Dra. Viviane Leite Dias de Mattos
Coorientadora: Profa. Dra. Andréa Cristina Konrath

Rio Grande, setembro, 2023

Liara Kurtz da Veiga

“ANÁLISE DA INFLUÊNCIA DAS VARIÁVEIS METEOROLÓGICAS NA DEMANDA DE ELETRICIDADE”

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal do Rio Grande - FURG, como requisito parcial para obtenção do Grau de Mestre. Área de concentração: Modelagem Computacional.

Aprovada em:

BANCA EXAMINADORA

Documento assinado digitalmente
 **VIVIANE LEITE DIAS DE MATTOS**
Data: 02/10/2023 16:51:34-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Viviane Leite Dias de Mattos (Orientadora – FURG)

 Documento assinado digitalmente
ANDREA CRISTINA KONRATH
Data: 02/10/2023 15:50:36-0300
CPF: ***.170.610-**
Verifique as assinaturas em <https://v.ufsc.br>

Profa. Dra. Andréa Cristina Konrath (Coorientadora – UFSC)

Documento assinado digitalmente
 **TIAGO DA CRUZ ASMUS**
Data: 02/10/2023 15:15:29-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Tiago da Cruz Asmus (FURG)

Documento assinado digitalmente
 **DIANA FRANCISCA ADAMATTI**
Data: 02/10/2023 09:21:27-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Diana Francisca Adamatti (FURG)

Documento assinado digitalmente
 **Luiz Ricardo Nakamura**
Data: 29/09/2023 17:26:06-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Luiz Ricardo Nakamura (UFLA)

Rio Grande - RS
2023

Ficha Catalográfica

V426a Veiga, Liara Kurtz da.

Análise da influência das variáveis meteorológicas na demanda de eletricidade / Liara Kurtz da Veiga. – 2023.

103 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Modelagem Computacional, Rio Grande/RS, 2023.

Orientadora: Dra. Viviane Leite Dias de Mattos.

Coorientador: Dra. Andréa Cristina Konrath.

1. Patamares de Energia 2. *Box-Jenkins* 3. Variáveis Meteorológicas 4. SARIMAX I. Mattos, Viviane Leite Dias de II. Konrath, Andréa Cristina III. Título.

CDU 004

Catálogo na Fonte: Bibliotecário José Paulo dos Santos CRB 10/2344

AGRADECIMENTOS

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento da bolsa e a Universidade Federal do Rio Grande (FURG) e ao Programa de Pós-Graduação em Modelagem Computacional, os quais tornaram possível receber os subsídios a fim de proporcionar a minha continuidade nos estudos e aprimoramento da carreira profissional.

Agradeço à orientadora Viviane Leite Dias de Mattos, que guiou as minhas tarefas e tornou possível a resolução deste trabalho.

Agradeço à coorientadora Andréa Cristina Konrath, que me ajudou nas tomadas de decisão e contribuiu de diversas formas na elaboração desta pesquisa.

Ao colega Eng. Eletricista Felipe Germanos, que me ajudou na aquisição do banco de dados.

Aos meus familiares por todo apoio recebido, em especial à minha mãe, Vanda Laci Kurtz, e ao meu namorado, Evandro Cesca Mezzari. Também agradeço à minha gata Lupita pela companhia desde a graduação.

RESUMO

A energia elétrica foi fundamental para o surgimento de novas tecnologias, as quais tiveram um forte desenvolvimento nos últimos anos, aumentando extremamente a demanda por eletricidade. As modificações no clima apresentam projeções de aumento de eventos extremos, e mudanças relevantes nas variáveis meteorológicas podem ocorrer. Portanto, o objetivo do trabalho foi analisar a influência das variáveis meteorológicas na máxima demanda instantânea de energia elétrica de uma localidade no Rio Grande do Sul, Brasil. A partir dos dados de demanda de carga elétrica, foram constituídas sete séries temporais de demanda mensal, de janeiro de 2009 até agosto de 2021, as quais foram analisadas quanto a: medidas resumo, tendência e sazonalidade; esta última foi detectada com base na análise da variância dos dados. Assim, foram aplicados modelos SARIMA na imputação e modelagem dos dados de demanda, bem como, na imputação das séries meteorológicas. A demanda foi relacionada às variáveis meteorológicas por modelos SARIMAX com o método *stepwise backward* para seleção de variáveis. A capacidade preditiva dos modelos foi analisada pelos indicadores: RMSE, MAPE e MAE, cujos melhores índices tiveram a análise complementada em relação aos intervalos de confiança. Os resultados mostraram séries estacionárias de demanda mais correlacionadas com variáveis do tipo temperatura. Os modelos para os dias úteis revelaram que, durante o período da madrugada e as primeiras horas da manhã, exceto às segundas-feiras, houve influência da temperatura mínima na demanda de eletricidade. Ao longo do horário comercial, incluindo segundas-feiras, a demanda foi levemente influenciada pela temperatura mínima e também pela temperatura mínima associada à umidade. Aos finais de semana, quase todas as variáveis e conjunto de variáveis meteorológicas utilizados apresentaram influência na demanda de energia elétrica. Os índices de parcimonia e capacidade preditiva dos SARIMA mostraram que os modelos com ajuste razoável aos dados possuem maior potencial de atingir melhor desempenho na predição, em relação a modelos de ajuste ótimo, em vista das possíveis ocorrências de alterações no padrão da variabilidade das séries. Há exceções, e, neste trabalho, foi o patamar da carga média dos sábados, série com ajuste ótimo e melhor capacidade preditiva, além de ser a única a melhorar sua parcimonia com a entrada das variáveis exógenas. Este estudo envolveu os conceitos de estatística, engenharia e meio ambiente, a fim de melhorar os subsídios para o planejamento energético, no intuito de proporcionar o aprimoramento das condições de fornecimento de energia elétrica à população consumidora.

Palavras-chaves: Patamares de Energia, Box-Jenkins, Variáveis Meteorológicas, SARIMAX

ABSTRACT

Electrical energy was fundamental for the emergence of new technologies, which have had a strong development in recent years, greatly increasing the demand for electricity. Climate changes present projections of increased extreme events, and relevant changes in meteorological variables may occur. Therefore, the objective of this study was to analyze the influence of meteorological variables on the maximum instantaneous demand for electricity in a locality in Rio Grande do Sul, Brazil. From the electric load demand data, seven time series of monthly demand were created, from January 2009 to August 2021, which were analyzed in terms of: summary measures, trend and seasonality; the latter was detected based on the analysis of variance of the data. Thus, SARIMA models were applied in the imputation and modeling of demand data, as well as in the imputation of meteorological series. Demand was related to meteorological variables through SARIMAX models, with the stepwise backward method for variable selection. The predictive capacity of the models was analyzed by the indicators: RMSE, MAPE and MAE, whose best indices had the analysis complemented in relation to confidence intervals. The results showed stationary demand series more correlated with temperature-type variables. The models for weekdays revealed that, during the pre-dawn and the early hours of the morning, except on Mondays, there was an influence of the minimum temperature on the electricity demand. Throughout business hours, including Mondays, demand was slightly influenced by the minimum temperature and also the minimum temperature associated with humidity. During weekends, almost all the variables and ensembles of meteorological variables used had an influence on the demand electricity. The parsimony indices and predictive capacity of SARIMA showed that models with reasonable fit to the data have greater potential to achieve better performance in prediction, in relation to optimal fit models, in view of the possible occurrences of changes in pattern of variability of the series. There are exceptions, and, in this work, it was the average load level on Saturdays, a series with an optimal fit and better predictive capacity, in addition to being the only one to improve its parsimony with the input of exogenous variables. This study involved the concepts of statistics, engineering and the environment, in order to improve subsidies for energy planning, with the intention of provide the improvement of the conditions of supply of electric energy to the consumer population.

Keywords: Energy levels, Box-Jenkins, Meteorological Variables, SARIMAX

ÍNDICE

1. INTRODUÇÃO	15
1.1. Apresentação do tema	15
1.2. Justificativa.....	18
1.3. Objetivos	19
1.3.1. Objetivo Geral	19
1.3.2. Objetivos Específicos.....	19
2. FUNDAMENTAÇÃO TEÓRICA.....	20
2.1. Horizontes dos modelos de previsão na operação energética	20
2.2. Patamares da demanda de energia elétrica no Brasil.....	21
2.3. Dados de carga elétrica na prática: abordagem corretiva	22
2.4. Dados meteorológicos brutos: problemas e alternativas de solução	24
2.5. A relação das variáveis meteorológicas com a demanda de energia elétrica	27
2.6. Tendência, Sazonalidade e Estacionariedade das Séries Temporais	29
2.6.1. Tendência.....	29
2.6.2. Sazonalidade.....	30
2.6.3. Estacionariedade.....	31
2.7. Modelagem Box-Jenkins e Box-Tiao de séries temporais	32
2.7.1. Modelos Box-Jenkins.....	32
2.7.2. Modelos Box-Tiao	37
2.7.3. Validação e Intervalo de Confiança dos Modelos.....	39
3. METODOLOGIA.....	42
3.1. Origem e Preparo dos dados de estudo.....	42
3.2. Modelagem.....	47
4. RESULTADOS E DISCUSSÃO	50
4.1. Análise exploratória dos dados de demanda de energia elétrica.....	50
4.2. Relação entre os dados meteorológicos e de carga elétrica	53
4.3. Imputação dos dados de carga elétrica	57
4.4. Estacionariedade das séries e ordem das variáveis exógenas	60
4.5. Imputação dos dados meteorológicos.....	63
4.6. Modelos	64
4.6.1. Patamar de Carga Leve dos Dias Úteis (UL)	64
4.6.2. Patamar de Carga Média dos Dias Úteis (UM)	67
4.6.3. Patamar de Carga Pesada dos Dias Úteis (UP)	70
4.6.4. Patamar de Carga Média dos Sábados (SM)	72

4.6.5.	Patamar de Carga Pesada dos Sábados (SP).....	75
4.6.6.	Patamar de Carga Média dos Domingos (DM)	79
4.6.7.	Patamar de Carga Pesada dos Domingos (DP)	82
4.6.8.	Análise dos Resultados	85
5.	CONSIDERAÇÕES FINAIS.....	87
6.	REFERÊNCIAS.....	90
APÊNDICE 1 – Medidas resumo das séries		98
APÊNDICE 2 – Coeficientes de correlação.....		100
APÊNDICE 3 – Resultados dos testes de Estacionariedade		101

LISTA DE FIGURAS

Figura 2.1 - Curva de carga diária e tipos de agregação da carga	21
Figura 2.2 – (a) Ausência e (b) Presença de normalidade nos dados	30
Figura 2.3 – (a) Ausência e (b) presença de estacionariedade nas séries	32
Figura 2.4 - Etapas de configuração do modelo ARIMA para previsão de séries temporais	34
Figura 3.1 - Esquema das etapas utilizadas para aquisição e tratamento dos dados de estudo	42
Figura 3.2 - Localização da Estação Meteorológica do INMET e da Subestação da CEEE	43
Figura 3.3 - Esquema das etapas utilizadas nas modelagem SARIMAX.....	48
Figura 4.1 – Tendência e subséries mensais dos patamares dos dias úteis: jan./2009 – dez./2014 ...	51
Figura 4.2 – Tendência e subséries mensais dos patamares dos sábados: jan./2009 – dez./2014.....	52
Figura 4.3 – Tendência e subséries mensais dos patamares dos domingos: jan./2009 – dez./2014 ..	52
Figura 4.4 – Distribuições dos dados referentes aos dias úteis: jan./2009 – ago./2021.....	54
Figura 4.5 – Distribuições dos dados referentes aos sábados: jan./2009 – ago./2021	55
Figura 4.6 – Distribuições dos dados referentes aos domingos: jan./2009 – ago./2021	55
Figura 4.7 - Séries dos patamares de carga elétrica: jan./2009 - ago./2021	60
Figura 4.8 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar UL ...	65
Figura 4.9 – Validação do patamar de carga leve dos dias úteis: dados reais e simulados	66
Figura 4.10 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar UM	67
Figura 4.11 - Validação do patamar de carga média dos dias úteis: dados reais e simulados	69
Figura 4.12 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar UP .	71
Figura 4.13 - Validação do patamar de carga pesada dos dias úteis: dados reais e simulados	72
Figura 4.14 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar SM.	73
Figura 4.15 - Validação do patamar de carga média dos sábados: dados reais e simulados.....	75
Figura 4.16 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar SP ..	76
Figura 4.17 - Validação do patamar de carga pesada dos sábados: dados reais e simulados.....	78
Figura 4.18 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar DM	79
Figura 4.19 - Validação do patamar de carga média dos domingos: dados reais e simulados	81
Figura 4.20 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar DP .	82
Figura 4.21 - Validação do patamar de carga pesada dos domingos: dados reais e simulados	84

LISTA DE TABELAS

Tabela 2.1. Coeficientes de correlação amostrais de Pearson, Spearman e Kendall	28
Tabela 2.2. Características identificadas nas séries temporais estacionárias	31
Tabela 2.3. Características da Função Autocorrelação e da Função Autocorrelação Parcial	35
Tabela 2.4. Critérios de informação	36
Tabela 2.5. Indicadores de qualidade de modelos quanto à capacidade preditiva	39
Tabela 3.1. Intervalos dos patamares de carga nos meses janeiro, fevereiro e março	43
Tabela 3.2. Descrição das variáveis meteorológicas consideradas neste trabalho	44
Tabela 3.3. Configurações do modelo SARIMAX para previsão de séries temporais	49
Tabela 4.1. Medidas resumo dos patamares de carga elétrica: jan./2009 - ago./2021	50
Tabela 4.2. Resultados do Teste de Cox-Stuart: jan./2009 - dez./2014 ($\alpha = 0,05$)	51
Tabela 4.3. Resultados da ANOVA em blocos: jan./2009 - dez./2014 ($\alpha = 0,05$)	53
Tabela 4.4. Resultados do Teste de Friedman: jan./2009 - dez./2014 ($\alpha = 0,05$)	53
Tabela 4.5. Resultados da análise de correlação entre as séries meteorológicas e de carga elétrica	56
Tabela 4.6. Dados faltantes e modelos SARIMA aplicados nas séries de carga elétrica	57
Tabela 4.7. Resultados do teste HEGY para estacionariedade (D) das séries analisadas	60
Tabela 4.8. Ordem e critérios de entrada ao Xreg da função Arima	61
Tabela 4.9. Modelos SARIMA aplicados nas séries meteorológicas	64
Tabela 4.10. Patamar UL: Modelos SARIMA Candidatos e Seleccionados ($\alpha = 0,05$)	65
Tabela 4.11. Patamar UL: sequência de Teste do Modelo com os regressores meteorológicos	66
Tabela 4.12. Patamar UL: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)	67
Tabela 4.13. Patamar UM: Modelos SARIMA Candidatos e Seleccionados ($\alpha = 0,05$)	68
Tabela 4.14. Patamar UM: sequência de Teste 1 do Modelo com os regressores meteorológicos ...	68
Tabela 4.15. Patamar UM: sequência de Teste 2 do Modelo com os regressores meteorológicos ...	69
Tabela 4.16. Patamar UM: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)	70
Tabela 4.17. Patamar UP: Modelos SARIMA Candidatos e Seleccionados ($\alpha = 0,05$)	71
Tabela 4.18. Patamar UP: sequência de Teste do Modelo com os regressores meteorológicos	72
Tabela 4.19. Patamar SM: Modelos SARIMA Candidatos e Seleccionados ($\alpha = 0,05$)	73
Tabela 4.20. Patamar SM: sequência de Teste do Modelo com os regressores meteorológicos	74
Tabela 4.21. Patamar SM: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)	75
Tabela 4.22. Patamar SP: Modelos SARIMA Candidatos e Seleccionados ($\alpha = 0,05$)	76
Tabela 4.23. Patamar SP: sequência de Teste do Modelo com os regressores meteorológicos	77

Tabela 4.24. Patamar SP: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$) ...	79
Tabela 4.25. Patamar DM: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)	80
Tabela 4.26. Patamar DM: sequência de Teste do Modelo com os regressores meteorológicos	80
Tabela 4.27. Patamar DM: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$) .	81
Tabela 4.28. Patamar DP: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)	83
Tabela 4.29. Patamar DP: sequência de Teste do Modelo com os regressores meteorológicos	83
Tabela 4.30. Patamar DP: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)...	84

LISTA DE SÍMBOLOS

R^2	Coeficiente de determinação
r_p	Coeficiente de correlação de Pearson
r_s	Coeficiente de correlação de Spearman
n	Número total de observações
E_{95}	Quartil de 95%
S	Desvio padrão amostral
d_i	Diferença entre ordenações de dados
n_c	Número de pares concordantes
n_d	Número de pares discordantes
$E[.]$	Esperança matemática
var	Variância
p	Número de defasagens de um processo autorregressivo
P	Número de defasagens de um processo autorregressivo sazonal
q	Número de defasagens de um processo de médias móveis
Q	Número de defasagens de um processo de médias móveis sazonal
d	Grau de diferenciação
D	Grau de diferenciação sazonal
c	Constante
Y_t	Série temporal estocástica
y'_t	Série temporal estocástica diferenciada
$K-L$	Kullback-Leibler
$\log(L)$	Log-verossimilhança
Z_t	Série temporal estocástica diferenciada uma vez
w	Número de parâmetros de um modelo
Y_{t-i}	Série temporal estocástica defasada no tempo
β_i	Coeficiente de da série temporal estocástica defasada no tempo
w_j	Variáveis exógenas
r	Número máximo de variáveis exógenas
ε_{t-j}	Termo de erro defasado no tempo
$N(0, \sigma^2)$	Distribuição normal

$\hat{Y}_t$	Série temporal estocástica estimada
H	Número total de observações na validação da série temporal
U	Estatística U de Theil
E_F	Quartil associado a uma probabilidade avaliada conforme um modelo específico
F_x	Modelo com número de parâmetros determinado
$P[.]$	Probabilidade
$1 - \alpha$	Nível de confiança
Z	z-score
Y	valor observado da série temporal

Símbolos gregos

τ	Estatística Tau de Kendall
$\bar{Y}$	Média amostral
γ	Covariância
σ^2	Variância
σ	Desvio Padrão
S_Y	Desvio Padrão amostral
ϕ_p	Coeficiente autorregressivo
θ_q	Coeficiente de médias móveis
ε_t	Ruído branco
ω_j	Coeficiente das variáveis exógenas
θ_j	Coeficiente do termo de erro defasado no tempo
ϑ_w	Todos os parâmetros de um modelo específico
λ_i	Limite inferior de confiança
λ_s	Limite superior de confiança

Sub índices

k	Defasagem temporal
t	Instante de tempo
i	Defasagem temporal da série estocástica
j	Defasagem temporal do termo de erro

LISTA DE ABREVIATURAS

INMET	Instituto Nacional de Meteorologia
SIN	Sistema Interligado Nacional
SEB	Setor Elétrico Brasileiro
EPE	Empresa de Pesquisa Energética
PDO	Programação Diária de Operação
ONS	Operador Nacional do Sistema Elétrico
ANEEL	Agência Nacional de Energia Elétrica
SEP	Sistema Elétrico de Potência
SSA	Singular Spectrum Analysis
MAPE	Erro Percentual Absoluto Médio
EIA	Administração de Informações de Energia dos Estados Unidos
MICE	Multivariate Imputation by Chained Equation
ML	Machine Learning
MAE	Erro Médio Absoluto
IQD	Inverso do Quadrado da Distância
NS	Coefficiente de Nash-Sutcliffe
IPD	Inverso da Potência da Distância
RMSE	Raiz do Erro Quadrático Médio
ANOVA	Análise de Variância
USP	Universidade de São Paulo
DF	Dickey-Fuller
ADF	Dickey-Fuller Aumentado
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
PP	Phillips-Perron
LMC	Leybourne-McCabe
OCSB	Osborn-Chui-Smith-Birchenhall
HEGY	Hylleberg-Engle-Granger-Yoo
CH	Canova-Hansen
AIC	Critério de Informação de Akaike
AICc	Critério de Informação de Akaike corrigido

BIC	Critério Bayesiano de Schwarz
HQ	Hannan-Quinn
ARCH	Autoregressive Conditional Heteroskedasticity
AR	Autorregressivo
MA	Médias Móveis
ARMA	Autorregressivo e de Médias Móveis
ARIMA	Autorregressivo Integrado e de Médias Móveis
SARIMA	Autorregressivo Integrado e de Médias Móveis Sazonal
ARX	Autorregressivo com fator Exógeno
ARMAX	Autorregressivo e de Médias Móveis com fator Exógeno
ARIMAX	Autorregressivo Integrado e de Médias Móveis com fator Exógeno
SARIMAX	Autorregressivo Integrado e de Médias Móveis Sazonal com fator Exógeno
FS	Forward Selection
BE	Backward Elimination
SR	Stepwise Regression
SRMP	Stepwise Regression on Model Prediction
IC	Intervalo de Confiança
MW	Megawatt
TW	Terawatt
kW	Quilowatt
CEEE	Companhia Estadual de Energia Elétrica
UL	Carga Leve dos Dias Úteis
UM	Carga Média dos Dias Úteis
UP	Carga Pesada dos Dias Úteis
SM	Carga Média dos Sábados
SP	Carga Pesada dos Sábados
DM	Carga Média dos Domingos
DP	Carga Pesada dos Domingos
BDMEP	Banco de Dados Meteorológicos para Ensino e Pesquisa
NA	Not Available
OMS	Organização Mundial da Saúde

1. INTRODUÇÃO

1.1. Apresentação do tema

A energia elétrica foi desenvolvida por meio de muitos estudos científicos ao longo dos séculos. Um acontecimento marcante, na história do setor, foi a invenção da primeira lâmpada elétrica incandescente, em 1879, por Thomas Edison e, um dos maiores cientistas em termos de eletricidade foi Nikola Tesla, que teve uma fundamental contribuição nos conhecimentos sobre corrente alternada e sistemas complexos e distribuição de energia. Essencial durante toda Segunda Revolução Industrial, a energia elétrica, aos poucos, foi se tornando uma alternativa da energia à vapor, juntamente com o petróleo (FARIAS; SELLITTO, 2011; MAIA, 2017).

No Brasil, o início dos serviços do Setor Elétrico Brasileiro (SEB), em substituição aos processos a vapor e tração animal, modernizou a infraestrutura das cidades na competitividade das atividades industriais dos países. No século XIX, com a chegada da Light no país, empresa que procurava aproveitar o potencial hidráulico brasileiro, a participação por fonte de energia na matriz energética nacional se reverteu a favor da hidroeletricidade. Assim, com características únicas em relação a outros países, o Brasil desenvolveu seu setor elétrico com uma forte base hídrica, buscando aproveitar as condições favoráveis dos seus rios (ALVES, 2008).

No mundo, a eletricidade foi fundamental para o surgimento de novas tecnologias, ao longo do século XX, também conhecidas como tecnologias digitais, que revolucionaram a indústria, a economia e a sociedade, nas quais novas formas de armazenamento e de difusão da informação foram reveladas (RIBEIRO, 2022). A consequência da disseminação das novas tecnologias caracterizou a origem da Terceira Revolução, mais voltada para o âmbito técnico-científico, onde avanços científicos levam ao desenvolvimento de novas técnicas e novos equipamentos que, por sua vez, retroalimentam o avanço da ciência (CHAVES, 2017).

Com o forte desenvolvimento da tecnologia, equipamentos elétricos estão, cada vez mais, sendo utilizados no dia-a-dia e, associado ao fato do mundo estar mais populoso, o resultado é um aumento extremo da demanda por eletricidade nos últimos anos. Por exemplo, na Turquia, a demanda bruta anual de eletricidade aumentou de 15,6 TWh no ano de 1975 para 246,4 TWh no ano de 2013 (GUNAY, 2016; VU *et al.*, 2014).

A popularização da informática e o desenvolvimento da internet e redes sociais, nos últimos 20 anos, aumentou a dependência de energia elétrica por parte da população, em razão da propagação do livre comércio e serviços online atingirem o cotidiano das pessoas e, inevitavelmente, afetarem

aqueles que possuem menos contato com recursos tecnológicos. Nos últimos tempos, a pandemia de COVID-19 agravou tal dependência, provocando o confinamento das pessoas em suas residências e fazendo muitas trabalharem em casa, com o auxílio da internet e suas plataformas digitais, atividade também denominada como *home office* (LIPPO; ROCHA, 2021).

Para garantir o acesso à energia elétrica é necessário investir em planejamento energético e, segundo Alves (2008), isso deve ser feito desde a geração até o consumo final e, a má condução ou até mesmo a falta, seja da operação ou expansão do sistema, pode resultar em elevados prejuízos a todos os consumidores, tanto sob o ponto de vista financeiro quanto em relação à ocorrência de racionamentos, bem como no tocante à perda da qualidade de serviços. Por estes motivos, conhecer com antecedência o comportamento da carga é fundamental no planejamento, análise e operação dos sistemas de potência, de forma a garantir um fornecimento de energia elétrica ininterrupto, confiável, seguro e econômico (MIRASGEDIS *et al.*, 2006).

Importante, antes de mencionar a demanda de energia elétrica no texto a seguir, é ressaltar os diferentes conceitos de demanda e consumo: define-se demanda como a média das potências instantâneas (potências consumidas no instante), solicitadas pela unidade consumidora, seja residência, indústria, ponto comercial e etc., integralizadas em intervalos de 15 (quinze) minutos e, por sua vez, o consumo é a quantidade de energia ativa efetivamente medida no período de faturamento (SILVA, 2007).

Entretanto, destaca-se que, neste trabalho, os conceitos de demanda e consumo de energia elétrica foram considerados equivalentes, em razão da grande maioria das situações ocorrer o atendimento total da demanda, que, conseqüentemente, se torna igual ao consumo de energia elétrica.

Alves (2008) afirma que existem vários fatores que influenciam o comportamento da demanda, que podem ser classificados em quatro categorias: econômicos, sazonais, climáticos e aleatórios:

- Fatores econômicos: influenciam diretamente na demanda. Exemplo: uma população que apresenta elevação no seu consumo per capita e que antes possuía acesso restrito à energia, buscará aumentar o seu conforto, seja com a maior utilização de equipamentos elétricos existentes em suas residências, ou na compra de novos equipamentos consumidores de energia elétrica.
- Fatores sazonais: diretamente ligados a efeitos de estações do ano, determinando a variação da demanda em função do período do ano. Exemplo: no verão, uma maior quantidade de equipamentos de condicionamento de ar é acionada, e no inverno, há uma busca pela elevação da temperatura ambiente que demanda uma maior quantidade

de energia pelo uso de equipamentos de aquecimento. No Brasil, a sazonalidade apresenta uma variação conforme diferentes fatores regionais. Mais especificamente, na região sul do país, existe uma característica mais voltada para as condições de temperatura ao longo das quatro estações do ano, por exemplo: os meses de junho a setembro correspondem ao período do inverno, quando é verificada a redução da temperatura nesta região. Também podem ser classificados como fatores sazonais: reduções de demanda de energia por paradas de manutenção realizada anualmente pelas indústrias, em feriados prologados, dentre outros.

- Fatores climáticos: as mudanças no clima podem desencadear significativos impactos nos mais variados setores da sociedade. Espera-se que as mudanças climáticas alterem a temperatura média e, conseqüentemente, aumentem a variabilidade dos eventos extremos associados à precipitação, podendo desencadear inundações e secas mais intensas e frequentes. Extremos climáticos associados não só à precipitação, como também à temperatura, podem afetar setores que envolvem diretamente o consumo de energia (SANTOS, 2009).
- Fatores aleatórios: como exemplos podem ser citados distúrbios no Sistema Interligado Nacional (SIN) ocasionados pelo tipo de operação que um conjunto de pequenos consumidores ou um único grande consumidor pode provocar na estabilidade do sistema, devido ao seu mercado de atuação ficar mais aquecido.

Em relação aos fatores climáticos, estudos relatam a associação entre as mudanças climáticas e as alterações no meio ambiente por parte dos seres humanos: Santos (2009) afirma que, na América Latina é observada uma grande variedade de modificações no clima, que pode ser resultante não só da variabilidade climática natural, como também da interferência humana no sistema; Hooper *et al.* (2005) asseguram que o desmatamento e o mau uso dos ecossistemas alteram o clima regional contribuindo para perturbações no sistema global. Um exemplo de alteração regional do clima são as áreas urbanas que, segundo Assis (2005), geralmente abrigam maior parte das atividades econômico-financeiras, sociais e culturais, concentrando grandes grupos populacionais, resultando em zonas de grande demanda e consumo de energia, água e matérias-primas.

O mesmo autor afirma que, apesar de ser polêmica a ideia da influência da urbanização na mudança climática global, alguns autores concordam sobre o impacto do ambiente antrópico no clima, pelo menos no nível microclimático. De acordo com Vu *et al.* (2014), um ponto importante de ser investigado para o planejamento das cidades corresponde ao parcelamento das áreas urbanas, que

alteram o ambiente de vida em vários aspectos, inclusive em relação ao conforto térmico, podendo acarretar uma maior ou menor necessidade de energia elétrica.

Conforme Apadula *et al.* (2012), a desregulamentação dos mercados de eletricidade e as evidências das mudanças climáticas em curso promoveram um aumento do interesse científico no estudo das relações entre o consumo de energia elétrica e o clima. Como a frequência de eventos extremos tende a aumentar no futuro e mudanças relevantes nas variáveis meteorológicas podem ocorrer de acordo com as projeções climáticas, existe uma grande importância em investigar especificamente a influência das condições meteorológicas na demanda de eletricidade. Pardo *et al.* (2002) afirma que, de um número significativo de estudos analisados, a maioria explora o papel da variabilidade do clima e do consumo de energia modelando a demanda em função de fatores climáticos sazonais.

Acerca dos modelos, existe também a possibilidade de modelar a demanda de eletricidade com base na identificação de padrões na variabilidade dos dados da própria demanda. Neste caso, a regressão linear corresponde à técnica de modelagem e um exemplo de aplicação são os modelos Box-Jenkins (1970). Desta forma, de acordo com Pereira (2021), a relação do comportamento da demanda com os fatores climáticos pode ser estabelecida considerando as variáveis meteorológicas como exógenas aos modelos Box-Jenkins, configurando, assim, a modelagem Box-Tiao (1975).

1.2. Justificativa

Verifica-se no planejamento um importante processo que visa a minimização dos custos inerentes a toda cadeia do mercado de energia, desde a sua geração até o consumo, além de proporcionar a máxima confiabilidade sistêmica. No entanto, a falta deste, ou até mesmo, um planejamento mal conduzido, seja ele na expansão do sistema ou da operação do sistema, poderia trazer elevados prejuízos a todos os consumidores, resultando em custos mais elevados, novos riscos de racionamentos e na perda da qualidade de serviços (ALVES, 2008).

É vital para um país a capacidade de fornecer eletricidade igual à demanda. Se a capacidade de geração de um país for inferior à demanda bruta, a indústria dependente de eletricidade é afetada negativamente e ocorrem apagões; por outro lado, uma capacidade de geração superior à demanda resulta em usinas com capacidade ociosa de trabalho, representando um desperdício de recursos econômicos (GUNAY, 2016).

A fim de manter o equilíbrio entre a demanda e oferta de eletricidade, é extremamente relevante a realização da previsão precisa da demanda futura. Além disso, com a tendência de aumento dos extremos climáticos, significativas mudanças nas variáveis meteorológicas tendem a

ocorrer de acordo com as projeções climáticas, razão pela qual existe uma grande importância em investigar especificamente a influência das condições meteorológicas na demanda de eletricidade (APADULA *et al.*, 20120; VU *et al.*, 2014).

1.3. Objetivos

1.3.1. Objetivo Geral

Analisar a influência das variáveis meteorológicas na máxima demanda instantânea de energia elétrica de uma região específica, a fim de aprimorar a qualidade do ajuste, bem como, a capacidade preditiva dos modelos no setor.

1.3.2. Objetivos Específicos

- Realizar um levantamento bibliográfico para identificar possíveis variáveis meteorológicas que influenciam na demanda de energia elétrica;
- Definir a base de dados e realizar o tratamento e a análise exploratória das informações meteorológicas e de demanda de eletricidade;
- Identificar as variáveis meteorológicas de maior influência na demanda de energia elétrica;
- Efetuar a modelagem da demanda de energia elétrica com a aplicação de modelos Box-Jenkins, a fim de explicar o comportamento da máxima demanda instantânea em função de seus valores e erros passados;
- Utilizar as variáveis meteorológicas como regressoras dos modelos Box-Jenkins;
- Realizar a validação dos modelos encontrados.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Horizontes dos modelos de previsão na operação energética

Os modelos de previsão de carga elétrica têm sido amplamente explorados e formulados nas últimas décadas em diferentes escalas temporais e espaciais. A complexidade da operação energética exige diferentes tipos de análises de desempenho em sua logística (geração, transmissão, distribuição e carga, além da questão dos usos múltiplos), que, por sua vez, levam em consideração os efeitos do planejamento de médio prazo, de curto prazo e da programação diária (APADULA *et al.*, 2012; PARDO *et al.*, 2002).

No caso do planejamento de longo prazo, existe uma vinculação com a expansão do sistema, envolvendo estudos de identificação e quantificação de potenciais energéticos mais econômicos, que são de competência da Empresa de Pesquisa Energética (EPE), conforme a legislação brasileira (BRASIL, 2004).

Por sua vez, o planejamento energético de médio prazo abrange um horizonte de dois a cinco anos, utiliza uma metodologia que trabalha com intervalos mensais dos dados e está principalmente relacionado com o planejamento do fornecimento de combustíveis, operação de manutenção de concessionárias e negociações de contratos com outras empresas do setor energético e demais setores (ALVES, 2008; APADULA *et al.*, 2012).

No planejamento energético de curto prazo, segundo Alves (2008), o horizonte de estudo compreende um ano à frente, com intervalo semanal para o primeiro mês e mensal nos meses seguintes.

Em relação à Programação Diária de Operação (PDO) das usinas, o mesmo autor afirma que, o período previsto corresponde a duas semanas, com intervalos horários dos dados, e são consideradas, detalhadamente, as restrições de cada usina geradora, juntamente com suas características de eletricidade. Na PDO são estabelecidos ajustes na curva da demanda de energia conforme as condições meteorológicas previstas, cujas informações meteorológicas provêm das estações de medição do INMET, do Ministério da Defesa, dos sistemas estaduais de meteorologia e da Agência Nacional das Águas (ONS, 2008a).

2.2. Patamares da demanda de energia elétrica no Brasil

Segundo Lima (2008), desde os grandes motores industriais até os equipamentos eletrodomésticos, todos são projetados e construídos para funcionar dentro de certos limites de tensão elétrica e, equipamentos operando fora de tais limites podem apresentar comportamento insatisfatório ou, até mesmo, ser danificados. Desta forma, torna-se imprescindível o estabelecimento de níveis de tensão de energia elétrica a serem observados pelo Operador Nacional do Sistema Elétrico (ONS), concessionárias e permissionárias de serviços públicos de distribuição de energia elétrica.

Os níveis de tensão, regulamentados pela Agência Nacional de Energia Elétrica (ANEEL, 2001), impõem a todos os setores relacionados à operação dos sistemas elétricos (planejamento/pré-operação, operação em tempo real e pós-operação) o atendimento de diversas regras, garantindo a realização adequada da operação (LIMA, 2008).

Desta forma, segundo Farias (2009), foram estabelecidos períodos, ou patamares de carga elétrica de três tipos (leve, média e pesada), nos quais as informações horárias foram substituídas por intervalos agregados de tempo dentro de um mesmo dia.

A carga elétrica apresenta uma variação contínua, oscilando entre o patamar de carga leve passando por um período de carga média até alcançar a carga pesada. Após o término do período de carga pesada, ocorre novamente um período de carga média, finalizando o ciclo. A curva de carga diária, disposta na Figura 2.1, representa o comportamento da sociedade, em termos de consumo de energia (SACRAMENTO *et al.*, 2012; FARIAS, 2009).

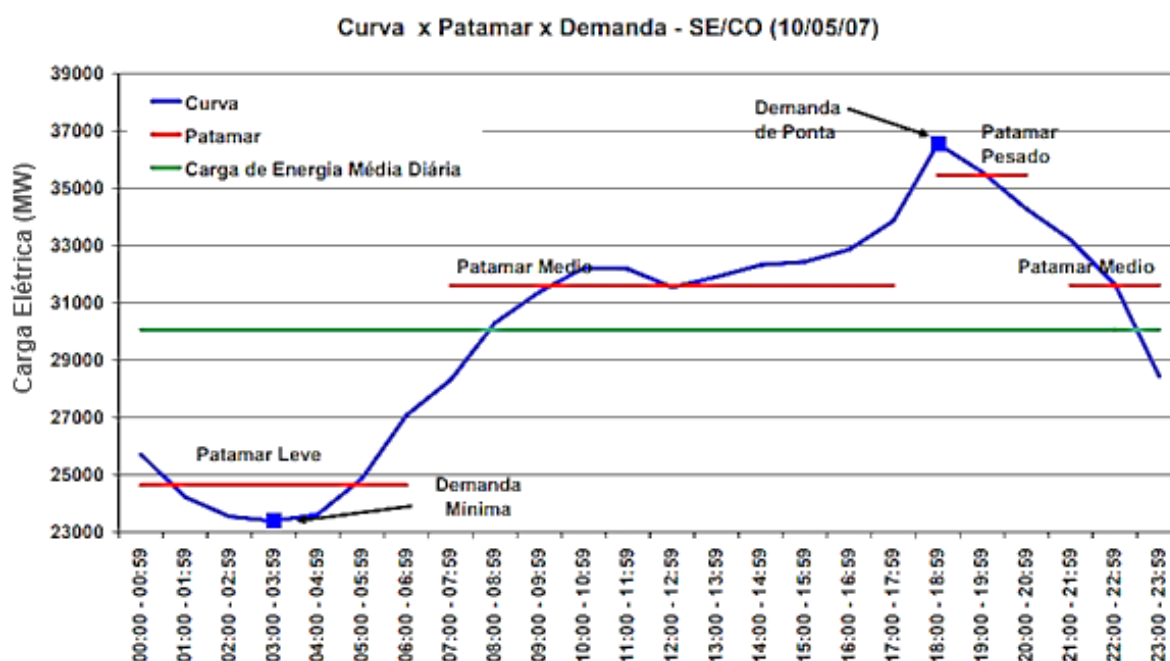


Figura 2.1 - Curva de carga diária e tipos de agregação da carga (FARIAS, 2009)

O exemplo apresentado na Figura 2.1 mostra que os horários de intervalo dos patamares foram definidos considerando os "níveis" da curva de carga durante o dia, sendo assim, o patamar leve representa os períodos de carga mais baixa, o patamar de carga média corresponde aos intervalos onde a carga atinge "níveis" médios, e o patamar de carga pesada, ocorre quando os níveis são os mais elevados (FARIAS, 2009).

Assim, conforme Lima (2008), os valores entre os limites de tensão elétrica são definidos para cada período de carga, podendo apresentar valores iguais em períodos de carga diferentes. A tarefa do controle de tensão possui como principal objetivo impedir que as tensões dos barramentos elétricos, sob sua responsabilidade, atinjam ou, até mesmo, ultrapassem valores não permitidos. Tais valores são definidos tanto pelas características dos equipamentos presentes no Sistema Elétrico de Potência (SEP) quanto pelos estudos realizados no planejamento/pré-operação.

O mesmo autor afirma que, nos casos onde os barramentos próximos aos centros de carga apresentam altos índices de tensão, ou seja, acima do limite superior permitido, há possibilidade de ocorrer danos aos equipamentos de consumidores e também nas subestações do sistema (para-raios, disjuntores e, transformadores, dentre outros). Por outro lado, as baixas tensões também podem causar danos em motores, interromper processos industriais e, até mesmo, conduzir a operação a uma situação de insegurança no caso de alteração do sistema elétrico, podendo resultar no desligamento de uma linha de transmissão ou a entrada repentina de um grande bloco de carga, comprometendo o fornecimento de energia ao consumidor final.

2.3. Dados de carga elétrica na prática: abordagem corretiva

A contínua oferta de eletricidade ao consumidor é, sobretudo, dependente do prévio conhecimento sobre o comportamento da carga elétrica por parte do setor energético (Mirasgedis *et al.*, 2006). As previsões deste comportamento podem ser feitas por meio de análise e modelagem de séries temporais, porém tais métodos exigem séries completas e consistentes de dados, algo que não ocorre com muita frequência no cotidiano.

Segundo Barbosa *et al.* (2018), falhas ou fraudes no processo de medição e transmissão de energia resultam em problemas de leitura incorreta ou perda de dados, gerando informações corrompidas ou faltantes, as quais, por sua vez, são novamente medidas e transmitidas à central, causando inconsistências ao sistema.

Conforme Correia (2018), o preenchimento de falhas na base de dados é de essencial importância para previsão de séries temporais, uma vez que, a ausência de dados em uma série de tempo pode comprometer a estimativa e a previsão, as quais são baseadas no histórico de informações.

Porém, realizar este preenchimento, ou seja, imputar dados, não costuma ser uma tarefa fácil, devido às estimativas possuírem a necessidade de manter as principais características dos dados originais. Uma outra questão importante está associada à escolha de um método inadequado de imputação, o qual pode resultar em equívocos nas conclusões das análises (BARBOSA *et al.*, 2018; CORREIA, 2018).

Como solução de tais adversidades, surgiram técnicas estatísticas de imputação de dados, as quais preenchem lacunas ou anomalias presentes nas séries e geram intervalos temporais com dados consistentes e aptos à aplicação em posteriores análises.

Barbosa *et al.* (2018) propôs um algoritmo adaptativo para imputar dados de redes inteligentes de energia, também conhecidas como *Smart Grids*, uma vez que, problemas de perda ou leitura incorreta dos dados resultam em dados corrompidos ou faltantes que são reproduzidos à central, gerando inconsistência. O algoritmo *AdaptS* foi constituído a partir da avaliação de desempenho de outros dois algoritmos: *Spline* e *Singular Spectrum Analysis* (SSA). Os resultados comprovaram que o *AdaptS* produziu valores mais precisos para os dados ausentes, maximizando assim a sua acurácia média em relação aos dados originais.

Correia (2018) também direcionou seu estudo para as *Smart Grids*, que contam com medidores inteligentes em conjunto com um sistema de coleta automática de dados, os quais representam a totalidade das novas instalações no Brasil. Porém, a regulação que determina regras para as leituras ainda não acompanhou esta modernização e continua com sistema de medição convencional, realizado por meio da atuação de leituristas. Então, a fim de solucionar o problema exposto, o autor propôs uma nova metodologia de elaboração do faturamento na ausência de leituras, de forma a considerar consumos anteriores, com aplicação de técnicas estatísticas para definir o valor a ser faturado. Na imputação de dados faltantes foram aplicados métodos de interpolação linear e método *Spline*. Já na previsão da quantidade de energia elétrica consumida em kWh para cada cliente de uma distribuidora, foram propostos modelos de Suavização Exponencial de Holt e Modelos Box-Jenkins. A avaliação dos modelos foi feita com base no Erro Médio Absoluto Percentual (MAPE). Os melhores resultados foram obtidos utilizando o preenchimento da dados faltantes por interpolação linear e a previsão de valores não medidos pelo método da Suavização Exponencial de Holt.

Ruggles *et al.* (2020) utilizaram um método para criar conjuntos completos de dados de séries temporais a partir de um conjunto de registros de demanda pertencentes à Administração de Informações de Energia dos Estados Unidos (EIA). Primeiramente, os autores desenvolveram oito algoritmos de triagem de valores anômalos para identificar as discrepâncias mais extremas, e em muitos destes foram aplicadas técnicas de mediana e intervalo interquartil. Após, para imputação de

dados foi utilizada uma técnica de Imputação Múltipla por Equações Encadeadas (MICE), que consiste, essencialmente, em uma série de regressões lineares executadas em sequência e repetidas até que as estimativas para valores ausentes tenham convergido o suficiente. A técnica foi avaliada com a marcação intencional de dados plausíveis como ausentes, imputando tais dados e comparando estas imputações com os valores verdadeiros por meio do MAPE. Em geral, o MAPE dos valores imputados foi de 3,5% e a metodologia proposta demonstrou uma concordância razoável e uma ligeira vantagem para lacunas maiores de dados em relação ao método utilizado pelo EIA, onde os valores da hora e do dia anterior são usados para substituir dados ausentes ou anômalos.

Com uma abordagem mais complexa, Khan *et al.* (2020) propuseram um modelo híbrido de previsão de dados de consumo de energia elétrica para Coreia do Sul, baseado em técnicas de aprendizado de máquina (*Machine Learning*), o qual também foi utilizado para imputar uma grande lacuna, que representava 21,26% do total dos dados. As técnicas ML utilizadas foram: *XGBoost* - *Extreme Gradient Boosting* - um método que aprende em sequência ajustando a árvore de regressão atual aos erros da última árvore; *CatBoost* que corresponde a uma biblioteca de aumento de gradiente que pode gerenciar dados categóricos, e; *Random Forest*, que consiste em uma abordagem de aprendizado conjunto para classificação, regressão e outras obrigações a serem executadas na construção da árvore de decisão. Na avaliação da qualidade do ajuste do modelo, foram utilizados: o Erro Absoluto Médio (MAE) e o coeficiente de determinação (R^2). O modelo proposto com dados imputados mostrou-se muito melhor em relação ao mesmo modelo com dados ausentes e, em relação ao desempenho, os resultados da previsão foram melhores do que técnicas com algoritmos individuais.

Todos os estudos mencionados acima possuem como foco resolver os problemas de ausências e falhas nas séries de demanda de energia elétrica, aplicando técnicas estatísticas de imputação de dados.

2.4. Dados meteorológicos brutos: problemas e alternativas de solução

Segundo Xavier (2016), as informações recolhidas em intervalos iguais de tempo durante um período mais longo são caracterizadas como séries temporais, ou seja, tanto os dados de demanda quanto os dados meteorológicos coletados para o presente estudo possuem esta definição.

Fante e Neto (2016) afirmaram que uma das condições fundamentais para estudos e pesquisas no âmbito das séries temporais meteorológicas corresponde à consistência de dados, que na prática é raramente possível, devido a diversos fatores, como: ausência de dados em determinados locais, falhas na série histórica, mudanças inadequadas dos locais em que os instrumentos de medição estavam instalados, ou, até mesmo, o fechamento das estações meteorológicas.

A consequência de tais fatores, conforme Silva e Jardim (2017), corresponde à ausência ou inconsistência de dados de estações ou postos pluviométricos no território nacional e a alternativa para corrigir as falhas existentes é dada por meio de procedimentos estatísticos. A seguir serão abordados alguns destes procedimentos aplicados para dados horários.

Baba *et al.* (2014) pesquisaram uma metodologia para automatizar e dar suporte na identificação e correção de dados meteorológicos inconsistentes, utilizando técnicas estatísticas e um banco de dados espacial. Para realização dos testes foi utilizado um conjunto de dados horários referentes ao mês de novembro de 2009 distribuídos em nove estações automáticas no Paraná. A variável utilizada nos testes foi a temperatura e os métodos testados foram: a correlação de Pearson (1932), para verificar quais estações apresentaram dados mais correlacionados entre si, combinada com a regressão linear, para estimar um valor na estação de interesse com base nas estações mais correlacionadas; e dois métodos que consideram o aspecto espacial: o inverso do quadrado da distância (IQD) e o método geoestatístico de krigagem, ambos utilizando um raio limite de distância para selecionar as estações mais próximas do ponto de interesse. Apesar da regressão linear utilizar, em média, um menor número amostral de estações de apoio, foram obtidos resultados semelhantes de erros e estimativas quando comparados aos obtidos por krigagem e IQD.

Em relação a dados de precipitação, Brubacher *et al.* (2020), com base na revisão de métodos de preenchimento de falhas e de interpolação espacial, afirmam que a regressão linear múltipla costuma ser muito utilizada e produz bons resultados sempre que a variável a ser estimada depende fortemente de dois ou mais atributos, ou variáveis preditivas. De outro ponto de vista, os interpoladores espaciais apresentam vantagem no contexto espacial, devido à utilização da localização para ponderar o peso de cada estação na estimativa. Uma das técnicas mais simples utilizadas, porém com razoáveis resultados para uma ampla variedade de dados, corresponde ao inverso da potência da distância, IPD, podendo ser quadrática ou não. Esta técnica costuma ser adotada em estudos de clima, para interpolar precipitação diária em bacias hidrográficas ou parâmetros de equações de chuvas intensas em locais desprovidos de dados pluviométricos, como é o caso do estudo de Flores *et al.* (2017). Estes autores utilizaram o método espacial do inverso do quadrado da distância, IQD, para o preenchimento de dados horários de precipitação na cidade de

Lajeado/RS, no período entre 2004 e 2016. Neste processo foram utilizadas oito estações meteorológicas e a avaliação de desempenho do método foi realizada com base no coeficiente de Nash-Sutcliffe (NS) e o quartil de 95% (E_{95}). A metodologia deste estudo permitiu a realização de uma análise integrada das características de chuvas extremas.

Ventura *et al.* (2016), por sua vez, compararam o desempenho dos métodos: média simples, média móvel, regressão linear simples e regressão linear múltipla. Tal comparação foi baseada nos resultados dos testes que envolveram as variáveis: temperatura, umidade relativa, ponto de orvalho, pressão atmosférica e radiação solar. Os dados tiveram intervalo definido a cada hora, durante um período de 90 dias, em três cidades brasileiras: Porto Alegre, Manaus e Rio de Janeiro. Os resultados mostraram que a regressão linear múltipla teve ótimo desempenho para a maioria das variáveis e a média simples modelou de forma precisa a pressão atmosférica. Porém, a radiação atmosférica não apresentou bons resultados para nenhum dos métodos testados.

A radiação atmosférica, segundo Oliveira (2017) e Guarnieri (2007) *apud* Machado e Rampinelli (2018), apresenta particularidades, como a descontinuidade e a natureza estocástica. Por consequência, seus dados possuem variabilidade elevada e as incertezas em sua estimativa são grandes, devido à dependência de parâmetros como cobertura de nuvens e à imprecisão dos métodos utilizados tanto para medição quanto para estimativa de dados. Os autores empregaram métodos de replicação e persistência no preenchimento de lacunas de dados pertencentes à cidade de Araranguá/SC, durante um ano, com periodicidade de cinco minutos. Com as análises, os autores constataram que o método de persistência, equivalente a uma média aritmética simples do instante anterior e posterior daquele que se deseja estimar, forneceu valores de radiação com certo grau de confiabilidade, sendo considerado um método eficaz.

Silva *et al.* (2019) avaliaram a técnica de imputação de dados, ou seja, preenchimento de falhas, via *software R* em dados horários de radiação solar global de Arapiraca, no estado de Alagoas, durante o período de 2008 a 2016. Os autores realizaram o processo de imputação múltipla pelo pacote *mice* (BUUREN; OUDSHOORN, 2011), através do método *midastouch*, equivalente à média preditiva ponderada, e utilizaram diversos indicadores para verificar ajustes de sub ou superestimativa da amostra, tais como: coeficiente de determinação (R^2), coeficiente de correlação de Pearson (r_p), Índice de Concordância, Raiz Quadrada do Erro Quadrático Médio (RMSE) e, Erro Absoluto Médio (MAE), dentre outros. Os resultados foram satisfatórios, provando ser um método muito útil e, além disso, pode ser devidamente aplicado em qualquer série temporal de dados climáticos.

2.5. A relação das variáveis meteorológicas com a demanda de energia elétrica

As variáveis meteorológicas, como temperatura, precipitação, umidade, insolação, velocidade do vento e umidade podem ter impactos significativos na demanda de eletricidade. Ao todo, a interdependência das variáveis meteorológicas tende a ser muito complexa, porém, de forma geral, a temperatura costuma ser a variável mais prevalente e com maior influência na demanda, uma vez que, as variações de temperatura impõem a utilização de aparelhos de aquecimento no inverno e ar condicionado no verão (KIM *et al.*, 2019; VU *et al.*, 2014).

Segundo Alves (2008), os equipamentos reguladores de temperatura são utilizados quando há necessidade de melhorar o conforto térmico do corpo humano em um ambiente, e para que haja este conforto, a sensação térmica precisa constar dentro de uma faixa de índices, os chamados índices de conforto térmico, cuja variação é dada, em sua maior parte, pelas variáveis temperatura ambiente e umidade relativa do ar. Conforme Apadula *et al.* (2012), a umidade relativa do ar apresenta um efeito nos índices de demanda de energia quando combinada com temperaturas mornas e quentes, porque a temperatura percebida tende a ser maior nestas condições de ambiente e, conseqüentemente, o uso de aparelhos de refrigeração aumenta.

Por sua vez, Mirasgedis *et al.* (2016) afirmam que a experiência de muitas concessionárias indica os seguintes elementos meteorológicos com maior efeito na demanda: temperatura, umidade, velocidade do vento e precipitação, em ordem decrescente de importância. Apesar desta afirmativa, os autores optaram por aplicar uma análise da influência de determinados parâmetros meteorológicos na demanda de energia elétrica.

Portanto, apesar de reconhecida uma certa hierarquia padrão para as variáveis meteorológicas mais influentes na demanda, os fatores climáticos locais podem apresentar características muito peculiares, provocando alterações no comportamento padrão, por exemplo: em regiões de inverno característico, a velocidade do vento é importante para mensuração de índices de conforto, devido ao aumento da sensação de frio de um corpo quando exposto ao vento, causando o desconforto térmico do indivíduo (ALVES, 2008). Sendo assim, existe muita importância em avaliar a relação das variáveis meteorológicas com a demanda de eletricidade de uma região específica de interesse.

Muitas vezes, tal relação costuma ser avaliada com base nas análises dos coeficientes de correlação de Pearson, Spearman (1904) ou Kendall (1938), apresentados na Tabela 2.1.

O coeficiente de correlação de Pearson é utilizado para expressar quanto uma associação de variáveis quantitativas pode ser descrita por uma função linear. De outro lado, tanto o coeficiente de correlação de Kendall, como o coeficiente de correlação de Spearman, são medidas de correlação não paramétrica para dados com mensuração ordinais. Porém, o coeficiente de correlação de Spearman

indica o quão bem uma relação pode ser descrita em uma função monotônica (função constante, crescente ou decrescente) que não necessariamente é linear e o coeficiente de correlação de Kendall representa a probabilidade de ocorrência de pares de dados que são concordantes nas variáveis analisadas (CAMPOS *et al.*, 2021).

Tabela 2.1. Coeficientes de correlação amostrais de Pearson, Spearman e Kendall

Coeficiente de Correlação	Definição
Pearson	$r_p = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{S_x} \right) \left(\frac{Y_i - \bar{Y}}{S_y} \right)$
Spearman	$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$
Kendall	$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)}$

Fonte: Elaborado pela autora, baseado em Falcão (2012).

Obs.: a variável n corresponde ao número total de observações das séries consideradas; X_i , aos dados da primeira série; $\bar{X}$, à média dos valores de X ; S_x , ao desvio padrão amostral dos valores de X ; Y_i , aos dados da segunda série analisada; $\bar{Y}$, à média dos valores de Y ; S_y , ao desvio padrão amostral dos valores de Y ; d_i equivale à diferença entre as ordenações de X e Y ; n_c , ao número de pares concordantes; n_d , ao número de pares discordantes. Maiores detalhes podem ser consultados no estudo de Falcão (2012).

Nos casos de séries temporais com tendência, é necessário cautela antes de aplicar os coeficientes de correlação, devido à possibilidade de gerar resultados com correlações espúrias, que ocorrem quando uma alta correlação existe nos dados em função da tendência de ambas as variáveis e não, devido à semelhança na variabilidade de seus dados. Por este motivo, algumas vezes, aplica-se a correlação cruzada, que utiliza séries estacionárias durante o processo, descartando efeitos de tendência nos dados. Outra possibilidade consiste em utilizar coeficientes de correlação simples aplicados a séries estacionárias. Em geral, o método mais amplamente utilizado é o coeficiente de correlação de Pearson tanto para descrever o padrão de relacionamento entre variáveis como para fazer inferências válidas para a população a partir de dados amostrais (CHEN; POPOVIC, 2002).

2.6. Tendência, Sazonalidade e Estacionariedade das Séries Temporais

Morettin e Toloi (2018) afirmam que, antes da tentativa de modelar e prever uma determinada série de tempo, é desejável ter uma ideia preliminar de seus dados, ou seja, fazer uma análise exploratória, para poder descrever algumas das suas principais propriedades. Conforme Lima (2008), geralmente, as séries temporais são compostas por quatro componentes distintas:

- Tendência: corresponde a um comportamento específico dos dados ao decorrer do tempo, exibindo constantes crescimentos ou declínios em sucessivos intervalos de tempo;
- Sazonalidade: são variações específicas dos dados em curtos períodos de tempo, geralmente dentro de um ano;
- Ciclicidade: equivale a um certo grau de regularidade nas variações a longo prazo;
- Variações Aleatórias: que são ruídos nos dados das séries decorrentes de fatores não identificados.

Conforme Xavier (2016), as principais fontes de variação na maioria das séries temporais são: a tendência e a sazonalidade, e ambas costumam ser investigadas frequentemente durante a execução das análises exploratórias de dados.

2.6.1. Tendência

A fim de examinar a tendência em uma série, podem ser aplicadas algumas técnicas, como mínimos quadrados, que possui uma análise de fácil entendimento, uma vez que, apresenta uma reta de regressão com um valor numérico positivo ou negativo para o coeficiente angular de regressão, podendo ser um indicativo de tendência significativa. Por sua vez, a confiabilidade do coeficiente de regressão pode ser verificada por um teste de hipótese, neste caso, é utilizado o teste t de Student, julgando, assim, a confiabilidade da tendência encontrada (LIMA, 2008).

Entretanto, o método dos mínimos quadrados, assim como outros métodos que buscam ajustar os dados em função do tempo, como a regressão polinomial, são paramétricos, ou seja, exigem que os dados da série a ser analisada apresentem uma distribuição conhecida ou previamente especificada (NAGUETTINI; PINTO, 2007). Do mesmo modo, além do teste t de Student, também existem outros testes paramétricos para verificar a significância da tendência, por exemplo: teste Z , teste t pareado e teste t para duas amostras (LIMA, 2008; USP, 2023).

De outro lado, existem métodos não-paramétricos de estimativa da tendência, que buscam suavizar os valores da série ao redor de um ponto, para estimar a tendência naquele ponto. Neste caso,

podem ser citados os métodos de suavização por médias móveis, suavização por medianas móveis e também o método de *lowess* (FELICIANI *et al.*, 2012; CARVALHO; KUPFER, 2007).

Conforme Morettin e Toloi (2006), a maioria dos testes não-paramétricos para estimativa de tendência são baseados na independência das observações, devido à suposição de que os dados constituem uma amostra provinda de uma população. Dentre os testes não paramétricos, os mais utilizados para detectar tendências estão: Wald-Wolfowitz (1943), o teste de Run, Cox-Stuart (1955), Mann-Kendall e o teste de Spearman (LIMA, 2008; XAVIER, 2016).

2.6.2. Sazonalidade

A sazonalidade de uma série temporal pode ser classificada em: determinística, quando o padrão sazonal é regular e estável no tempo, podendo ser prevista perfeitamente a partir de dados de meses anteriores, ou estocástica, quando o padrão sazonal varia com o tempo. Os métodos de regressão são ótimos para séries que apresentam sazonalidade determinística e o método de médias móveis é mais apropriado quando a componente sazonal possui comportamento estocástico (MORETTIN; TOLOI, 2006; XAVIER, 2016).

Na estimativa da sazonalidade podem ser aplicados testes paramétricos, como o teste F associado a uma análise de variância (ANOVA). Entretanto, para aplicação desse teste é necessária a validação do modelo aditivo, incluindo a normalidade dos resíduos, ou seja, devem apresentar comportamento semelhante à normal padrão (Figura 2.2b).

Morettin e Toloi (2006) afirmam que também podem ser utilizados testes não-paramétricos de inferência, como Kruskal-Wallis, para amostras independentes, e o teste de Friedman, para amostras dependentes ou relacionadas, neste último, os meses são considerados “tratamentos” e os anos são avaliados como “blocos”.

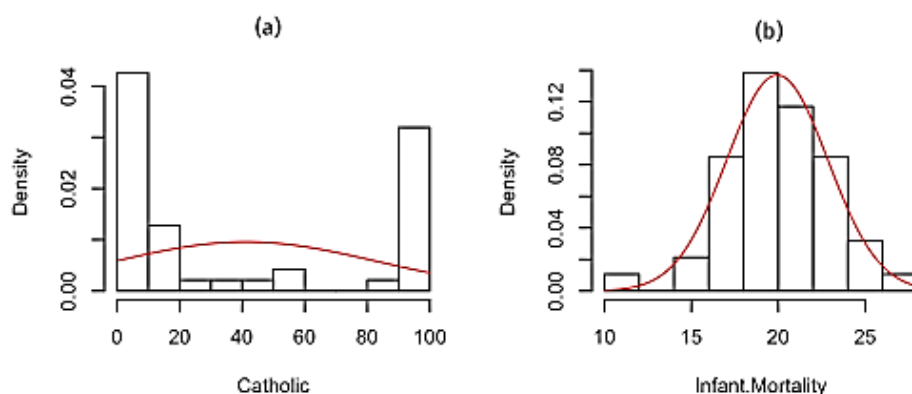


Figura 2.2 – (a) Ausência e (b) Presença de normalidade nos dados (LEITE, 2018)

2.6.3. Estacionariedade

Conforme Chatfield (2000), os métodos clássicos costumam apresentar desempenho satisfatório quando as características de tendência e sazonalidade possuem uma variação predominantemente regular. Por outro lado, tais métodos não tendem a funcionar com o mesmo nível de eficiência quando a tendência e/ou os efeitos sazonais mudam ao longo do tempo, ou quando existe correlação entre sucessivos valores nas flutuações irregulares das séries temporais. O mesmo autor afirma que a maioria das séries temporais são estocásticas (ou aleatórias), cujos valores futuros podem ser parcialmente determinados por valores passados. Estas séries podem ser representadas por meio de modelos denominados processos estocásticos.

Segundo Pereira (2021), uma importante propriedade a ser considerada na modelagem de um processo estocástico corresponde à estacionariedade dos dados. Morettin e Toloi (2018) afirmam que uma série estocástica é estacionária se a sua média, variância e covariância (em diferentes fases) permanecem as mesmas (Tabela 2.2), conforme ilustrado na Figura 2.3b.

Tabela 2.2. Características identificadas nas séries temporais estacionárias

Propriedade	Expressão
Média	$E(Y_t) = \mu$
Variância	$var(Y_t) = (Y_t - \mu)^2 = \sigma^2$
Covariância	$\gamma_k = E[(Y_t - \mu)(Y_{t-k} - \mu)]$

Fonte: Elaborado pela autora, baseado em Valer, 2016.

Obs.: a variável Y_t é uma série temporal estocástica; γ_k a covariância na defasagem k , é a covariância entre os valores Y_t e Y_{t-k} . Portanto, uma amostra de dados é estacionária quando seus parâmetros característicos se mantêm fixos e constantes ao longo do período de sua observação (VALER, 2016).

Por outro lado, conforme Xavier (2016), muitas vezes, as séries apresentam algum formato de não-estacionariedade, quando a média e/ou variância variam com o tempo. Nos casos mais raros e complexos, Morettin e Toloi (2018) afirmam que, embora possa existir tendência explosiva dos dados, na maioria das situações ocorre tendência linear, na qual a série flutua ao redor de uma reta, com inclinação positiva ou negativa (Figura 2.3a).

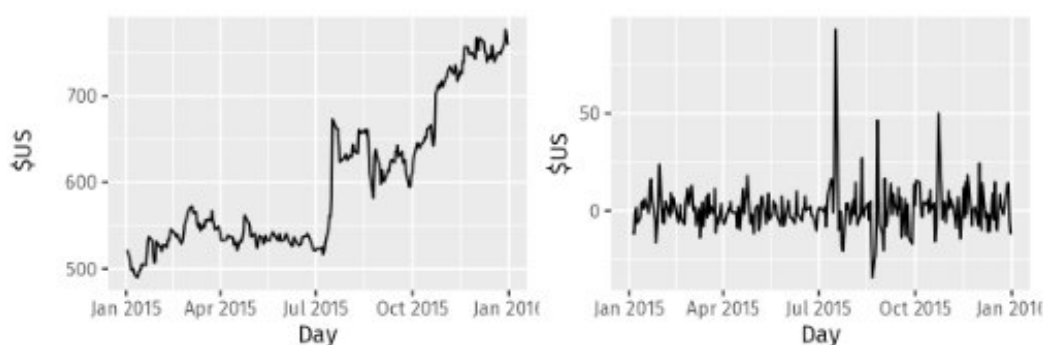


Figura 2.3 – (a) Ausência e (b) presença de estacionariedade nas séries (HYNDMAN; ATHANASOPOULOS, 2021)

De acordo com Camelo (2018), existem diferentes métodos para identificação da estacionariedade de uma série temporal: análise gráfica, comparação da média e variância em diferentes períodos de tempo da série temporal, observação das funções de autocorrelação e, também, aplicação de testes de raiz unitária.

Segundo Xavier (2016), deve-se utilizar mais do que um teste de raiz unitária de modo a certificar a estacionariedade de uma série. Alguns dos testes mais utilizados são os testes de Dickey-Fuller (DF) e Dickey-Fuller Aumentado (ADF, DICKEY; FULLER, 1981), Phillips-Perron (PP, PHILLIPS; PERRON, 1988), Kwiatkowski-Phillips-Schmidt-Shin (KPSS, KWIATKOWSKI *et al.*, 1992) e o teste Leybourne-McCabe (LMC, LEYBOURNE; McCABE, 1994), mais detalhes sobre os testes podem ser consultados no trabalho de Fracaro (2018). Também existem testes de raiz unitária aplicados à verificação da estacionariedade em componentes sazonais das séries, dos quais o mais comumente utilizado é o teste Osborn-Chui-Smith-Birchenhall (OCSB, OSBORN *et al.*, 1988), mas também costumam ser aplicados, com a mesma finalidade, os testes: Hylleberg-Engle-Granger-Yoo (HEGY, HYLLEBERG *et al.*, 1990) e Canova-Hansen (CH, CANOVA; HANSEN, 1995), maiores informações podem ser encontradas no estudo de Sieben (2017).

2.7. Modelagem Box-Jenkins e Box-Tiao de séries temporais

2.7.1. Modelos Box-Jenkins

No início de 1970, George Box e Gwilym Jenkins formularam um dos modelos de séries temporais mais importantes para previsão, trata-se do modelo Auto-Regressivo Integrado de Médias Móveis (ARIMA), também conhecido como modelo de Box-Jenkins (BOX *et al.*, 2015), o qual, por sua vez, é considerado um modelo estatístico e univariado, ou seja, possui a entrada de padrões identificados no comportamento da própria variável preditora ao longo do tempo para realização de sua projeção futura.

De acordo com Gujarati e Porter (2010), o modelo ARIMA possui características flexíveis, permitindo ajustes e adaptações em seus parâmetros além de ter a capacidade de descrever tanto processos lineares estacionários, como também, não-estacionários (CAMELO, 2018).

Em processos lineares estacionários, é possível utilizar basicamente três tipos de modelos:

- Processo Auto Regressivo de ordem p , representado por AR (p);
- Processo de Médias Móveis de ordem q , representado por MA (q);
- Processo Auto Regressivo e de Médias Móveis de ordem p e q , representado por ARMA (p, q).

Nos casos de séries lineares não-estacionárias, é necessário torná-las estacionárias e, para realizar este procedimento, deve-se aplicar um número determinado de diferenças entre seus dados de ordem (d). Geralmente, são realizadas uma ou duas diferenciações (CAMELO, 2018).

Nestes processos, o modelo é comumente representado por ARIMA (p, d, q), que pode ser descrito conforme mostra a Eq. (2.7.1):

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \quad (2.7.1)$$

em que: y'_t representa a série diferenciada (pode ter sido diferenciada mais de uma vez), os “preditores” no lado direito incluem valores defasados e erros defasados de y_t (distanciados no tempo). Este modelo pode ser chamado de ARIMA (p, d, q), onde: p representa a ordem da parte autoregressiva; d , o grau de diferenciação envolvido, e; q , a ordem da parte de médias móveis. Maiores detalhes sobre a equação podem ser conferidos em Hyndman e Athanasopoulos (2021).

O modelo ARIMA (p, d, q) pode ainda, ter incluído um operador sazonal, sendo representado então por SARIMA (p, d, q)(P, D, Q) $_s$, em que a ordem P, D e Q representam componentes sazonais relativos aos termos Auto-Regressivo, Integrado, e Média Móvel, respectivamente.

Para realizar a construção de um modelo Auto-Regressivo Integrado de Médias Móveis (ARIMA) devem ser aplicadas as etapas exibidas na Figura 2.4, as quais apresentam um comportamento cíclico até a aquisição do modelo apto à previsão.

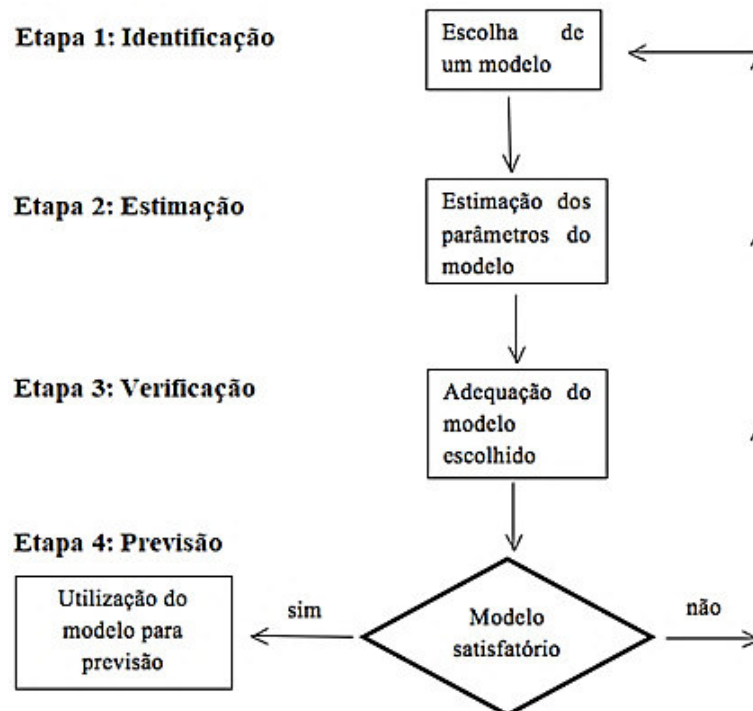


Figura 2.4 - Etapas de configuração do modelo ARIMA para previsão de séries temporais (CAMELO, 2018)

Na etapa de identificação, após a definição da ordem de integração da série, passa-se a trabalhar com os dados estacionários, uma vez que, os dados das séries não-estacionárias podem ser diferenciados uma ou mais vezes, com base no procedimento da Eq. (2.7.2):

$$Z_t = \Delta Y_t = Y_t - Y_{t-1}, \quad (2.7.2)$$

em que: Y_t é a observação Y , no período t da série Y_t sem diferenciação; Y_{t-1} é a observação Y , no período $t-1$ da série Y_t sem diferenciação; $\Delta Y_t = Z_t$ é a observação Z , no período t , pertencente à série Z_t com os dados da série Y_t diferenciados à primeira vez.

Em seguida, inicia-se a fase de identificação dos filtros AR e MA, ou seja, dos valores das ordens p e q do modelo ARIMA. Esta última pode ser realizada com base na observação das características da Função Autocorrelação (FAC) e da Função Autocorrelação Parcial (FACP) dos dados da série analisada.

As funções FAC e FACP, quando representadas graficamente, são comumente chamadas de correlogramas, os quais também são conhecidos como diagramas de autocorrelação (representam a autocorrelação da amostra versus os intervalos de tempo considerados na amostragem). Ao examinar os picos dos correlogramas FAC e FACP em cada *lag* (termo que significa defasagem em inglês)

poderão apontar quais são significativos e, desta forma, indicar as ordens de p e q , uma vez que um pico significativo se estende para além dos limites do nível de significância utilizado, indicando que a correlação para este *lag* é significativa. Desta forma, com base nas análises descritas na Tabela 2.3, podem ser identificados os termos AR e MA dos modelos candidatos ARIMA.

Tabela 2.3. Características da Função Autocorrelação e da Função Autocorrelação Parcial

Modelo	Padrão Típico da FAC	Padrão Típico da FACP
AR (p)	Declina exponencialmente ou com padrão de onda senoidal, ou ambos.	Picos significativos através das defasagens p .
MA (q)	Picos significativos através das defasagens q .	Declina exponencialmente.
ARMA (p, q)	Declina exponencialmente.	Declina exponencialmente.

Fonte: Camelo, 2018.

Segundo Pereira (2021), após a identificação dos modelos candidatos, é feita a estimação dos parâmetros do modelo a ser selecionado, a qual pode ser obtida por diferentes maneiras: mínimos quadrados ordinários, mínimos quadrados generalizados, máxima verossimilhança condicional ou máxima verossimilhança exata, entre outros.

Dentre estas opções, é importante ressaltar que não existem modelos verdadeiros, mas sim apenas modelos aproximados da realidade, que causam perda de informações (EMILIANO, 2010) e, de acordo com Bueno (2011), o melhor modelo será o mais parcimonioso, ou seja, aquele que apresenta um menor número de parâmetros no modelo e, com isso, seja capaz de gerar estimativas com menores imprecisões do que modelos que apresentem mais parâmetros inseridos.

A parcimonia é uma propriedade que pode ser verificada pelos critérios de informação. Dentre os mais explorados pela bibliografia estão: o critério de informação de Akaike (AIC), Akaike corrigido (AICc), Critério Bayesiano de Schwarz (BIC) e de Hannan-Quinn (HQ), os quais estão dispostos na Tabela 2.4.

Tabela 2.4. Critérios de informação

Critérios de Informação	Definição	Fonte
AIC	$AIC = -2\log(L) + 2w$	Akaike (1974)
AICc	$AIC_c = -2\log(L) + \frac{2nw}{n - w - 1}$	Sugiura (1978) Hurvich; Tsai (1989)
BIC	$BIC = -2\log(L) + w \log(n)$	Akaike (1978) Schwarz (1978)
HQ	$HQ = -2\log(L) + 2w \log(n)$	Hannan; Quinn (1979)

Fonte: Elaborado pela autora, baseado em Guerra, 2015.

Obs.: o primeiro termo consiste numa estimativa da distância K-L por meio de $\log(L)$, que representa a log-verossimilhança do modelo estimado, e $2w$ é uma estimativa para o viés, em que w representa o número de parâmetros do modelo, e n equivale ao número de observações. Tais critérios são baseados na distância Kullback-Leibler (K-L) entre o modelo candidato e o verdadeiro e empregam a estimativa do log da máxima verossimilhança, penalizando os ajustes dos modelos pelo número de parâmetros estimados (GUERRA, 2015; PEREIRA, 2021).

Após a seleção do modelo candidato mais parcimonioso, é necessário avaliar a aderência deste modelo aos dados da série em questão. Uma das formas é avaliar se os resíduos do modelo apresentam um comportamento de ruído branco e, desta forma, atendam os seguintes pressupostos: são independentes, ou seja, não são autocorrelacionados, possuem média zero e variância constante, em outras palavras, apresentam homoscedasticidade e ainda devem seguir uma distribuição normal. Na situação de confirmação dos pressupostos, significa que o modelo é adequado aos dados da série analisada, caso contrário, existem informações nos dados que o modelo não conseguiu abranger, tornando-o insuficiente para ser aplicado na previsão. Sendo assim, é importante testar outras possibilidades para o modelo estimado (BUENO, 2011; CAMELO, 2018; PEREIRA, 2021).

Na análise dos resíduos, a identificação dos pressupostos, em geral, é realizada por meio de testes de inferência aplicados aos resíduos. Na verificação da normalidade, conforme Cardoso *et al.* (2023), destacam-se os testes de Doornick-Hansen, Kolmogorov-Smirnov, Kolmogorov-Smirnov Lilliefors, Shapiro-Wilk (1965) e Jarque-Bera (1986). Em relação à homoscedasticidade, são salientados na literatura, os testes de Goldfeld-Quandt, Breusch-Pagan, McLeod-Li e o teste de Heterocedasticidade Condicional Autoregressiva (ARCH), e; para independência dos resíduos, há relevância nos testes de: Durbin-Watson, Box-Pierce, Breusch-Godfrey e Ljung-Box (1978).

2.7.2. Modelos Box-Tiao

Outra metodologia de previsão de séries temporais, que utiliza técnicas de Auto-Regressão, é a modelagem Box-Tiao, a qual é comumente conhecida por ARIMAX, abreviação procedente do termo Auto-Regressivo Integrado de Médias Móveis e entradas Exógenas, que representa uma expansão do modelo ARIMA. O modelo ARIMAX, considerado multivariado, acrescenta um componente linear em função das observações das covariáveis (também chamadas de variáveis exógenas). A principal diferença entre os dois modelos é que o ARIMAX possui, além dos parâmetros auto-regressivos e de médias móveis, a entrada exógena (BENNETT; STEWART; LU, 2014).

O modelo ARIMAX pode ser entendido como a combinação dos modelos Auto-Regressivo AR (p), Integrado (d), Média-Móvel MA (q) e Exógeno X (r), simbolizado por ARIMAX (p, d, q, r), pois, a obtenção dos termos p, d, q é semelhante à descrição realizada no modelo ARIMA; já o termo r representa a quantidade de variáveis exógenas utilizadas na configuração do ARIMAX. Da mesma forma que o SARIMA representa uma versão sazonal do modelo ARIMA, o mesmo pode se dizer em relação ao ARIMAX, cuja versão sazonal é simbolizada por SARIMAX(p, d, q, r)(P, D, Q)_s. Uma forma simplificada desse modelo está descrita na Eq. (2.7.3).

Ressalta-se que as técnicas utilizadas no modelo ARIMA para identificar se os resíduos tem o comportamento de ruído branco, como, por exemplo, a aplicação dos critérios de informação, também podem ser utilizados para identificar a viabilidade do modelo ARIMAX em representar os dados observados da série analisada.

$$Y_t = c + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=1}^r \omega_j w_j + \sum_{j=1}^q (\theta_j \varepsilon_{t-j}) + \varepsilon_t, \quad (2.7.3)$$

em que: Y_t é a série da variável no tempo t que se pretende projetar; c é uma constante; Y_{t-i} é a série da variável dependente defasada por i passos de tempo; β_i é o coeficiente de Y_{t-i} ; p é o número máximo de intervalos de tempo; w_j representa as variáveis exógenas; o termo ω_j representa os coeficientes das variáveis exógenas; r é o número máximo de variáveis exógenas; θ_j é o coeficiente do termo ε_{t-j} que, por sua vez, representa o erro no tempo t defasado de j ; ε_t é a componente de erro do modelo, com $\varepsilon_t \sim N(0, \sigma^2)$, um ruído branco. Os coeficientes dos modelos são estimados por regressão, maiores detalhes sobre a Eq. (2.7.3) podem ser obtidos em (YOGARAJAH *et al.*, 2013).

Segundo Morais Jr. (2011), o sucesso de um modelo depende diretamente da seleção ou escolha das variáveis, entradas e saídas, porque esta vai garantir uma boa confiabilidade, além de manter as características de predição do modelo.

Desta forma, existem diferentes técnicas de seleção e definição de variáveis exógenas para entrada de modelos do tipo ARX, ARMAX, SARIMAX: métodos de busca exaustiva, que consideram todas as combinações possíveis entre as variáveis e são recomendados em sistemas com poucas variáveis de entrada, em vista do alto custo computacional e de um tempo de convergência elevado; e, métodos evolutivos ou clássicos, que consistem em propor um modelo dividido em etapas e a cada etapa é introduzida uma nova variável, as etapas acabam quando não ocorre mais melhora significativa no modelo ou quando todas as variáveis já foram incluídas. Alguns modelos utilizam a remoção de variáveis ao invés de adição e outros combinam as duas técnicas.

Dentre os métodos de seleção de variáveis, segundo Morais Jr. (2011), podem ser destacados:

- Adição sequencial ou *Forward Selection (FS)*: a construção do modelo se dá em etapas e a cada etapa uma variável é introduzida;
- Seleção por Eliminação ou *Backward Elimination (BE)*: inicia-se com um modelo que considera todas as variáveis candidatas e realiza sucessivas remoções de variáveis do modelo com o objetivo de obter o melhor subconjunto de variáveis exógenas;
- Regressão por Etapas ou *Stepwise Regression (SR)*: consiste em uma combinação dos métodos anteriores e é aplicado para selecionar as variáveis que mais influenciam no conjunto de saída, é um processo iterativo que adiciona (*forward*) e remove variáveis (*backward*) a partir de um critério de seleção, como o teste F de inferência, o erro quadrático total e o critério de informação de Akaike (ALVES *et al.*, 2013);
- Regressão por etapas nas predições do modelo *Stepwise Regression on Model Predictions (SRMP)*: constrói um modelo de forma gradual, a fim de analisar a importância de cada uma das variáveis individualmente. Nesta técnica, as variáveis são divididas em dois conjuntos: o conjunto básico, que inicia vazio e finaliza comportando as variáveis que compõem o modelo final; e, o conjunto não-básico, o qual inicia com todas as variáveis candidatas, sofre o processo de eliminação (*backward*), e finaliza não comportando as variáveis selecionadas para o modelo final. O deslocamento das variáveis é determinado com base no efeito que as mesmas exercem na capacidade preditiva do modelo final.

As variáveis exógenas também podem ser selecionadas por processos de busca aleatória. Dentre os mais conhecidos está o Algoritmo Genérico, muito utilizado em situações de busca exaustiva impraticável, ou seja, quando há um grande número de atributos a serem selecionados.

Outras técnicas também podem ser verificadas, como Análise de Componentes Principais, Decomposição de Valores Singulares e análises de Sensibilidade. Mais detalhes podem ser conferidos no estudo de Morais Jr. (2011).

2.7.3. Validação e Intervalo de Confiança dos Modelos

Segundo Montgomery *et al.* (2012), assim como a modelagem Box-Jenkins, todos os modelos de regressão são utilizados para previsão e controle de dados, requerendo uma validação de modo a verificar sua eficiência na produção de respostas sobre o conjunto de dados.

Na validação ocorre a comparação entre os dados simulados pelo modelo e os dados observados da série analisada, e a diferença entre seus valores é denominado erro numérico de predição. Então, para realizar a mensuração dos erros de predição, ao longo de um período temporal, são aplicadas medidas de acurácia, que são indicadores de qualidade e avaliam a capacidade preditiva do modelo (Tabela 2.5).

Tabela 2.5. Indicadores de qualidade de modelos quanto à capacidade preditiva

<i>Medidas Estatísticas de Erros</i>	
Medida	Expressão
MAE	$MAE = \frac{1}{H} \sum_{t=1}^H (Y_t - \hat{Y}_t) $
MAPE	$MAPE = \left(\frac{1}{H} \sum_{t=1}^H \left \frac{Y_t - \hat{Y}_t}{Y_t} \right \right) \cdot 100$
RMSE	$RMSE = \left[\sum_{t=1}^H \frac{Y_t - \hat{Y}_t}{H} \right]^{\frac{1}{2}}$
U de Theil	$U = \sqrt{\frac{\sum_{t=1}^{H-1} \left(\frac{\hat{Y}_{t+1} - Y_{t+1}}{Y_t} \right)^2}{\sum_{t=1}^{H-1} \left(\frac{Y_{t+1} - Y_t}{Y_t} \right)^2}}$

Fonte: Elaborado pela autora, baseado em Pereira, 2021.

Obs.: Y_t são as observações na série temporal; $\hat{Y}_t$ são os valores previstos para a série; e, H é o número total de observações no período de validação da série.

Conforme Camelo (2018), em relação às técnicas estatísticas das medidas de acurácia, o Erro Médio Absoluto (MAE) mede o valor médio absoluto dos erros entre as séries temporais (observada e prevista). A Raiz do Erro Quadrático Médio (RMSE) representa as diferenças individuais quadráticas entre as séries temporais (observada e prevista). No RMSE, a proporção entre as variáveis observadas e previstas ocorrem na mesma grandeza, ou seja, se a variável analisada apresentar grandes valores na série, então os erros nas variáveis ajustadas também serão grandes. Por outro lado, os resultados obtidos no RMSE utilizam os quadrados dos resíduos, logo os resíduos possuem um peso maior em sua mensuração. Por este motivo, alguns autores como Morettin e Toloí (2018) consideram o RMSE como a melhor medida de acurácia (PEREIRA, 2021).

Por sua vez, o Erro Médio Absoluto Percentual (MAPE) possui rápida interpretação em seus resultados, por ser apresentados em termos percentuais (%), mas existe uma desvantagem quando a série analisada apresentar valores muito pequenos. Neste caso, qualquer discrepância pode fazer o MAPE alcançar valores superiores a 100%, cujo valor não é aceito nas análises de acurácia dos modelos.

Além desses indicadores de qualidade, segundo Levenbach (2017), existe uma técnica desenvolvido Theil (1966) que também pode ser utilizada, a qual é denominada U de Theil, cujo objetivo é testar se o modelo apresenta uma previsão melhor em relação à previsão ingênua, ou trivial, a qual ocorre quando a estimativa do valor futuro é igual ao valor atual (CARVALHO, 2005).

A capacidade preditiva dos modelos também pode ser testada por meio de índices de confiabilidade dos valores previstos e, neste âmbito, um dos quantificadores mais utilizados corresponde ao intervalo de confiança (IC) dos modelos. A construção de intervalos de confiança serve para sintetizar as incertezas decorrentes da estimação de parâmetros baseados em amostras finitas. Desta forma, as análises feitas a partir da construção de intervalos de confiança possibilitam extrair informações a respeito dos aspectos práticos do setor em que se está trabalhando, além de proporcionar tomadas de decisão com a minimização de riscos em condições de incerteza do analista (SILVA *et al.* 2011; FERREIRA, 2013).

A estimação de intervalos para um quantil E_F , associado a uma probabilidade avaliada de acordo com um dado modelo com w parâmetros, $F_x(x|\vartheta_1, \vartheta_2, \dots, \vartheta_w)$, em que ϑ_w designam os parâmetros, consiste em identificar os limites inferior λ_i e superior λ_s , tornando válida a Eq. (2.7.4).

$$P(\lambda_1 \leq E_F \leq \lambda_2) = 1 - \alpha, \quad (2.7.4)$$

em que: a expressão $1 - \alpha$, equivale ao nível de confiança. Como E_F é uma grandeza populacional e não uma variável aleatória, interpreta-se a Equação 2.7.4 como: a probabilidade do intervalo $[\lambda_i, \lambda_s]$

conter o verdadeiro, embora desconhecido, valor de E_F é $100(1 - \alpha)\%$. A equação reflete as incertezas decorrentes da estimação de parâmetros $\vartheta_1, \vartheta_2, \dots, \vartheta_w$ e, conseqüentemente, de E_F , frente às flutuações amostrais, não incluindo as incertezas associadas à seleção da lei de probabilidades $F_x(x|\vartheta_1, \vartheta_2, \dots, \vartheta_w)$. A princípio supõe-se verdadeira a função de distribuição de probabilidades que governa o comportamento da variável aleatória E .

Conforme Silva *et al.* (2011), a construção de intervalos de confiança para os quantis de uma variável aleatória pode ser feita por meio de três abordagens:

- Aplicação da teoria assintótica de amostragem, na qual para uma dada probabilidade de não-excedência, F , os quantis em torno da estimativa $\hat{X}_F$ são aproximadamente distribuídos de acordo com a lei Normal, ou seja, contendo normalidade nos dados, e desvio padrão dependente da dimensão da amostra e do método de estimação dos parâmetros de distribuição.
- Técnica de Monte Carlo para, a partir do conjunto de estimativas dos parâmetros $F_x(x|\vartheta_1, \vartheta_2, \dots, \vartheta_w)$, obter M (com M muito grande) amostras sintéticas com dimensão igual à da amostra original. Tais amostras sintéticas são utilizadas no cálculo de M curvas de quantis que, depois de ordenadas para uma dada probabilidade de não-excedência, F , permitem a construção do IC: $100(1 - \alpha)\%$ para X_F .
- Técnica de re-amostragem (*o bootstrap* na designação inglesa), a qual permite obter um grande número de amostras (M) a partir de re-amostragens com recomposição da amostra original. Há semelhança com a técnica de Monte Carlo, visto que, as amostras obtidas por re-amostragem são utilizadas na estimativa das M curvas de quantis que, depois de ordenadas para uma dada probabilidade de não-excedência F , permitem a construção do IC: $100(1 - \alpha)\%$ para X_F .

A partir das técnicas revisadas na bibliografia, foram selecionados os métodos para realização do presente estudo, os quais estão descritos no próximo capítulo.

3. METODOLOGIA

Neste capítulo, serão abordados os aspectos metodológicos adotados para o desenvolvimento da pesquisa, tais como: origem e organização dos dados e as técnicas de modelagem utilizadas. Em relação aos recursos computacionais foi utilizada a planilha eletrônica *Microsoft Excel* (2019) e para análises estatísticas, representações gráficas e a estimação de modelos de previsão, foi utilizado o *software R* sob a interface *RStudio* versão 4.1.0 (R DEVELOPMENT CORE TEAM, 2022). O nível de significância considerado neste estudo foi de 5% ($\alpha = 0,05$).

3.1. Origem e Preparo dos dados de estudo

As etapas desenvolvidas neste item, dispostas na Figura 3.1, foram necessárias para adquirir e efetuar o tratamento dos dados, a fim de possibilitar a aplicação destas informações na entrada dos modelos.



Figura 3.1 - Esquema das etapas utilizadas para aquisição e tratamento dos dados de estudo

Os dados de eletricidade foram do tipo máxima demanda instantânea, em Megawatt (MW), e estão vinculados à carga verificada de referência do barramento pertencente à subestação de energia elétrica da Figura 3.2, localizada na região da campanha do Rio Grande do Sul, cujas informações foram fornecidas pela Companhia Estadual de Energia Elétrica (CEEE).

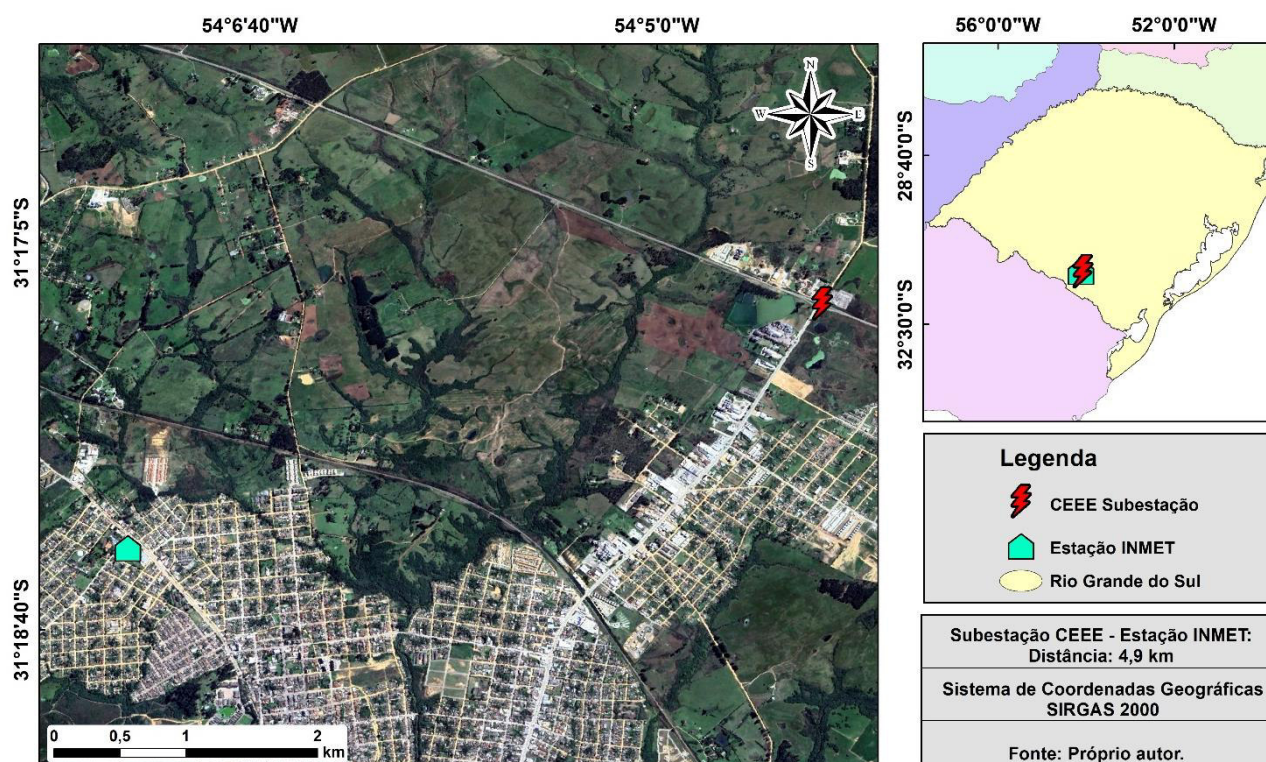


Figura 3.2 - Localização da Estação Meteorológica do INMET e da Subestação da CEEE

Na Tabela 3.1 são apresentados os horários de ocorrência dos patamares de carga de energia leve, média e pesada ao longo de uma semana no estado do Rio Grande do Sul. Ao todo, foram constituídas sete séries temporais de demanda mensal, referentes aos patamares de carga leve, média e pesada dos dias úteis (UL, UM e UP, respectivamente) e, para os sábados e domingos, os patamares de carga média e pesada (SM, SP e DM, DP, respectivamente), no período de janeiro de 2009 até agosto de 2021.

Tabela 3.1. Intervalos dos patamares de carga nos meses janeiro, fevereiro e março

Patamar de Carga	Dias úteis	Sábado	Domingo**
Leve	01:00h às 08:00h*	01:00h às 08:00h	01:00h às 09:00h
Média	09:00h às 17:00h	09:00h às 17:00h	10:00h às 17:00h
Pesada	19:00h às 22:00h	19:00h às 22:00h	19:00h às 22:00h

* Não considera segundas-feiras. ** Nos demais meses do ano: horários de domingo são equivalentes aos de sábado.

Fonte: Elaborado pela autora, 2023.

Em relação aos dados meteorológicos (Tabela 3.2), foram retirados dados horários da estação meteorológica automática do INMET (BDMEP, 2022), localizada a cerca de 4,9 km de distância da subestação da CEEE utilizada no presente estudo.

Tabela 3.2. Descrição das variáveis meteorológicas consideradas neste trabalho

Variável	Abreviação	Descrição	Unidade
Precipitação	Pc	Precipitação total horária	mm
Pressão	Press	Pressão atmosférica ao nível da estação, horária	mB
Radiação	Rad	Radiação global	kJ/m ²
Temperatura	Temp	Temperatura do ar, bulbo seco, horária	°C
Ponto de Orvalho	Orv	Temperatura do ponto de orvalho	°C
Temperatura Máxima	Tmáx	Temperatura máxima da hora anterior	°C
Temperatura Mínima	Tmín	Temperatura mínima da hora anterior	°C
Umidade	UR	Umidade relativa do ar, horaria	%
Rajada	Raj	Rajada máxima de vento, horária	m/s
Velocidade do Vento	VV	Velocidade do vento, horária	m/s

Fonte: Elaborado pela autora, 2023.

A fim de possibilitar a comparação entre as variáveis meteorológicas e os dados de carga máxima instantânea a cada mês, a organização dos dados meteorológicos foi realizada da seguinte maneira: para cada dia, separou-se os dados horários das variáveis meteorológicas em períodos idênticos aos patamares mencionados na Tabela 3.1 e, em seguida, para cada mês, foram extraídos: valores médios de pressão, radiação, temperatura, ponto de orvalho, umidade e velocidade do vento; valores máximos de precipitação, temperatura máxima e rajada, e; valores mínimos de temperatura mínima. Desta forma, foram obtidas as séries mensais dos dados meteorológicos referentes aos períodos de cada patamar de carga elétrica.

Na análise exploratória dos dados, primeiramente foram extraídas das séries meteorológicas e de carga elétrica, as medidas resumo de: valor mínimo, 1º quartil, mediana, média, 3º quartil, valor máximo e número de dados faltantes (*NA's*) provenientes da função *summary*, e com o auxílio da função *na.omit*, também foram extraídos índices de variância e desvio padrão, além dos índices de assimetria e curtose, resultantes dos comandos *skewness* e *kurtosis*, respectivamente, os quais pertencem ao pacote *e1071* (JOANES; GILL, 1998).

Nas séries de demanda de carga elétrica, foram identificados valores plausíveis em todos os patamares de carga, no período entre janeiro de 2009 e dezembro de 2014, período utilizado para análise de tendência e sazonalidade das séries de carga elétrica. Mais informações a respeito dos dados faltantes dos patamares de carga elétrica referentes aos dias úteis podem ser conferidas em Veiga *et al.* (2023).

A identificação da presença de tendência nas séries foi feita primeiramente, em termos de visualização gráfica, elaborando-se uma reta de regressão linear simples, com base no método de mínimos quadrados, através da inserção do argumento *geom_smooth(method=lm, se=FALSE)* nas especificações da função *ggplot*. Na sequência, a significância da tendência nas séries foi avaliada utilizando o teste de Cox-Stuart.

Por sua vez, a sazonalidade foi avaliada visualmente por gráficos de subséries mensais e, analiticamente, pela aplicação da análise de variância dos dados, ANOVA, em blocos, a qual apresenta as premissas normalidade, homoscedasticidade e independência dos resíduos. Tais características foram avaliadas com a aplicação dos testes de Shapiro-Wilk (1965), Teste ARCH e Ljung-Box, respectivamente. Nos casos de descumprimento destes requisitos, aplicou-se o teste de Friedman (Pereira *et al.*, 2021).

Na etapa seguinte, foi realizada comparação das variáveis meteorológicas entre si e também com a demanda de carga elétrica, para o período entre janeiro de 2009 e agosto de 2021, ainda que observações faltantes estivessem presentes, tanto nas séries meteorológicas, quanto na demanda de carga elétrica. Primeiramente, os dados foram colocados em termos de z-score (Eq. 3.1.1), a fim de retirar efeitos de grandeza das variáveis e possibilitar análise conforme a variabilidade dos dados.

$$Z = \frac{Y - \bar{Y}}{S_Y}, \quad (3.1.1)$$

em que: Y equivale à medição observada; $\bar{Y}$, à média da série de dados, e; S_Y , ao desvio padrão da série de dados (CURTIS *et al.*, 2016).

Em seguida, foram elaborados gráficos *boxplot* de cada série e a análise de correlação foi realizada utilizando o coeficiente de correlação de Pearson, pelo qual foram verificadas tanto as correlações entre variáveis meteorológicas, como também, entre cada variável meteorológica e a demanda de energia, dentro dos intervalos de patamar de carga elétrica considerados. Para facilitar a aquisição dos valores de correlação foram utilizadas as funções: *na.omit*, no intuito de omitir as observações faltantes e *kable* do pacote *knitr* (XIE, 2016), que plotou os valores de correlação sob o formato de matriz.

Em relação aos índices de correlação das séries meteorológicas, somente foram consideradas em posteriores análises: as variáveis com coeficiente de correlação de Pearson igual ou superior a 0,05 ($r_p \geq 0,05$) em relação às séries de demanda pertencentes ao respectivo patamar de carga elétrica.

Após a análise de correlação, realizou-se o preenchimento das sete séries de demanda de carga elétrica utilizando a modelagem Box-Jenkins para imputação dos dados, conforme descrição a seguir: primeiramente, buscou-se preencher as lacunas identificadas nas séries utilizando a função *auto.arima* (HYNDMAN; KHANDAKAR, 2008), a qual selecionou o modelo mais ajustado dentre os modelos candidatos, por meio da minimização do BIC (SILVEIRA, 2023), incrementando o argumento *ic = bic* nesta função, e também listou os modelos de saída, aplicando o argumento *trace = TRUE* e, forçou a utilização de todas as observações, aplicando *approximation = FALSE* (Wang. *et al.*, 2006). Sendo assim, a função na forma completa, contendo todos os argumentos, correspondeu à seguinte:

“`> auto.arima(dados, trace = TRUE, approximation = FALSE, ic = "bic")`”;

Os modelos de saída da função *auto.arima*, que apresentaram autocorrelação nos resíduos, foram substituídos por modelos selecionados com base na identificação da independência residual visualizada em gráficos de Função Autocorrelação (FAC) associada à aplicação do teste Ljung-Box.

Em posse dos dados de demanda de carga elétrica preenchidos, aplicou-se a análise de estacionariedade por meio da função *ndiffs* para parte não-sazonal e *nsdiffs* para parte sazonal das séries. Por sua vez, os dados meteorológicos foram analisados contendo observações faltantes, em vista da presença de lacunas em quantidades escassas, as quais foram consideradas insuficientes para afetar o grau de diferenciação das séries meteorológicas.

Os testes de hipótese aplicados na parte não-sazonal corresponderam a: ADF (DICKEY; FULLER, 1981), KPSS (KWIATKOWSKI *et al.*, 1992) e PP (PHILLIPS; PERRON, 1988) e foi considerado tanto dados em nível, como também, dados com tendência para a componente determinística na regressão, da seguinte maneira:

```
“> ndiffs(dados, alpha = 0.05, test = "adf", type = "level");
```

```
“> ndiffs(dados, alpha = 0.05, test = "adf", type = "trend");
```

Na parte sazonal, os testes aplicados foram: OCSB e HEGY utilizando no máximo duas diferenças sazonais por análise, conforme descrito a seguir:

```
“> nsdiffs(dados, alpha = 0.05, test = "ocsb", max.D = 2)”.
```

Na análise de estacionariedade, em detrimento da modelagem Box-Tiao, somente foram imputadas e consideradas exógenas aos modelos, as variáveis meteorológicas com séries de mesmo grau de diferenciação em relação às séries do patamar de carga elétrica correspondente (HYNDMAN; ATHANASOPOULOS, 2018).

Sobre a imputação dos dados das séries meteorológicas, Veiga *et al.* (2021) afirma que a grande maioria dos estudos considera apenas a similaridade temporal dos dados e a proximidade geográfica dos postos meteorológicos, indicando uma recomendação em utilizar dados de estações vizinhas, visto que, tais dados cumprem as prerrogativas temporal e espacial supracitadas. Porém, neste caso, assim como foi efetuado com os dados da estação de estudo, as informações dos postos meteorológicos adjacentes também precisariam ser reorganizadas em patamares, algo que demandaria muito tempo da parte do analista.

Desta forma, o método utilizado na imputação dos dados meteorológicos também correspondeu à modelagem Box-Jenkins, contendo as mesmas etapas empregadas aos dados de demanda de carga elétrica. A aplicação do mesmo método em variáveis de origem divergente se justifica pela obtenção de observações de preenchimento baseadas no comportamento das séries em si, ou seja, sem a necessidade de adquirir informações de demais estações meteorológicas.

3.2. Modelagem

A modelagem aplicada para relacionar a demanda de carga elétrica às variáveis meteorológicas correspondeu a modelos do tipo Auto-Regressivo Integrado de Médias Móveis Sazonal com entradas Exógenas – SARIMAX, cujas etapas desenvolvidas estão dispostas, sob forma esquemática, na Figura 3.3.

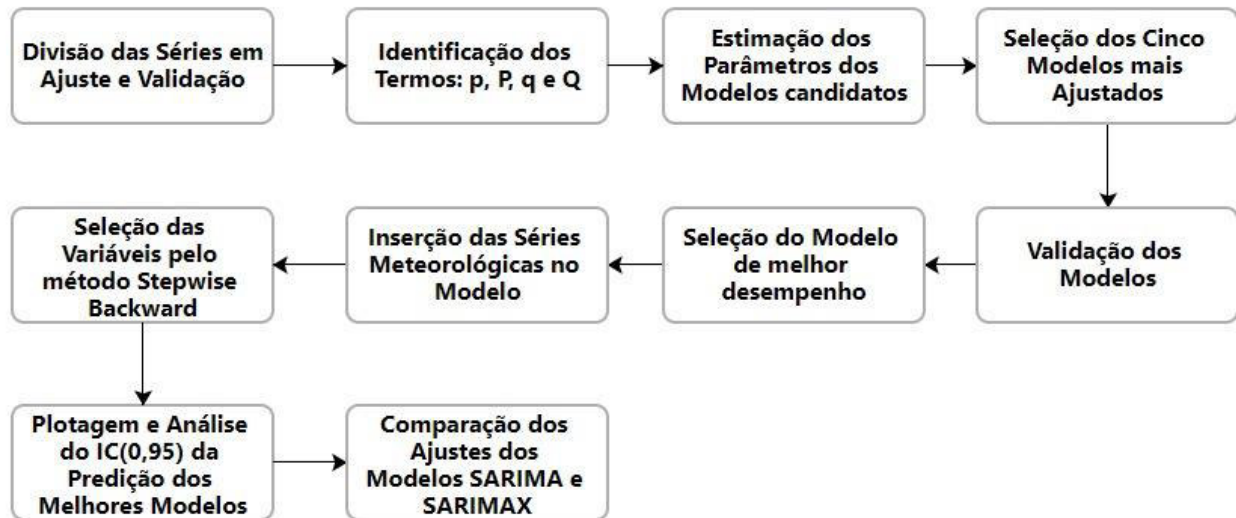


Figura 3.3 - Esquema das etapas utilizadas nas modelagem SARIMAX

Primeiramente, o conjunto de dados foi dividido entre período de ajuste: de janeiro de 2009 até dezembro de 2019, e; período de validação: de janeiro de 2020 até agosto de 2021; os quais também são conhecidos como, períodos de treino e teste: *training set /test set*.

Logo em seguida, foram construídos gráficos de Função Autocorrelação (FAC) e Função Autocorrelação Parcial (FACP) do período de ajuste, a fim de determinar os valores de p, q e P, Q dos modelos candidatos para análise da série de demanda de carga elétrica, os quais foram elegidos com base nos valores de:

- $P \pm 1, q \pm 1$ e $P \pm 1, Q \pm 1$;
- $p, q \pm 2$ e $P, Q \pm 2$, nos casos de interpretações gráficas mais amplas, e;
- $0 \leq p \leq 3, 0 \leq q \leq 3$ e $0 \leq P \leq 3, 0 \leq Q \leq 3$, nas situações onde não foi possível identificar o modelo inicial.

A partir da determinação dos modelos candidatos, estes foram hierarquizados, pela parcimônia, e escolhidos os cinco com menor valor do critério de informação Bayesiano - BIC, indicado por Hyndmann e Athanasopoulos (2021), e com independência identificada em seus resíduos, por meio do teste de hipótese Ljung-Box.

Dentre os cinco modelos mais ajustados aos dados de demanda de energia elétrica, selecionou-se, para análise da influência das variáveis meteorológicas, o modelo de melhor capacidade preditiva, ou seja, o modelo que obteve melhor desempenho e maior acurácia no período de validação, cuja verificação foi feita pelo indicador RMSE.

Na sequência, foi realizada a inserção de variáveis exógenas ao modelo SARIMA selecionado, por meio da adição de vetores e matrizes das séries meteorológicas no argumento *Xreg* da função *Arima*, a fim de torná-lo do tipo SARIMAX, conforme a descrição na Tabela 3.3.

O método de seleção de variáveis adotado foi *stepwise backward*, com a ordem de inserção inicial baseada nos índices de correlação das variáveis meteorológicas entre si e com a demanda de energia elétrica.

O desempenho na validação dos modelos SARIMAX foi verificado com base nos indicadores: RMSE, MAPE e MAE para cada passo da retirada de variáveis. Dentre os modelos com menores índices de erros, para cada patamar, foram plotados os valores preditos, e analisados quanto ao intervalo de confiança (IC) ao nível de 95%, elaborado a partir da teoria assintótica de amostragem.

Tabela 3.3. Configurações do modelo SARIMAX para previsão de séries temporais

<i>SARIMAX Time Series Forecasting</i>	
Função	<i>Arima(y, order = c(0, 0, 0), seasonal = c(0, 0, 0), xreg = NULL)</i>
Argumentos	Descrição
<i>y</i>	Uma série temporal univariada de classe <i>ts</i> .
<i>order</i>	Uma especificação da parte não-sazonal do modelo ARIMA com as três componentes: <i>p</i> , <i>d</i> , <i>q</i> .
<i>seasonal</i>	Uma especificação da parte sazonal do modelo ARIMA com as três componentes: <i>P</i> , <i>D</i> , <i>Q</i> .
<i>xreg</i>	Opcionalmente, um vetor ou matriz de regressores externos, que deve ter o mesmo número de linhas que <i>y</i> .

Fonte: Camelo, 2018

Os modelos SARIMAX de melhor desempenho na validação também tiveram o período de ajuste comparado aos seus respectivos modelos SARIMA, quanto ao valor do BIC e, também, quanto ao cumprimento dos pressupostos dos resíduos: normalidade, homoscedasticidade e independência, por meio da aplicação dos testes de hipótese: Jarque-Bera, teste ARCH e Ljung-Box, respectivamente.

A modelagem supracitada foi realizada separadamente para os patamares de carga leve, média e pesada dos dias úteis (UL, UM e UP, respectivamente) e, para os sábados e domingos, os patamares de carga média e pesada (SM, SP e DM, DP, respectivamente).

4. RESULTADOS E DISCUSSÃO

4.1. Análise exploratória dos dados de demanda de energia elétrica

Os resultados das medidas resumo das séries de demanda de carga elétrica estão dispostos na Tabela 4.1 e os resultados destas medidas de todas as séries analisadas, incluindo as séries meteorológicas, podem ser verificados no Apêndice 1.

A partir da Tabela 4.1 foi possível observar que o menor índice de demanda de eletricidade foi pertencente ao patamar de carga leve dos dias úteis (UL), correspondente a 13,62 MW, e o maior índice foi de 79,89 MW, encontrado no patamar de carga média dos dias úteis (UM). Os dias úteis também tiveram o maior índice médio, igual a 52,0 MW, referente à carga pesada.

Tabela 4.1. Medidas resumo dos patamares de carga elétrica: jan./2009 - ago./2021

Medidas resumo	Patamares de Carga Elétrica (MW)						
	Dias Úteis			Sábados		Domingos	
	Leve	Média	Pesada	Média	Pesada	Média	Pesada
Valor Mínimo	13,62	25,71	25,72	25,64	25,72	18,66	25,71
1° Quartil	23,96	43,23	46,72	36,87	45,99	33,17	43,43
Mediana	25,96	47,78	50,50	39,84	49,39	35,41	46,72
Média	26,94	50,12	52,00	41,47	49,77	37,17	47,42
3° Quartil	29,34	55,44	57,18	44,63	53,66	40,14	50,74
Valor Máximo	42,58	79,89	72,08	72,72	62,88	57,89	63,93
Variância	24,31	104,68	55,26	59,91	31,74	43,53	34,30
Desvio Padrão	4,93	10,23	7,43	7,74	5,63	6,60	5,86
Assimetria	0,68	0,65	0,10	1,30	-0,63	0,90	0,10
Curtose	0,95	-0,01	0,46	2,58	2,50	1,12	1,69
NA's	19	20	17	20	19	16	17

Fonte: Elaborado pela autora, 2023.

Na análise de tendência dos dados, entre janeiro de 2009 e dezembro de 2014, foi observada uma inclinação ascendente da reta de regressão linear, presente nos gráficos em todos os patamares de carga (Figuras 4.1, 4.2 e 4.3), indicando que a demanda de eletricidade aumenta com o decorrer do tempo. Este aumento acontece de forma significativa na maioria dos patamares, conforme indicado pelo teste Cox-Stuart (Tabela 4.2).

Entretanto, o teste de Cox-Stuart apontou uma exceção nos dados referentes aos patamares de carga média de sábados e domingos, onde a hipótese nula foi não pode ser rejeitada e, desta forma, não houve indicação tendência significativa nestes dados.

Tabela 4.2. Resultados do Teste de Cox-Stuart: jan./2009 - dez./2014 ($\alpha = 0,05$)

Patamar de Carga	Estatística do Teste	Graus de Liberdade	valor-p	Hipótese e Significância
UL	26	36	0,01133	H1: Significativo
UM	25	36	0,02882	H1: Significativo
UP	26	36	0,01133	H1: Significativo
SM	23	36	0,13250	H0: Não-Significativo
SP	31	36	0,00001	H1: Significativo
DM	23	36	0,13250	H0: Não-Significativo
DP	28	36	0,00119	H1: Significativo

Fonte: Elaborado pela autora, 2023.

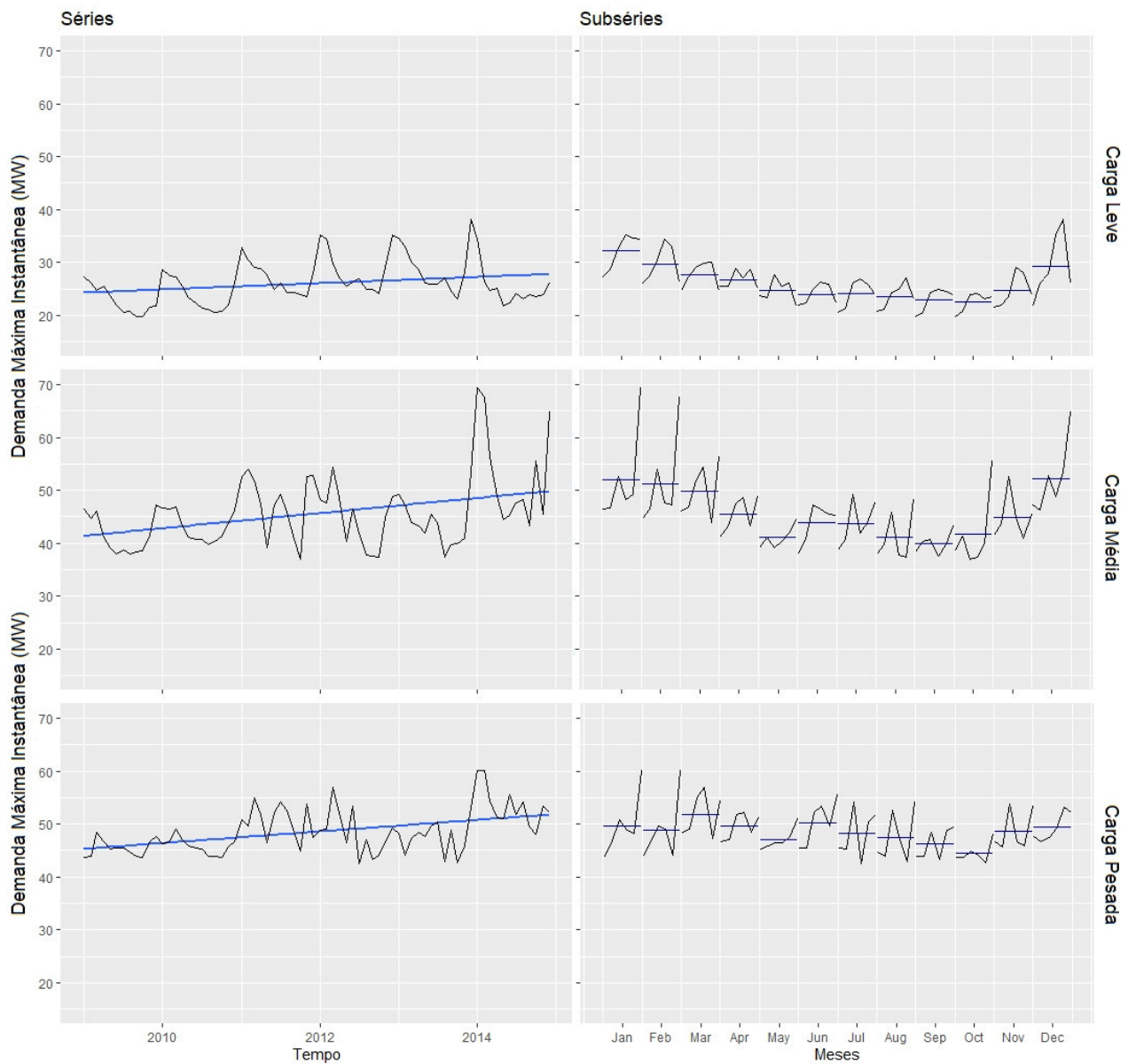


Figura 4.1 – Tendência e subséries mensais dos patamares dos dias úteis: jan./2009 – dez./2014

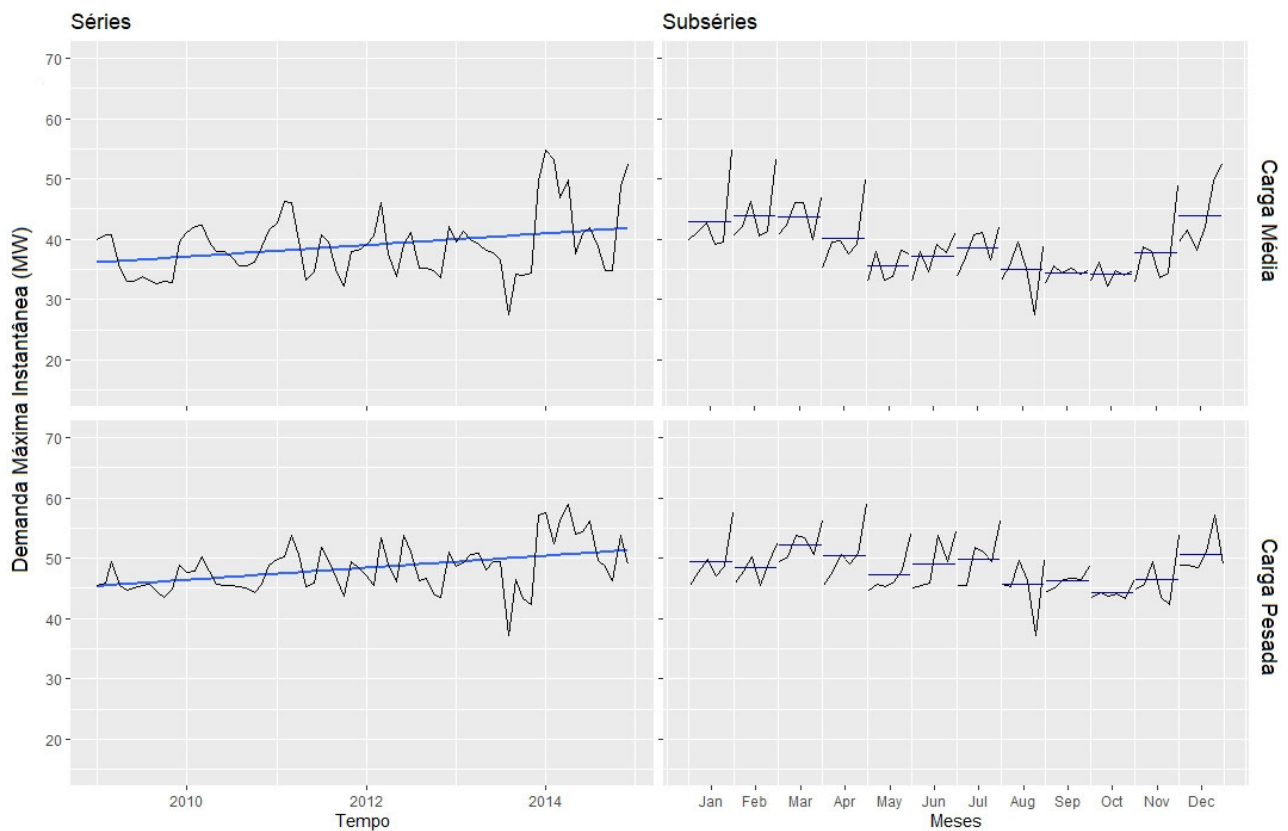


Figura 4.2 – Tendência e subséries mensais dos patamares dos sábados: jan./2009 – dez./2014

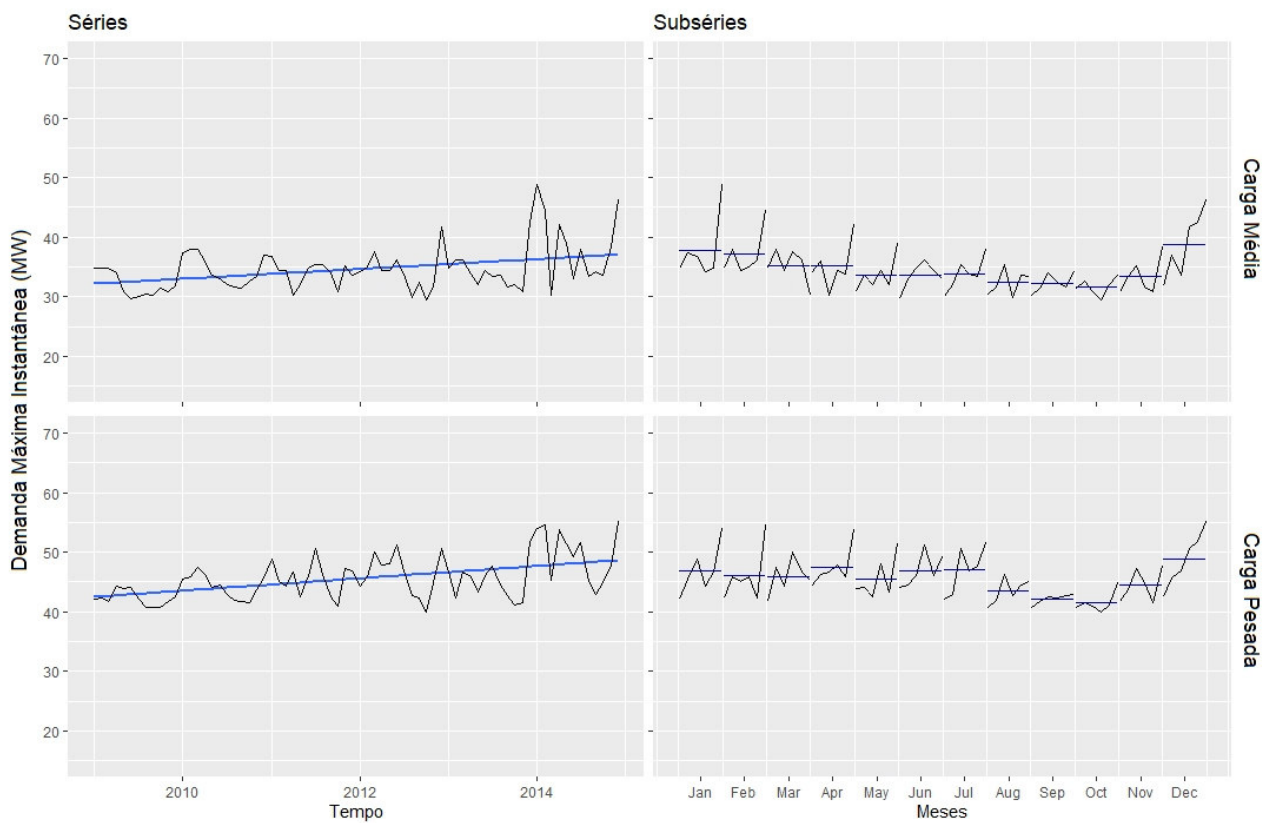


Figura 4.3 – Tendência e subséries mensais dos patamares dos domingos: jan./2009 – dez./2014

Na análise de sazonalidade, do ponto de vista gráfico, foi possível visualizar uma variação na média dos dados conforme o mês do ano em todos os patamares (Figuras 4.1, 4.2 e 4.3), mostrando-se importante verificar a hipótese da sazonalidade estar presente de forma significativa ao longo da variabilidade dos dados das séries de demanda de carga elétrica.

Do ponto de vista analítico, a ANOVA em blocos foi aplicada e validada nas séries de patamar de carga referente aos domingos (DM e DP), cumprindo os requisitos de normalidade, independência e homoscedasticidade. Desta forma, estas séries apresentaram sazonalidade significativa, conforme indicado na Tabela 4.3.

Tabela 4.3. Resultados da ANOVA em blocos: jan./2009 - dez./2014 ($\alpha = 0,05$)

Patamar de Carga	Estatística do teste	Fonte de Variação		Hipótese
		Bloco	Tratamento	
DM	F	7,06	4,13	H1: Significativo
	valor-p	3,66E-05	0,0002	
DP	F	14,06	5,59	H1: Significativo
	valor-p	7,43E-09	6,22E-06	

Fonte: Elaborado pela autora, 2023.

Para séries de demanda dos dias úteis e dos sábados (UL, UM, UP, SM e SP), a ANOVA em blocos não atendeu um ou mais requisitos necessários à análise. Sendo assim, o teste de Friedman (Tabela 4.4) foi aplicado e revelou influência significativa da sazonalidade na variabilidade dos dados nas cinco séries.

Tabela 4.4. Resultados do Teste de Friedman: jan./2009 - dez./2014 ($\alpha = 0,05$)

Patamar de Carga	Estatística do Teste	Graus de Liberdade	valor-p	Hipótese e Significância
UL	53,97	11	1,19E-07	H1: Significativo
UM	49,46	11	7,82E-07	H1: Significativo
UP	29,00	11	0,0023	H1: Significativo
SM	50,72	11	4,65E-07	H1: Significativo
SP	41,05	11	2,36E-05	H1: Significativo

Fonte: Elaborado pela autora, 2023.

4.2. Relação entre os dados meteorológicos e de carga elétrica

Os gráficos *boxplot* foram elaborados para todas as séries: meteorológicas e de carga elétrica; e estão dispostos nas Figuras 4.4, 4.5 e 4.6. A partir destes gráficos, foi possível notar que, em todos os patamares, as variáveis de temperatura apresentaram poucos casos de observações fora do intervalo entre o 1º quartil e o 3º quartil das séries. De outro lado, a precipitação teve numerosas observações

acima do 3º quartil, e as variáveis de vento apresentaram, em todas as situações, valores fora do intervalo entre o 1º e o 3º quartil mínimo e o máximo. Por sua vez, as informações da pressão estiveram próximas à mediana em todas as situações.

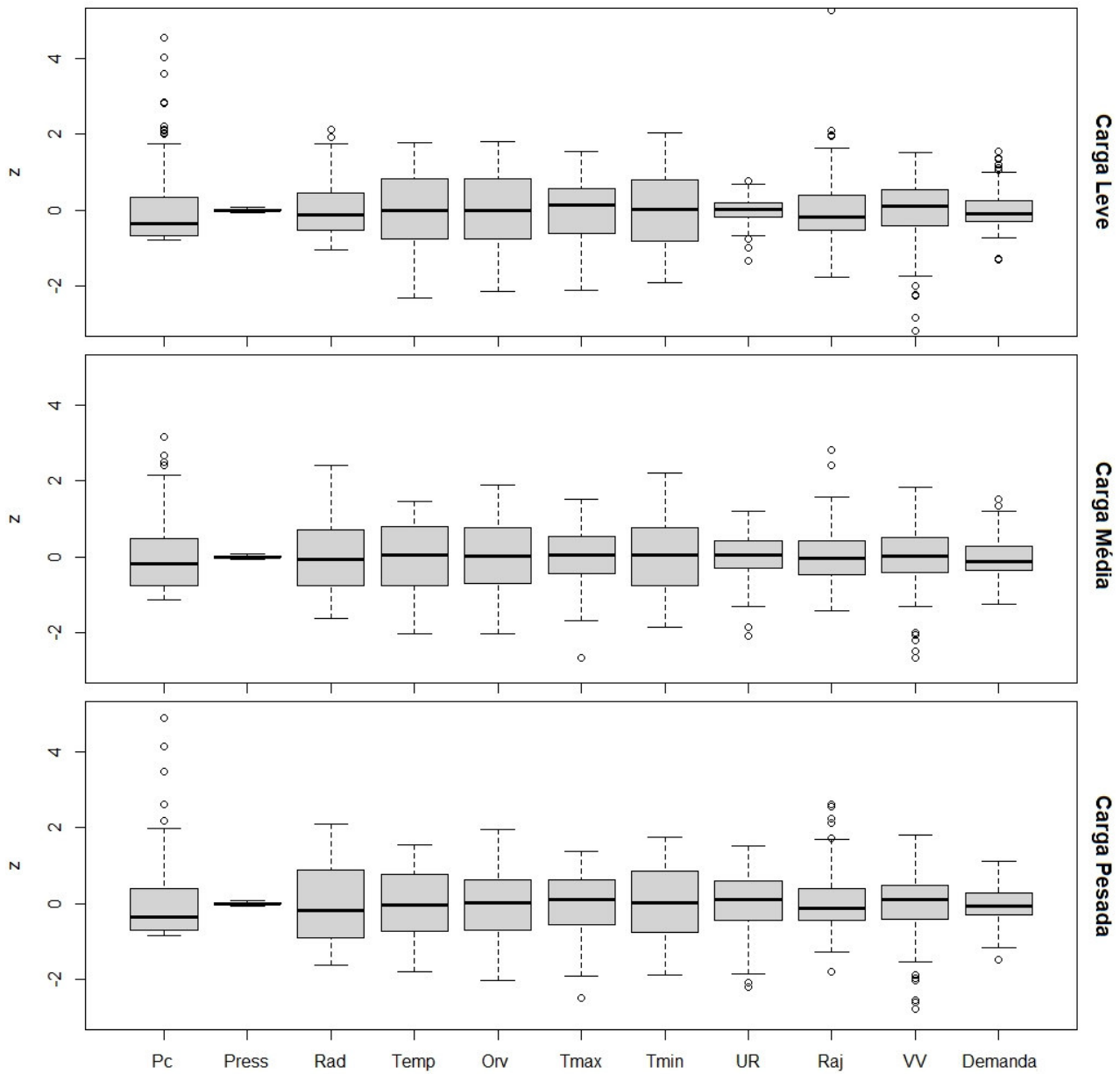


Figura 4.4 – Distribuições dos dados referentes aos dias úteis: jan./2009 – ago./2021

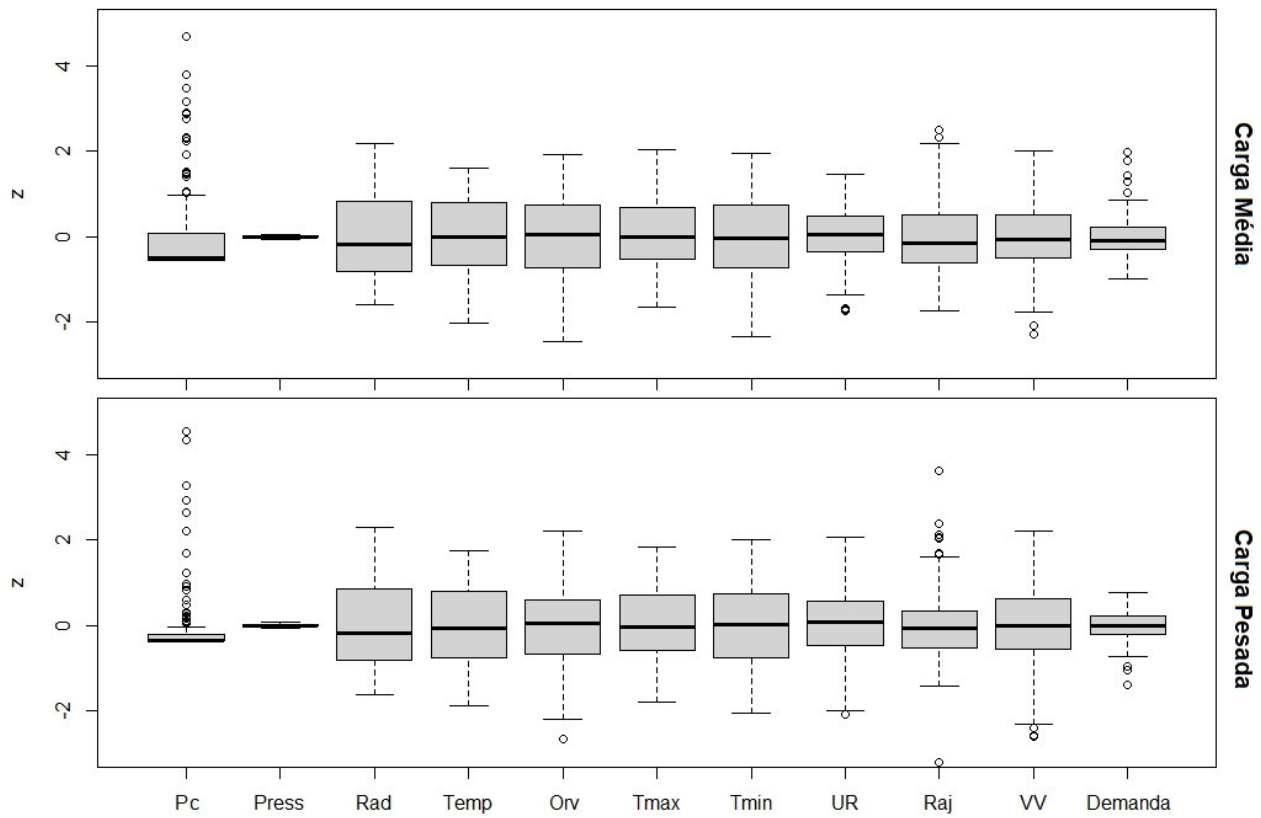


Figura 4.5 – Distribuições dos dados referentes aos sábados: jan./2009 – ago./2021

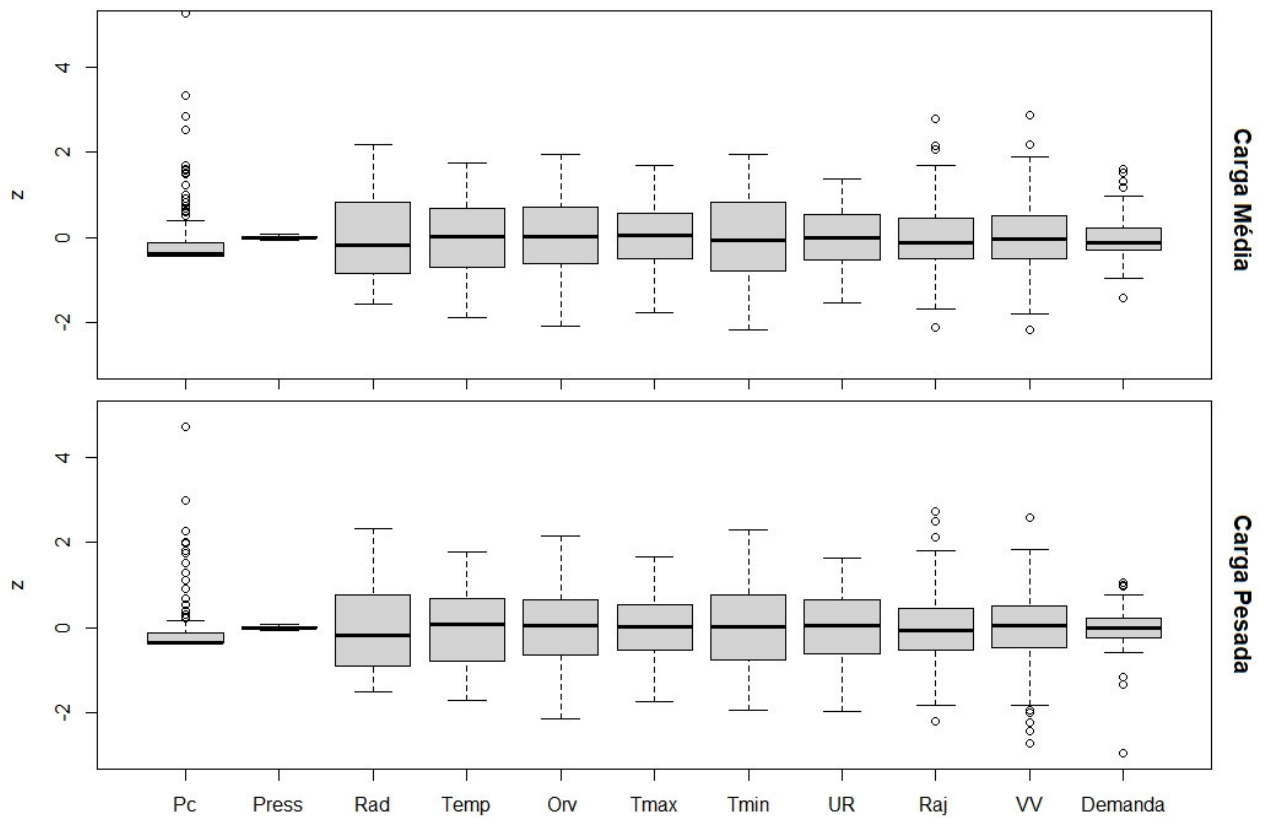


Figura 4.6 – Distribuições dos dados referentes aos domingos: jan./2009 – ago./2021

Os resultados para o coeficiente de correlação entre as variáveis meteorológicas e a demanda de carga elétrica estão dispostos na Tabela 4.5 e, de forma completa, os resultados da análise de correlação das variáveis meteorológicas entre si e também com a demanda de energia dos sete patamares analisados, estão disponíveis no Apêndice 2.

Com base nos resultados da análise de correlação, algumas séries de variáveis meteorológicas apresentaram o valor absoluto do coeficiente de correlação inferior a 0,05 em relação aos dados dos patamares de carga elétrica, e foram descartadas da aplicação em posteriores análises, as quais corresponderam a: precipitação nos patamares: UL, UM SP e DP; pressão em: UP e DP; radiação em: UP e SP, e; rajada no patamar SP. Seguindo tal critério, não apresentaram variáveis descartadas, os patamares de carga média dos sábados e domingos (SM e DM, respectivamente). Dentre as variáveis excluídas, apenas a rajada teve seus dados analisados quanto à estacionariedade, para fins de diagnóstico das variáveis de vento.

Tabela 4.5. Resultados da análise de correlação entre as séries meteorológicas e de carga elétrica

Variáveis Meteorológicas	Patamares de Carga Elétrica						
	UL	UM	UP	SM	SP	DM	DP
Pc	0,04*	0,02*	-0,08	-0,07	0,04*	-0,06	-0,04*
Press	-0,46	-0,33	-0,01*	-0,30	-0,05	-0,19	0,02*
Rad	-0,34	0,41	-0,04*	0,43	0,04*	0,36	0,08
Temp	0,61**	0,55	0,16	0,57	0,22	0,47	0,15
Orv	0,58	0,50	0,16	0,44	0,18	0,41	0,10
Tmáx	0,48	0,52	0,21	0,60	0,25	0,44	0,26
Tmín	0,54	0,56	0,17	0,42	0,15	0,41	0,08
UR	-0,21	-0,33	-0,07	-0,44	-0,17	-0,30	-0,09
Raj	-0,12	-0,34	-0,19	-0,17	0,02*	-0,16	-0,09
VV	-0,13	-0,11	-0,14	-0,15	-0,09	-0,05	-0,23

* Séries descartadas de análises posteriores; ** Maior correlação identificada entre séries meteorológicas e carga elétrica.

Fonte: Elaborado pela autora, 2023.

As séries dos patamares de carga pesada (UP, SP e DP) não se mostraram muito correlacionadas com as variáveis meteorológicas, com índices de correlação inferiores ou iguais a 0,26 nestes casos.

Entre as variáveis meteorológicas, os maiores índices de correlação foram identificados entre a temperatura e as suas derivações: ponto de orvalho, temperatura máxima e mínima (Temp, Orv, Tmáx e Tmín, respectivamente). Em seguida, índices de 0,30 até 0,59 foram observados entre as variáveis de vento: rajada e velocidade do vento.

A maior correlação entre as variáveis meteorológicas e a demanda de carga elétrica foi observada no patamar UL com a variável temperatura, cujo valor do coeficiente foi de 0,61. De forma

geral, em todos os patamares, as maiores correlações com a carga elétrica foram observadas nas variáveis de temperatura e suas derivações (Temp, Orv, Tmáx e Tmín).

4.3. Imputação dos dados de carga elétrica

Dados ausentes foram identificados em todos os patamares de carga elétrica ao longo de todo ano de 2015 e também no período entre setembro e dezembro de 2018. Anomalias não foram observadas nas séries dos patamares de carga média dos domingos (DM).

Conforme o resultado das análises de sazonalidade, tanto ANOVA quanto o teste de Friedman revelaram influência significativa da sazonalidade na variabilidade dos dados em todos os patamares de carga elétrica. Desta forma, os modelos do tipo SARIMA foram utilizados na modelagem dos dados e preenchimento das lacunas existentes nas séries de demanda de carga elétrica, conforme mostra a Tabela 4.6.

Tabela 4.6. Dados faltantes e modelos SARIMA aplicados nas séries de carga elétrica

Patamar de Carga	Dados Ausentes		Anomalias (<i>outliers</i>)	
	Observação ou Período	Modelo SARIMA de imputação	Observação ou Período	Modelo SARIMA de imputação
UL	Jan.-Dez. 2015 Set.-Dez. 2018	(2,1,2) (1,1,1) _[12] *** (2,1,2) (0,1,1) _[12]	Jul. 2016 Out.2016 Fev.2017	(1,1,2) (1,1,2) _[12] (1,1,2) (0,1,2) _[12] (1,1,2) (1,0,1) _[12]
UM	Jan.-Dez. 2015 Set.-Dez. 2018	(1,0,0) (0,1,0) _[12] * (1,1,1) (1,0,0) _[12]	Jul.-Out. 2016	(1,0,1) (0,1,2) _[12]
UP	Jan.-Dez. 2015 ** Set.-Dez. 2018	(1,1,1) (1,0,2) _[12] (2,0,2) (1,0,1) _[12]	Set. 2017**	(1,1,1) (1,0,0) _[12]
SM	Jan.-Dez. 2015 Set.-Dez. 2018	(1,0,0) (1,0,0) _[12] (1,0,1) (2,0,2) _[12]	Jul. 2016 Out.-Dez. 2016	(1,0,0) (1,0,0) _[12] (1,0,0) (1,0,0) _[12]
SP	Jan.-Dez. 2015 Set.-Dez. 2018	(2,1,2) (1,0,0) _[12] (2,0,2) (1,0,2) _[12] *	Jul. 2016 Nov. 2016 Fev. 2017	(2,1,2) (1,0,0) _[12] (2,1,2) (1,0,0) _[12] (1,1,2) (1,0,1) _[12]
DM	Jan.-Dez. 2015 ** Set.-Dez. 2018	(1,1,1) (1,0,1) _[12] (1,1,1) (2,0,0) _[12] *	-	
DP	Jan.-Dez. 2015 ** Set.-Dez. 2018	(1,0,1) (0,0,1) _[12] (2,0,1) (2,0,0) _[12]	Jul. 2016	(2,0,0) (1,0,1) _[12]

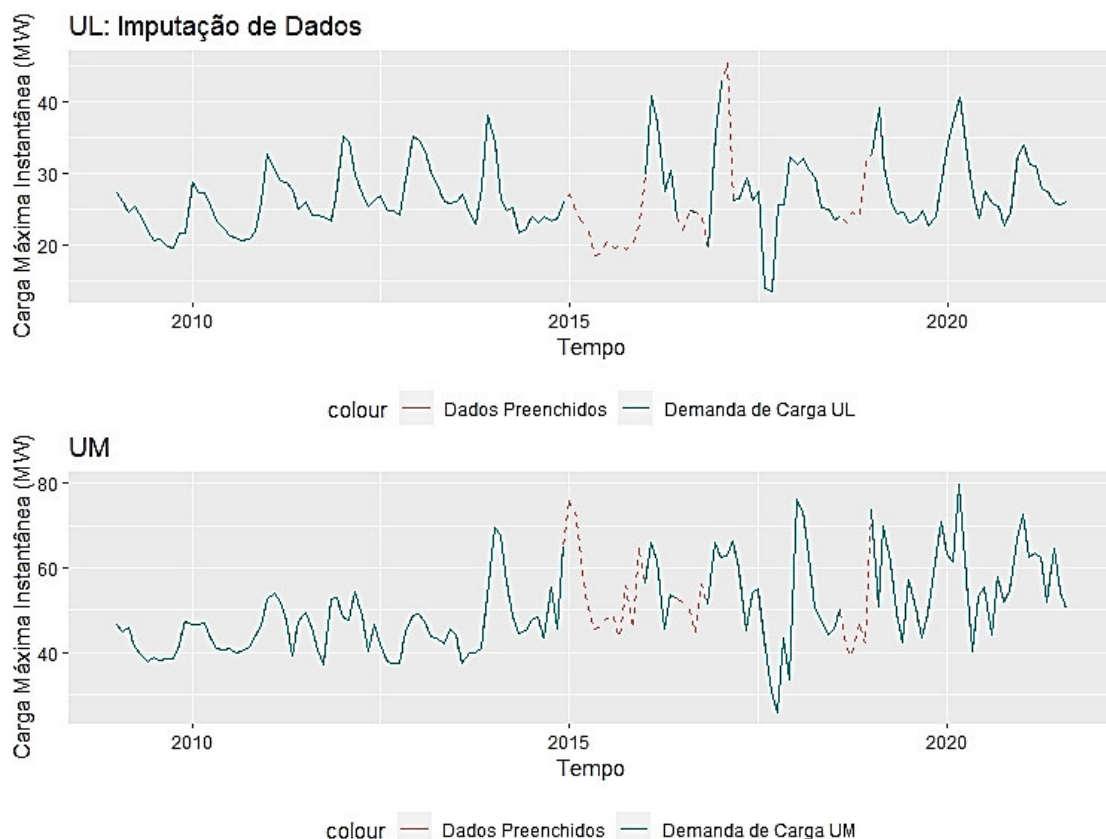
* Modelos de saída da *auto.arima*; ** Casos em que a saída da *auto.arima* foram modelos não-sazonais; *** Modelo selecionado por análise subjetiva gráfica.

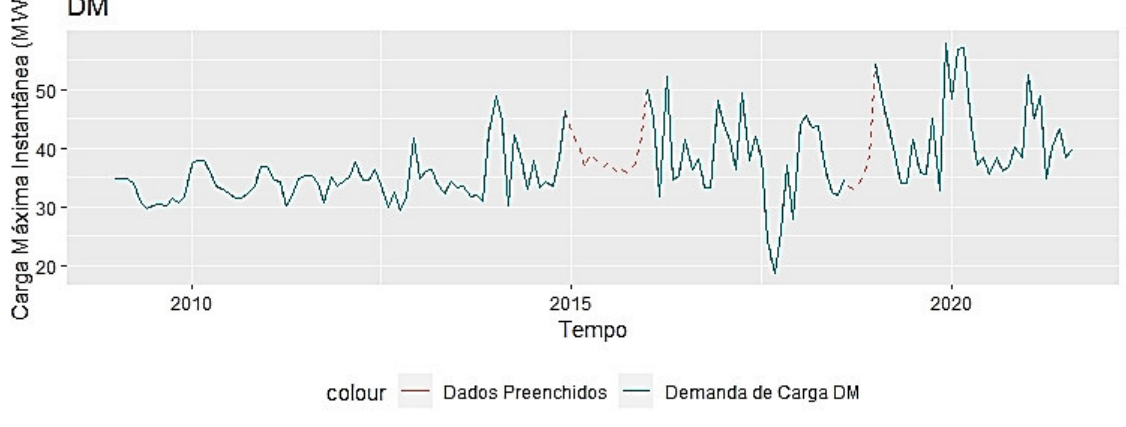
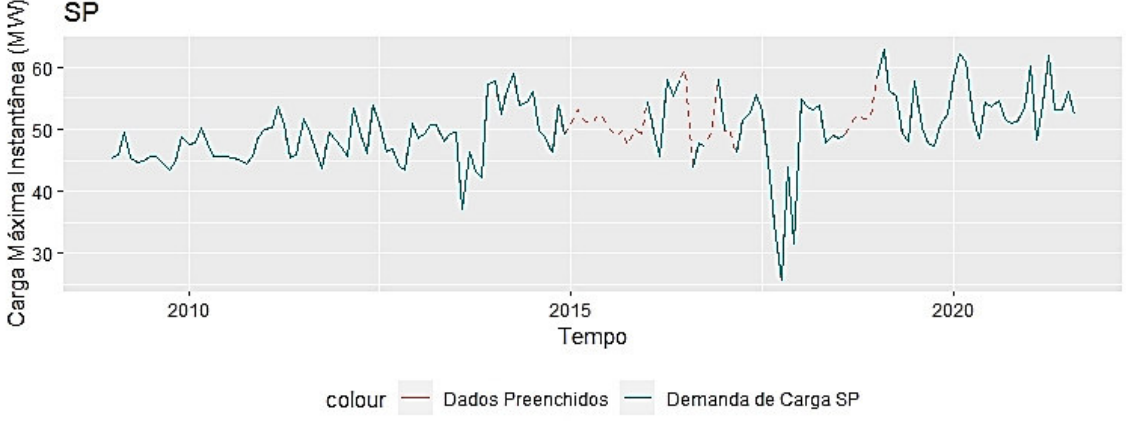
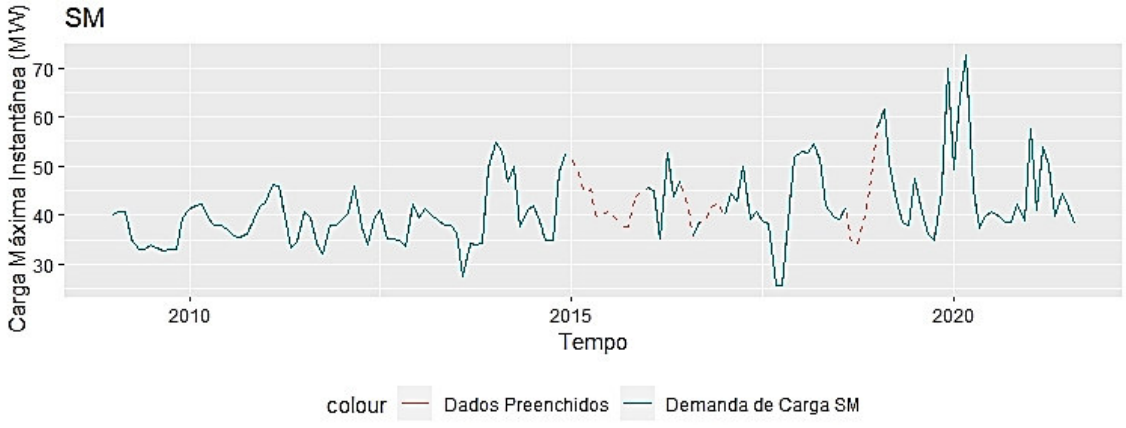
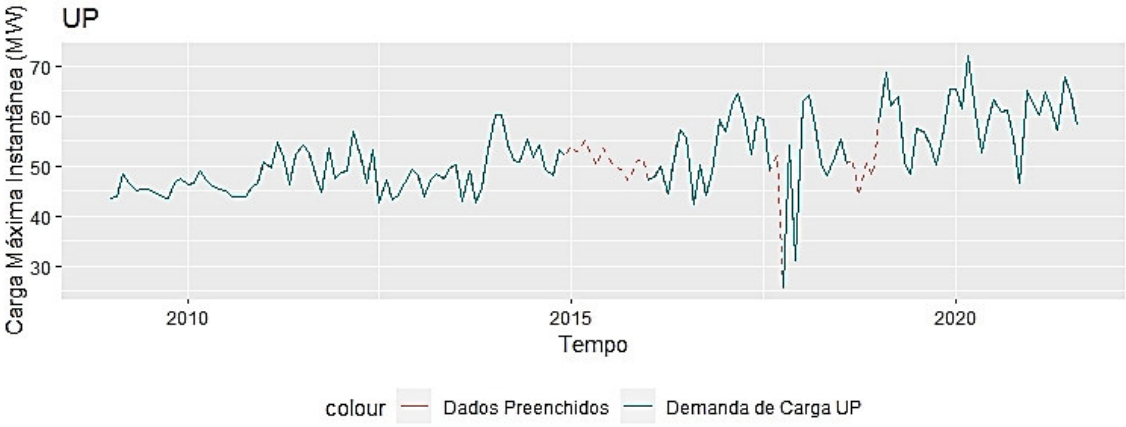
Fonte: Elaborado pela autora, 2023.

Na imputação dos dados, as previsões para o ano de 2015 do patamar UM e os quatro últimos meses de 2018 dos patamares SP e DM corresponderam à saída da função *auto.arima*. Por outro lado, nas previsões para 2015 dos patamares UP, DM e DP e para setembro de 2017 do patamar DP, a saída da função *auto.arima* foi um modelo ARIMA não-sazonal, mas como a análise exploratória dos dados indicou que todos os patamares apresentaram sazonalidade, a opção por um modelo sazonal, do tipo SARIMA, mostrou-se mais adequada, em vista de representar a influência dos meses na variabilidade dos dados.

Especificamente, para a previsão do ano de 2015 do patamar UL, além do modelo selecionado: $ARIMA(2,1,2)(1,1,1)_{[12]}$, houveram outros modelos candidatos aptos à imputação, porém todos os demais estimaram índices de carga elétrica muito baixos em relação aos limites inferior e superior do patamar UL. Então, com base na análise subjetiva gráfica, a escolha foi por um modelo com observações mais próximas aos limites desta série, no intuito de evitar grandes discrepâncias entre dados existentes e imputados. Nas demais séries, as saídas da função *auto.arima* foram modelos sazonais com resíduos autocorrelacionados, os quais foram substituídos por modelos de resíduos independentes e com o menor valor do BIC encontrado.

Na Figura 4.7 estão dispostas as séries de demanda de energia elétrica com os dados imputados, em pontilhado, baseados nos patamares de carga mencionados na Tabela 4.4.





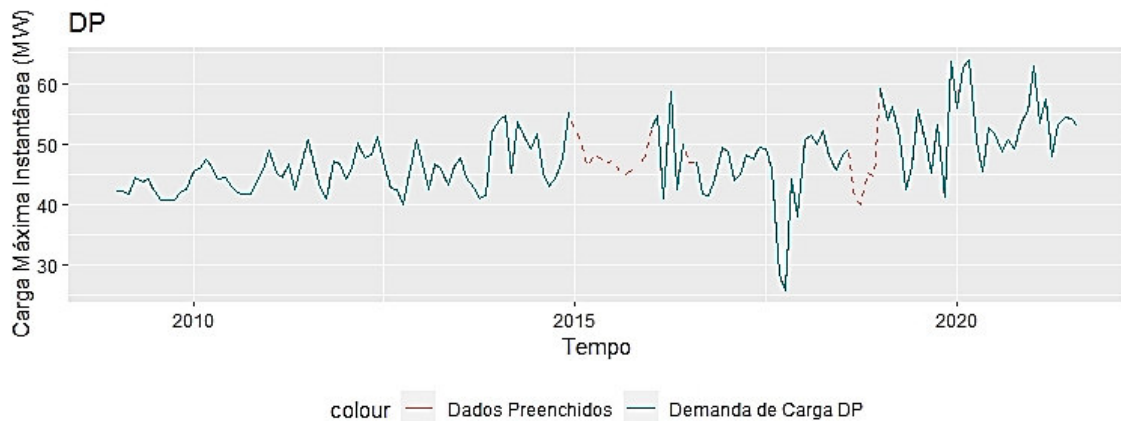


Figura 4.7 - Séries dos patamares de carga elétrica: jan./2009 - ago./2021

4.4. Estacionariedade das séries e ordem das variáveis exógenas

Os resultados dos testes de estacionariedade aplicados às séries dos patamares de carga elétrica e suas respectivas variáveis meteorológicas estão disponíveis no Apêndice 3 e na Tabela 4.7. A partir destes resultados foi possível observar que, todos os patamares de carga elétrica apresentaram um comportamento estacionário em nível, tal propriedade também foi observada na parte não-sazonal das séries meteorológicas consideradas. Por sua vez, o teste HEGY, em específico, evidenciou a presença de estacionariedade somente em primeira ou segunda diferença da parte sazonal de algumas séries meteorológicas, conforme as informações exibidas na Tabela 4.7.

Tabela 4.7. Resultados do teste HEGY para estacionariedade (D) das séries analisadas

Séries Temporais	Patamares de Carga Elétrica						
	UL	UM	UP	SM	SP	DM	DP
Demanda	0	0	0	0	0	0	0
Pc	-	-	2	0	-	2	-
Press	0	0	-	0	0	0	-
Rad	0	0	-	2	-	2	0
Temp	0	0	0	0	0	2	2
Orv	2	2	2	0	0	1	0
Tmáx	2	2	2	0	0	0	0
Tmín	0	0	0	0	0	2	2
UR	0	0	0	0	0	0	0
Raj	2	2	2	2	2	2	2
VV	2	2	2	2	2	2	2

Fonte: Elaborado pela autora, 2023.

Conforme mencionado anteriormente, as séries utilizadas nos modelos Box-Jenkins possuem como requisito apresentar o mesmo grau de diferenciação em seus dados, ou seja, as variáveis de entrada, bem como, as variáveis exógenas ao modelo necessitam apresentar o mesmo grau de diferenciação. Então, somente as séries das variáveis meteorológicas estacionárias foram testadas como variáveis exógenas aos modelos SARIMAX.

Desta forma, nos patamares dos dias úteis (UL, UM e UP), as variáveis de temperatura máxima e ponto de orvalho apresentaram estacionariedade somente em segunda diferença e foram eliminadas das etapas seguintes. Este mesmo critério excluiu a temperatura e a temperatura mínima nos patamares referentes aos domingos (DM e DP). As variáveis de vento: rajada e velocidade do vento, não apresentaram estacionariedade em nível nos sete patamares analisados e também foram descartadas de posteriores análises.

A partir da definição das variáveis exógenas aptas à modelagem, foi estabelecida a ordem de entrada ao argumento *Xreg*, considerando as variáveis do tipo temperatura como prioritárias, ou primárias, à entrada do modelo, em vista de maiores índices de correlação com a demanda de carga elétrica, além de outros critérios disponíveis na Tabela 4.8.

Tabela 4.8. Ordem e critérios de entrada ao *Xreg* da função *Arima*

Patamar de Carga	Ordem	Variável	Observações
UL	1	Temperatura	• Maior correlação com a demanda.
	X	Temperatura Mínima	• Segunda maior correlação com a demanda; • Utilizada alternativamente da temperatura.
	2	Radiação	• Segunda maior correlação com a demanda depois das variáveis de temperatura; • Baixa correlação com as mesmas em relação às demais variáveis.
	3	Umidade	• Terceira maior correlação com a demanda depois das variáveis de temperatura; • Baixa correlação com as mesmas.
	4	Pressão	• Maior correlação com a demanda depois das variáveis de temperatura; • Alta correlação com as mesmas.
UM	1	Temperatura Mínima	• Maior correlação com a demanda.
	X	Temperatura	• Segunda maior correlação com a demanda; • Utilizada alternativamente da temperatura mínima.

Patamar de Carga	Ordem	Variável	Observações
UM	2	Umidade	• Segunda maior correlação com a demanda depois das variáveis de temperatura; • Baixa correlação com as mesmas em relação às demais variáveis.
	3	Radiação	• Maior correlação com a demanda depois das variáveis de temperatura; • Alta correlação com as mesmas.
	4	Pressão	• Terceira maior correlação com a demanda depois das variáveis de temperatura; • Alta correlação com as mesmas.
UP	1	Temperatura Mínima	• Maior correlação com a demanda.
	X	Temperatura	• Segunda maior correlação com a demanda; • Utilizada alternativamente da temperatura mínima.
	2	Umidade	• Maior correlação com a demanda depois das variáveis de temperatura;
SM	1	Temperatura Máxima	• Maior correlação com a demanda.
	X	Temperatura	• Segunda maior correlação com a demanda; • Utilizada alternativamente da temperatura máxima.
	X	Ponto de Orvalho	• Terceira maior correlação com a demanda; • Utilizada alternativamente da temperatura máxima.
	X	Temperatura Mínima	• Quarta maior correlação com a demanda; • Utilizada alternativamente da temperatura máxima.
	2	Umidade	• Maior correlação com a demanda depois das variáveis de temperatura;
	3	Pressão	• Segunda maior correlação com a demanda depois das variáveis de temperatura;
	4	Precipitação	• Menor correlação com a demanda.
SP	1	Temperatura Máxima	• Maior correlação com a demanda.
	X	Temperatura	• Segunda maior correlação com a demanda; • Utilizada alternativamente da temperatura máxima.
	X	Ponto de Orvalho	• Terceira maior correlação com a demanda; • Utilizada alternativamente da temperatura máxima.

Patamar de Carga	Ordem	Variável	Observações
	X	Temperatura Mínima	• Quarta maior correlação com a demanda dentre as variáveis de temperatura; • Utilizada alternativamente da temperatura máxima.
	2	Umidade	• Maior correlação com a demanda depois das variáveis de temperatura;
	3	Pressão	• Menor correlação com a demanda;
DM	1	Temperatura Máxima	• Maior correlação com a demanda.
	2	Umidade	• Maior correlação com a demanda depois da variável de temperatura máxima.
	3	Pressão	• Menor correlação com a demanda;
DP	1	Temperatura Máxima	• Maior correlação com a demanda.
	X	Ponto de Orvalho	• Segunda maior correlação com a demanda; • Utilizada alternativamente da temperatura máxima.
	2	Umidade	• Maior correlação com a demanda depois das variáveis de temperatura;
	3	Radiação	• Menor correlação com a demanda;

Fonte: Elaborado pela autora, 2023.

4.5. Imputação dos dados meteorológicos

Após a organização dos dados meteorológicos, dentre as variáveis aptas à modelagem, todas apresentaram observações faltantes em maio e junho de 2016. Logo, conforme a metodologia, modelos SARIMA foram aplicados na imputação das séries meteorológicas, os quais podem ser conferidos na Tabela 4.9.

O modelo utilizado na imputação das observações meteorológicas de maio e junho de 2016 correspondeu à saída da função *auto.arima* no caso da temperatura nos patamares UL e SP e do ponto de orvalho no patamar DP. Por outro lado, modelos não-sazonais foram a saída desta função para a umidade nos patamares UL, SM, SP e DP, e para as variáveis radiação e precipitação nos patamares UL e SM, respectivamente.

Nos demais casos, assim como o procedimento aplicado nas séries de demanda de carga elétrica, para as variáveis meteorológicas as saídas da função *auto.arima* foram modelos sazonais com resíduos autocorrelacionados, os quais foram substituídos por modelos de resíduos independentes e com o menor valor do BIC encontrado.

Tabela 4.9. Modelos SARIMA aplicados nas séries meteorológicas

Patamar de Carga	Variável Meteorológica	Modelo SARIMA de imputação
UL	Temperatura	(0,0,0) (2,1,1) _[12] *
	Temperatura Mínima	(0,0,0) (0,2,2) _[12]
	Radiação**	(0,1,1) (1,0,2) _[12]
	Umidade**	(1,0,0) (1,0,0) _[12]
	Pressão	(0,0,0) (2,1,1) _[12]
UM	Temperatura	(0,0,2) (2,1,1) _[12]
	Temperatura Mínima	(0,0,0) (2,2,2) _[12]
	Umidade	(1,0,0) (1,1,1) _[12]
	Radiação	(0,0,0) (2,2,0) _[12]
	Pressão	(0,0,0) (1,1,2) _[12]
UP	Temperatura	(1,0,1) (2,1,1) _[12]
	Temperatura Mínima	(1,0,1) (1,1,1) _[12]
	Umidade	(2,0,2) (1,2,1) _[12]
SM	Temperatura	(2,0,0) (2,2,0) _[12]
	Temperatura Máxima	(2,0,0) (2,1,0) _[12]
	Temperatura Mínima	(0,0,0) (0,2,2) _[12]
	Ponto de Orvalho	(0,0,0) (2,1,1) _[12]
	Umidade**	(2,0,0) (1,1,2) _[12]
	Pressão	(1,0,0) (1,2,1) _[12]
	Precipitação**	(0,0,1) (1,1,2) _[12]
SP	Temperatura	(0,0,0) (0,1,1) _[12] *
	Temperatura Máxima	(0,0,2) (2,1,0) _[12]
	Temperatura Mínima	(2,0,1) (0,1,2) _[12]
	Ponto de Orvalho	(1,0,0) (2,1,1) _[12]
	Umidade**	(0,0,1) (1,0,1) _[12]
	Pressão	(0,0,0) (1,2,1) _[12]
DM	Temperatura Máxima	(0,0,0) (1,1,1) _[12]
	Umidade	(2,2,2) (1,2,0) _[12]
	Pressão	(0,0,0) (2,2,2) _[12]
DP	Temperatura Máxima	(0,0,0) (1,1,2) _[12]
	Ponto de Orvalho	(1,0,0) (2,1,0) _[12] *
	Umidade**	(2,2,2) (1,2,0) _[12]
	Radiação	(0,0,2) (2,1,2) _[12]

* Modelos de saída da *auto.arima*; ** Casos em que saída da *auto.arima* foram modelos não-sazonais.

Fonte: Elaborado pela autora, 2023.

4.6. Modelos

4.6.1. Patamar de Carga Leve dos Dias Úteis (UL)

Os gráficos da FAC e FACP das partes não-sazonal e sazonal do ajuste para carga leve dos dias úteis estão disponíveis na Figura 4.8. Na parte não-sazonal, identificou-se um declínio senoidal na FAC e uma queda brusca após a 2ª defasagem na FACP, configurando um processo autorregressivo

de 2ª ordem. Na parte sazonal, observou-se uma queda brusca após a 1ª defasagem na FAC e um leve declínio senoidal na FACP, sugerindo um processo de médias móveis de 1ª ordem. Assim, inicialmente, adotou-se o modelo candidato SARIMA(2,0,0)(0,0,1)_[12].

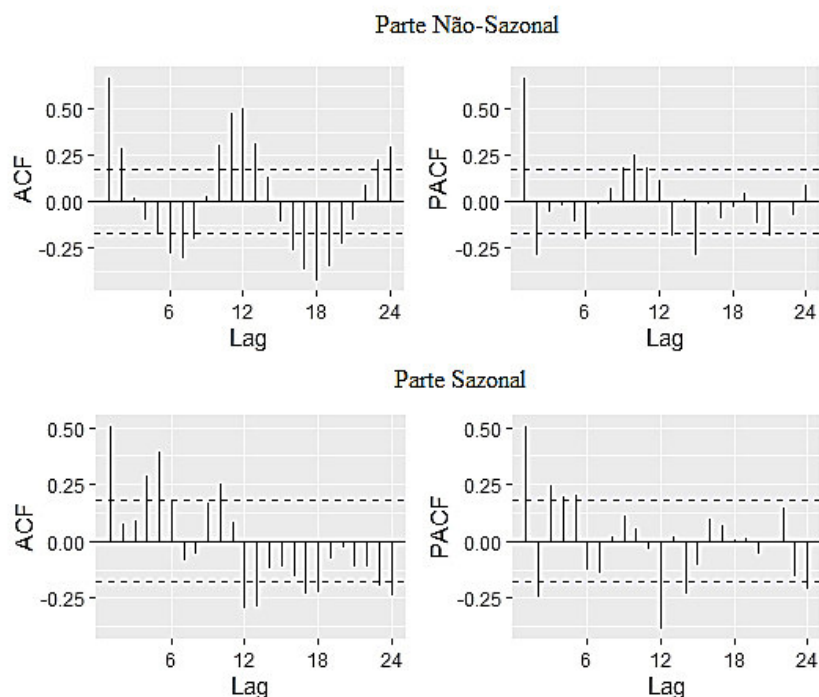


Figura 4.8 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar UL

Entre os trinta e seis modelos candidatos, com base nos valores de $p \pm 1$, $q \pm 1$ e $P \pm 1$, $Q \pm 1$, foram selecionados os cinco modelos, dispostos na Tabela 4.10 conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(3,0,0)(1,0,1)_[12].

Tabela 4.10. Patamar UL: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
(1,0,1) (1,0,1) _[12]	722,72	0,085	3,749
(2,0,0) (1,0,1) _[12]	723,68	0,073	3,748
(1,0,1) (1,0,2) _[12]	725,71	0,070	3,772
(1,0,1) (2,0,1) _[12]	726,27	0,076	3,770
(3,0,0) (1,0,1)_[12]	726,57	0,079	3,742

Fonte: Elaborado pela autora, 2023.

Por sua vez, o modelo SARIMA(3,0,0)(1,0,1)_[12] foi utilizado na sequência de modelagens com a inserção do argumento *Xreg* contendo as séries de variáveis meteorológicas do patamar correspondente, e aptas à aplicação na função *Arima* sob a ordem apresentada na Tabela 4.11.

Tabela 4.11. Patamar UL: sequência de Teste do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA		RMSE	MAPE	MAE
3,742	8,868	2,792	(3,0,0) (1,0,1)_[12]		3,742	8,868	2,792
			Xreg				
4,112	10,718	3,301	Temp Rad UR Press	Tmín Rad UR Press	4,057	10,206	3,164
4,010	10,187	3,166	Temp Rad UR	Tmín Rad UR	3,864	9,244	2,913
4,002	10,275	3,178	Temp Rad	Tmín Rad	3,800	9,066	2,846
3,844	9,551	2,981	Temp	Tmín	3,731	8,737	2,755

Fonte: Elaborado pela autora, 2023.

Durante a sequência de modelagens foi possível observar que somente a temperatura mínima melhorou o desempenho do modelo de entrada ao longo do período de validação, conforme os indicadores RMSE, MAPE e MAE. Entretanto, o acréscimo da temperatura mínima não foi capaz de inserir os dados reais dentro do intervalo de confiança do modelo, ao nível de 95% (Figura 4.9).

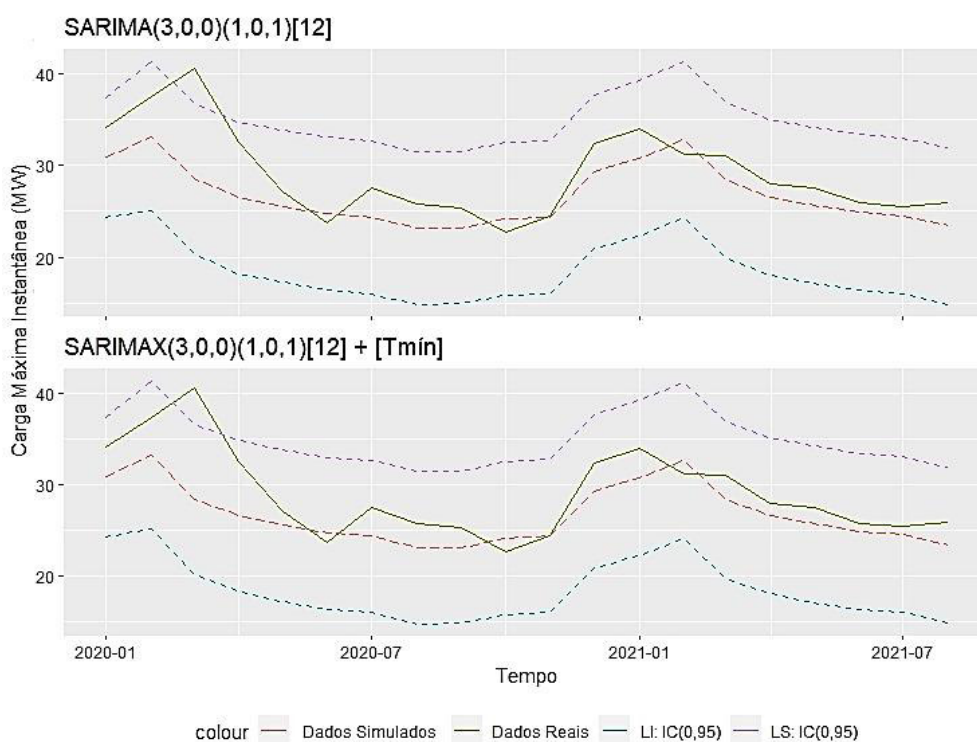


Figura 4.9 – Validação do patamar de carga leve dos dias úteis: dados reais e simulados

A propósito, na avaliação do intervalo de confiança do modelo SARIMA, independentemente das variáveis meteorológicas aplicadas ao *Xreg*, a observação de março de 2020 sempre esteve acima do limite superior do intervalo de confiança, ou seja, os dados reais da validação em nenhum momento estiveram totalmente inseridos dentro do IC, ao nível de 95%.

Em relação ao ajuste dos modelos SARIMA(3,0,0)(1,0,1)_[12] e SARIMAX(3,0,0)(1,0,1)_[12] + [Tmín], foi possível observar, por meio dos resultados expostos na Tabela 4.12, que o BIC teve um aumento razoável com a inserção dos regressores meteorológicos. Em relação aos resíduos, ao nível

de 5% de significância, os modelos estiveram de acordo com os critérios de independência e homoscedasticidade e, em nenhum caso foi identificada normalidade nas distribuições residuais.

Tabela 4.12. Patamar UL: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)

Modelos SARIMA/ SARIMAX	BIC	Resultados Residuais		
		Ljung-Box: valor-p (Lag 24)	Jarque-Bera: valor-p (df 2)	Teste ARCH: valor-p (Lag 24)
(3,0,0) (1,0,1) _[12]	726,57	0,079	< 2,2E-16	0,228
(3,0,0) (1,0,1) _[12] + [T _{mín}]	731,27	0,066	< 2,2E-16	0,212

Fonte: Elaborado pela autora, 2023.

4.6.2. Patamar de Carga Média dos Dias Úteis (UM)

Os gráficos da FAC e FACP das partes não-sazonal e sazonal do ajuste dos dados para carga média dos dias úteis estão disponíveis na Figura 4.10. Na parte não-sazonal, identificou-se um declínio senoidal na FAC e uma queda brusca após a 1ª defasagem na FACP, configurando um processo autorregressivo de 1ª ordem. Na parte sazonal, foi observada uma queda brusca após a 1ª defasagem na FAC, sugerindo um processo de médias móveis de 1ª ordem. Assim, inicialmente, adotou-se o modelo candidato SARIMA(1,0,0)(0,0,1)_[12].

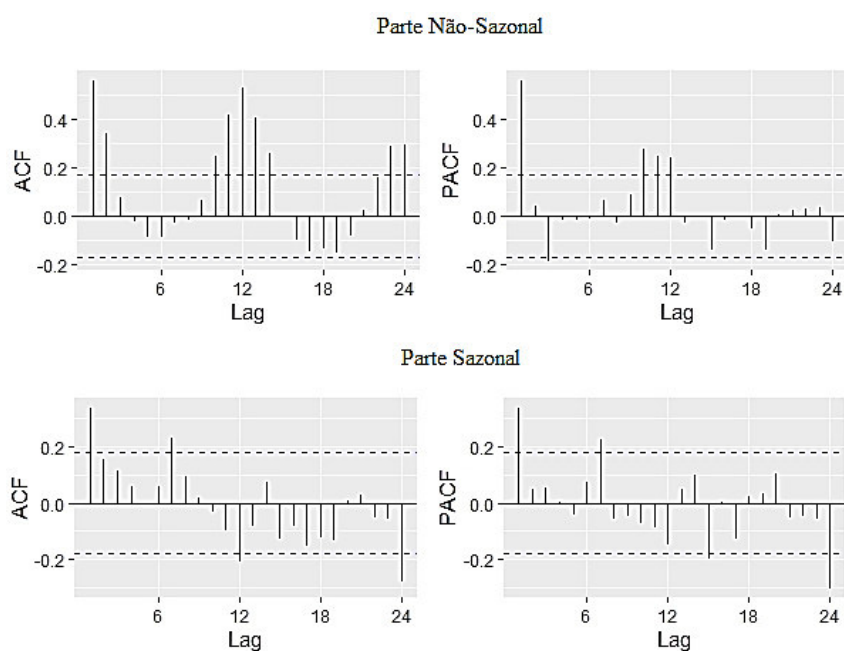


Figura 4.10 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar UM

Entre os trinta e seis modelos candidatos, com base nos valores de $p \pm 1$, $q \pm 1$ e $P \pm 1$, $Q \pm 1$, foram selecionados os cinco modelos, dispostos na Tabela 4.13, conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(2,0,0)(1,0,0)_[12].

Tabela 4.13. Patamar UM: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
(1,0,0) (1,0,0) _[12]	913,40	0,314	8,638
(1,0,0) (0,0,1) _[12]	916,34	0,121	9,820
(2,0,0) (1,0,0)_[12]	917,33	0,385	8,254
(1,0,1) (1,0,0)_[12]	917,47	0,368	8,255
(1,0,0) (1,0,1) _[12]	917,98	0,313	8,777

Fonte: Elaborado pela autora, 2023.

Porém, o modelo SARIMA(2,0,0)(1,0,0)_[12] não apresentou melhoras em seu desempenho para validação, durante a sequência de modelagens com a inserção do argumento *Xreg* contendo as variáveis meteorológicas do patamar correspondente, conforme mostra a Tabela 4.14.

Tabela 4.14. Patamar UM: sequência de Teste 1 do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA		RMSE	MAPE	MAE
8,254	11,255	6,712	(2,0,0) (1,0,0)_[12]		8,254	11,255	6,712
			Xreg				
8,711	12,618	7,427	Temp UR Rad Press	Tmín UR Rad Press	8,533	12,142	7,147
8,702	12,614	7,426	Temp UR Rad	Tmín UR Rad	8,444	12,021	7,084
8,657	12,508	7,377	Temp UR	Tmín UR	8,405	11,954	7,058
8,687	12,518	7,412	Temp	Tmín	8,363	11,768	7,038

Fonte: Elaborado pela autora, 2023.

Então, por este motivo, o modelo SARIMA(1,0,1)(1,0,0)_[12] foi escolhido, em vista do mesmo ter apresentado um desempenho mais próximo ao melhor modelo. Desta forma, a sequência de modelagens foi realizada e os resultados estão dispostos na Tabela 4.15.

Considerando a sequência de modelagens, foi possível notar que a utilização da temperatura mínima de forma individual e também em conjunto com a umidade relativa, melhorou o desempenho do modelo de entrada somente em relação RMSE, permanecendo os demais indicadores, MAPE e MAE, acima do índice de erros do modelo de entrada.

Tabela 4.15. Patamar UM: sequência de Teste 2 do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA (1,0,1) (1,0,0) _[12]		RMSE	MAPE	MAE
8,255	11,269	6,719			8,255	11,269	6,719
			Xreg				
8,593	12,420	7,286	Temp UR Rad Press	Tmín UR Rad Press	8,360	11,830	6,929
8,587	12,424	7,291	Temp UR Rad	Tmín UR Rad	8,293	11,763	6,904
8,550	12,339	7,258	Temp UR	Tmín UR	8,253	11,703	6,884
8,559	12,326	7,277	Temp	Tmín	8,209	11,527	6,875

Fonte: Elaborado pela autora, 2023.

Porém, dentre os modelos que melhoraram o RMSE, em nenhum houve inserção total dos dados reais no intervalo de confiança ao nível de 95% (Figura 4.11). O modelo acrescido do *Xreg* da temperatura mínima conseguiu inserir a observação de março de 2020 dentro do IC, mas ao mesmo tempo, retirou do IC a observação de junho de 2021. Este efeito também pode ser notado no modelo acrescido do conjunto da temperatura mínima com a umidade relativa, porém de forma mais acentuada, sendo assim, o modelo ficou “mais inserido” dentro do limite superior do IC na informação de março de 2020 e apresentou uma maior extrapolação deste limite em junho de 2021.

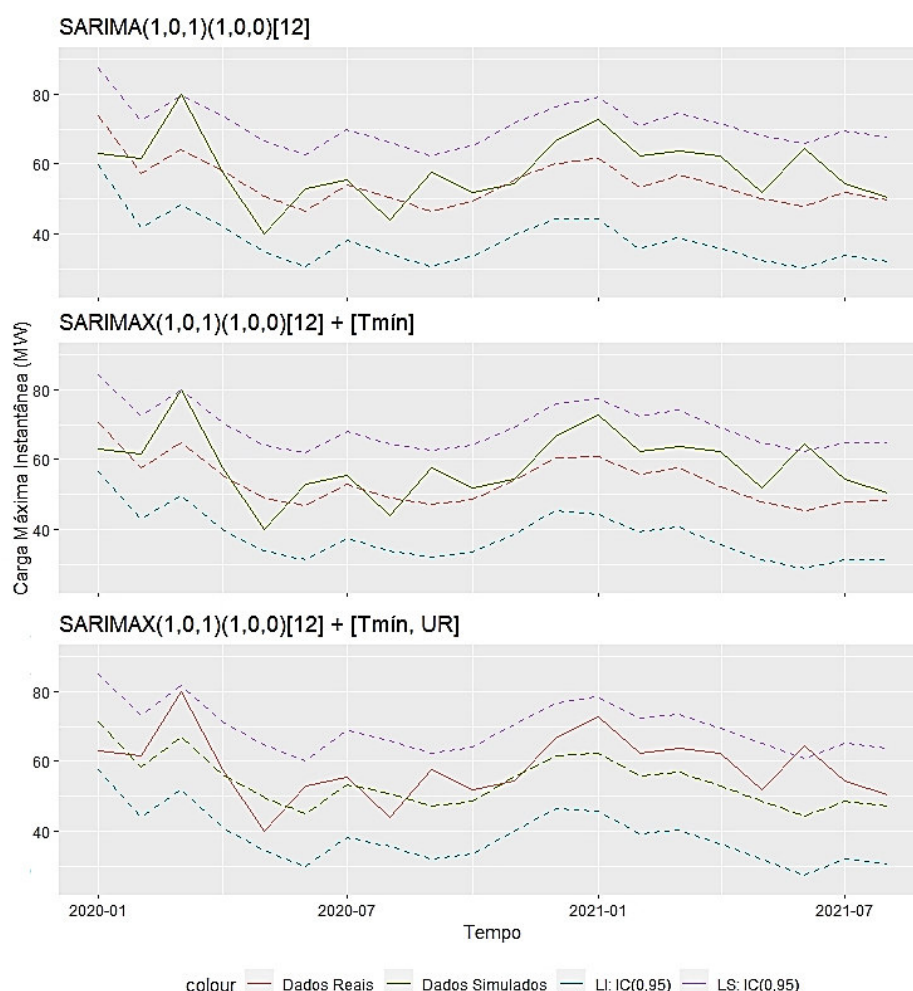


Figura 4.11 - Validação do patamar de carga média dos dias úteis: dados reais e simulados

Em relação ao ajuste do modelo SARIMA(1,0,1)(1,0,0)_[12] e dos modelos SARIMAX(1,0,1)(1,0,0)_[12] + [T_{mín}] e SARIMAX(1,0,1)(1,0,0)_[12] + [T_{mín}, UR], foi possível observar, por meio dos resultados expostos na Tabela 4.16, que o BIC teve um pequeno aumento no modelo com acréscimo dos regressores da temperatura mínima e um aumento razoável para o modelo com os fatores exógenos de temperatura mínima e umidade relativa.

Tabela 4.16. Patamar UM: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)

Modelos SARIMA/ SARIMAX	BIC	Resultados Residuais		
		Ljung-Box: valor-p (Lag 24)	Jarque-Bera: valor-p (df 2)	Teste ARCH: valor-p (Lag 24)
(1,0,1) (1,0,0) _[12]	917,47	0,368	2,33E-15	0,098
(1,0,1) (1,0,0) _[12] + [T _{mín}]	917,56	0,430	8,57E-09	0,005
(1,0,1) (1,0,0) _[12] + [T _{mín} , UR]	919,31	0,438	1,48E-06	0,003

Fonte: Elaborado pela autora, 2023.

A respeito dos resíduos, os modelos com a inserção das variáveis meteorológicas estiveram de acordo com os critérios de independência, porém, ao nível de 5% de significância, apresentaram evidências de heterocedasticidade, e em nenhum caso foi possível identificar normalidade nas distribuições residuais.

4.6.3. Patamar de Carga Pesada dos Dias Úteis (UP)

Os gráficos da FAC e FACP das partes não-sazonal e sazonal do ajuste para carga pesada dos dias úteis estão disponíveis na Figura 4.12. Na parte não-sazonal, identificou-se uma queda brusca após a 2ª defasagem no ACF e um declínio senoidal no PACF, configurando um processo de médias móveis de 2ª ordem. Na parte sazonal, observou-se um declínio senoidal na FAC e uma queda brusca após a 2ª defasagem na FACP, caracterizando um processo autorregressivo de 2ª ordem. Assim, inicialmente, adotou-se o modelo candidato SARIMA(0,0,2)(2,0,0)_[12].

Com base nos valores de $p \pm 2$, $q \pm 2$ e $P \pm 2$, $Q \pm 2$, aplicados em vista de maiores possibilidades de interpretações gráficas, foram selecionados os cinco modelos, dispostos na Tabela 4.17, conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(0,0,2)(1,0,1)_[12].

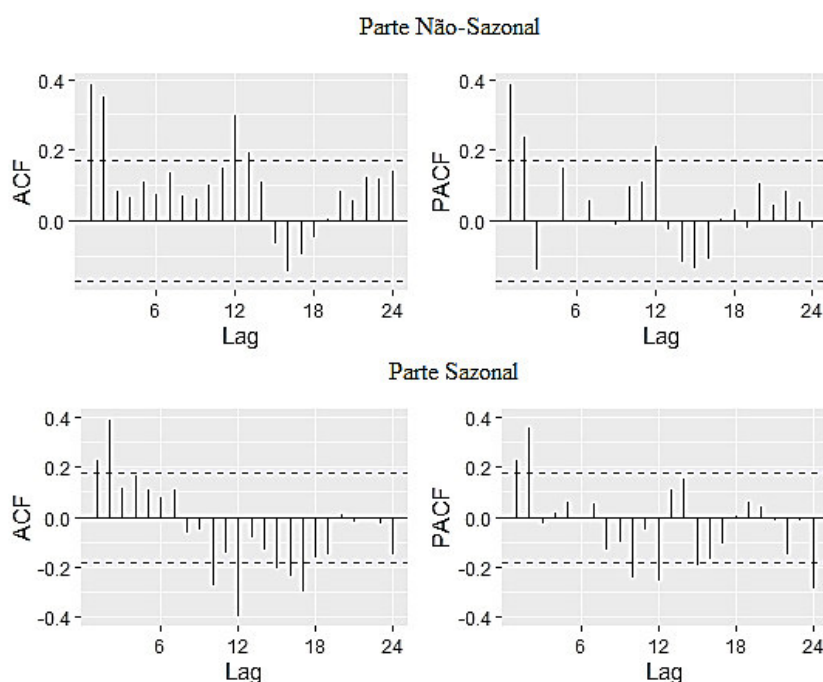


Figura 4.12 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar UP

Por sua vez, o modelo $SARIMA(0,0,2)(1,0,1)_{[12]}$ foi utilizado na sequência de modelagens com a inserção do argumento *Xreg* contendo as séries de variáveis meteorológicas do patamar correspondente e aptas à aplicação na função *Arima* sob a ordem apresentada na Tabela 4.18.

Tabela 4.17. Patamar UP: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
$(0,0,2) (1,0,0)_{[12]}$	834,15	0,933	9,568
$(0,0,2) (1,0,1)_{[12]}$	835,30	0,934	9,018
$(0,0,2) (0,0,1)_{[12]}$	835,69	0,903	10,076
$(0,0,2) (2,0,0)_{[12]}$	838,19	0,937	9,292
$(1,0,2) (1,0,0)_{[12]}$	838,34	0,943	9,316

Fonte: Elaborado pela autora, 2023.

A partir dos resultados da sequência de modelagens, dispostos na Tabela 4.18, foi possível observar que as variáveis exógenas não melhoraram o desempenho do modelo de entrada ao longo do período da validação, conforme os indicadores RMSE, MAPE e MAE. Além disso, os modelos gerados com os regressores das séries meteorológicas não conseguiram melhorar o modelo do ponto de vista do intervalo de confiança, ao nível de 95%.

Tabela 4.18. Patamar UP: sequência de Teste do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA		RMSE	MAPE	MAE
9,018	13,043	8,098	(0,0,2) (1,0,1)_[12]		9,018	13,043	8,098
			Xreg				
9,160	13,340	8,274	Temp UR	Tmín UR	9,141	13,298	8,247
9,179	13,353	8,283	Temp	Tmín	9,081	13,152	8,162

Fonte: Elaborado pela autora, 2023.

Desta forma, independentemente das variáveis aplicadas ao *Xreg*, as observações de março e dezembro de 2020 e também junho de 2021 permaneceram extrapolando o limite superior do IC, assim como no modelo de entrada SARIMA(0,0,2)(1,0,1)_[12], conforme mostra a Figura 4.13.

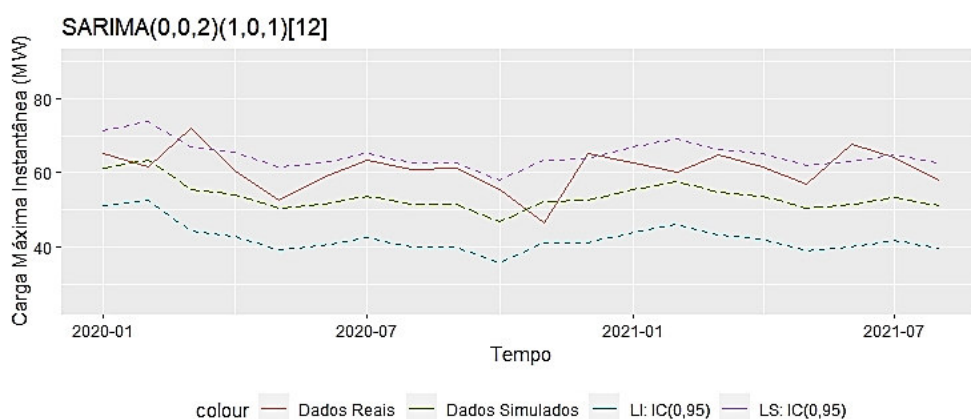


Figura 4.13 - Validação do patamar de carga pesada dos dias úteis: dados reais e simulados

Sobre as características do período de ajuste do modelo de entrada SARIMA(0,0,2)(1,0,1)_[12], o valor BIC, igual a 835,30, foi o segundo melhor encontrado na análise e, neste caso, se mostrou eficiente como critério de seleção do modelo de entrada. Em relação aos resíduos, ao nível de 5% de significância, mostraram-se independentes, com valor-p do Ljung-Box igual a 0,937, mas não apresentaram normalidade, visto que o teste Jarque-Bera resultou no valor-p de 7,2E-08, e nem homoscedasticidade, pelo teste ARCH, cujo valor-p correspondeu a 0,014.

4.6.4. Patamar de Carga Média dos Sábados (SM)

Os gráficos da FAC e FACP das partes não-sazonal e sazonal do ajuste para carga média dos sábados estão disponíveis na Figura 4.14. Na parte não-sazonal, identificou-se um declínio senoidal na FAC e uma queda brusca após a 1ª defasagem na FACP, configurando um processo autorregressivo de 1ª ordem. Na parte sazonal, observou-se uma queda brusca após a 1ª defasagem na FAC, sugerindo um processo de médias móveis de 1ª ordem. Assim, inicialmente, adotou-se o modelo candidato SARIMA(1,0,0)(0,0,1)_[12].

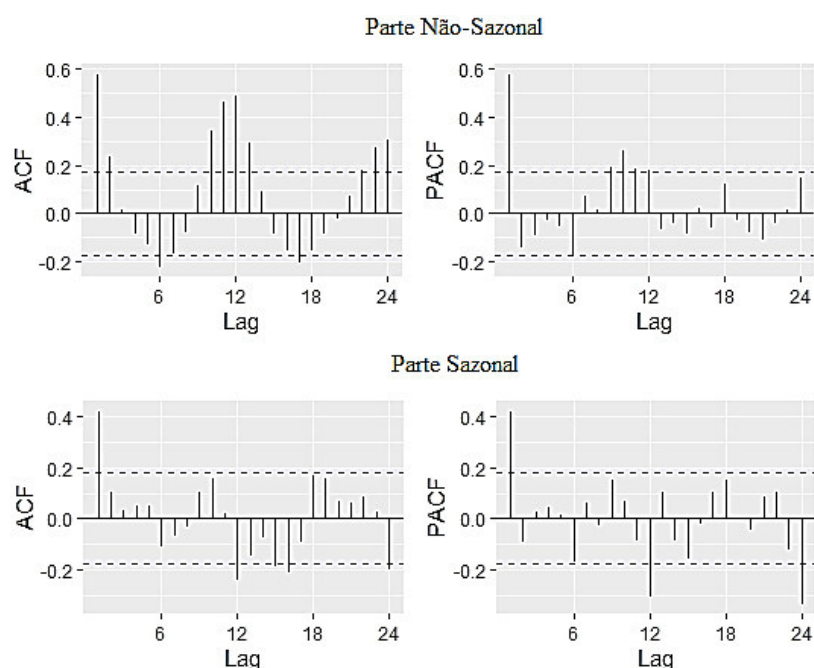


Figura 4.14 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar SM

Entre os trinta e seis modelos candidatos, com base nos valores de $p \pm 1$, $q \pm 1$ e $P \pm 1$, $Q \pm 1$, foram selecionados os cinco modelos, dispostos na Tabela 4.19, conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(1,0,0)(1,0,1)_[12].

Tabela 4.19. Patamar SM: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
(1,0,0) (1,0,1)_[12]	800,21	0,171	7,575
(1,0,0) (1,0,0) _[12]	802,26	0,079	8,150
(2,0,0) (1,0,1) _[12]	803,05	0,360	8,037
(1,0,1) (1,0,1) _[12]	803,36	0,298	7,904
(2,0,0) (1,0,0) _[12]	803,43	0,354	8,801

Fonte: Elaborado pela autora, 2023.

Por sua vez, o modelo SARIMA(1,0,0)(1,0,1)_[12] foi utilizado na sequência de modelagens com a inserção do argumento *Xreg* contendo as séries de variáveis meteorológicas do patamar correspondente e aptas à aplicação na função *Arima* sob a ordem apresentada na Tabela 4.20.

Durante a sequência de modelagens, em vista de uma maior correlação entre as variáveis exógenas: temperatura máxima e umidade relativa, a ordem de entrada no modelo deveria ser Tmax, Press, UR e Pc, e não a ordem que foi utilizada, correspondente a: Tmax, UR, Press e Pc; Porém, a ordem de inserção não apontou mudanças significativas nos resultados da validação.

Tabela 4.20. Patamar SM: sequência de Teste do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA		RMSE	MAPE	MAE
7,575	10,937	5,276	(1,0,0) (1,0,1)_[12]		7,575	10,937	5,276
			Xreg				
7,366	11,166	5,354	Tmáx UR Press Pc	Temp UR Press Pc	30,471	67,445	28,75
29,354	64,995	27,64	Tmáx UR Press	Temp UR Press	30,151	66,812	28,46
7,337	11,101	5,339	Tmáx UR	Temp UR	7,168	11,172	5,328
6,944	10,528	5,064	Tmáx	Temp	7,272	10,737	5,158
30,905	68,421	29,14	Orv UR Press Pc	Tmín UR Press Pc	31,230	69,106	29,38
30,386	67,344	28,68	Orv UR Press	Tmín UR Press	7,260	11,452	5,450
7,259	11,269	5,380	Orv UR	Tmín UR	7,312	11,509	5,487
7,573	10,629	5,152	Orv	Tmín	7,569	10,873	5,254

Fonte: Elaborado pela autora, 2023.

Os elevados índices dos indicadores, dispostos na Tabela 4.20, são resultados da produção de *NA* 's durante o seguinte processo no *software R*: *sqrt(diag(x\$var.coef))*, que gerou erros na previsão das observações para o período da validação.

Por outro lado, nos casos de sucesso, foi possível observar que o indicador RMSE apontou melhora do desempenho do modelo de entrada, ao longo do período da validação, utilizando quaisquer conjuntos de variáveis meteorológicas, desta forma, segundo o RMSE, a inserção das informações meteorológicas, aptas à modelagem ARIMA, melhoraram o desempenho do modelo SARIMA(1,0,0)(1,0,1)_[12] para validação. Todavia, somente a temperatura ou suas variações, adicionadas de forma individual ao *Xreg*, diminuíram todos os índices de indicadores de erros: RMSE, MAPE e MAE, e a temperatura máxima foi a variável que obteve os melhores índices.

Porém, nenhum modelo gerado com a adição dos regressores das séries meteorológicas conseguiu inserir os dados reais no intervalo de confiança de 95%. Sendo assim, independentemente das variáveis aplicadas ao *Xreg*, observações de janeiro, março e dezembro de 2020 permaneceram extrapolando os limites do IC.

Na Figura 4.15, acima estão expostos os dados reais e gerados pelo modelo SARIMA(1,0,0)(1,0,1)_[12], juntamente com o IC de 95%, e abaixo está o modelo em conjunto com os regressores da temperatura máxima, a qual obteve o melhor desempenho segundo os indicadores utilizados.

Em relação ao ajuste dos modelos SARIMA(1,0,0)(1,0,1)_[12] e SARIMAX(1,0,0)(1,0,1)_[12] + [Tmáx] foi possível observar, por meio dos resultados expostos na Tabela 4.21, que o BIC diminuiu com acréscimo dos regressores da temperatura máxima. Na análise residual, ambos os modelos estiveram de acordo com os critérios de independência e homoscedasticidade, porém, novamente, não houve identificação de normalidade nas distribuições.

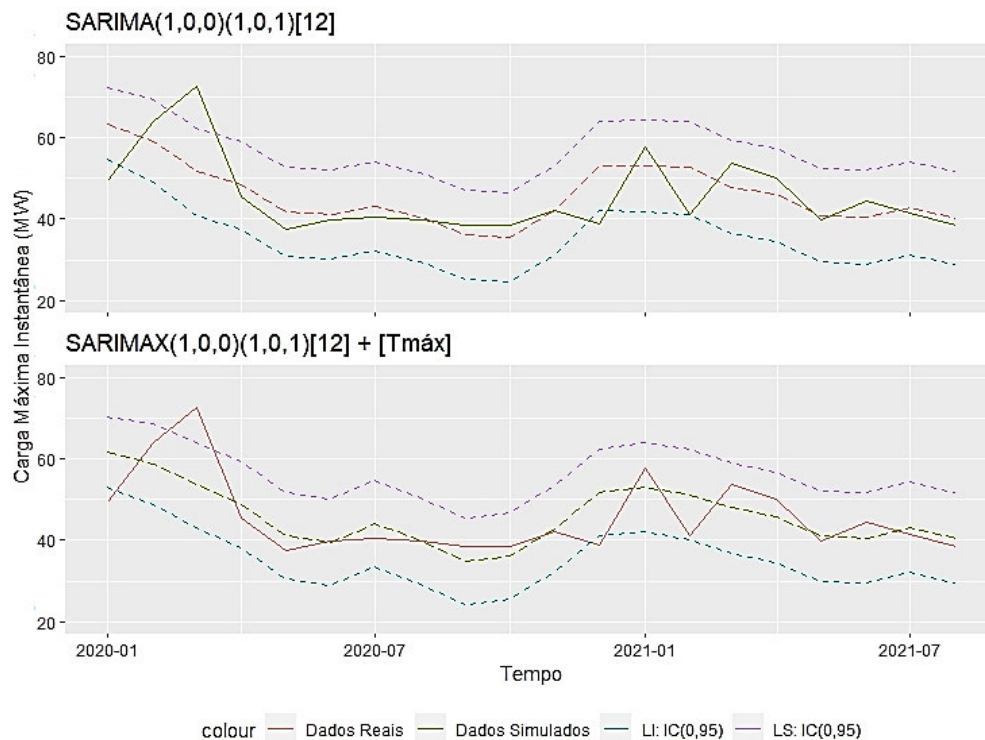


Figura 4.15 - Validação do patamar de carga média dos sábados: dados reais e simulados

Tabela 4.21. Patamar SM: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)

Modelos SARIMA/ SARIMAX	BIC	Resultados Residuais		
		Ljung-Box: valor-p (Lag 24)	Jarque-Bera: valor-p (df 2)	Teste ARCH: valor-p (Lag 24)
(1,0,0) (1,0,1) _[12]	800,21	0,171	2,2E-16	0,062
(1,0,0) (1,0,1) _[12] + [Tmáx]	797,2	0,095	2,2E-16	0,125

Fonte: Elaborado pela autora, 2023.

4.6.5. Patamar de Carga Pesada dos Sábados (SP)

Os gráficos da FAC e FACP das partes não-sazonal e sazonal do ajuste para carga pesada dos sábados estão disponíveis na Figura 4.16. Na parte não-sazonal, identificou-se um declínio senoidal na FAC e uma queda brusca após a 1ª defasagem na FACP, configurando um processo autorregressivo de 1ª ordem. Na parte sazonal, observou-se um declínio senoidal na FAC e uma queda brusca após a 2ª defasagem na FACP, sugerindo um processo autorregressivo de 2ª ordem. Assim inicialmente, adotou-se o modelo candidato SARIMA(1,0,0)(2,0,0)_[12].

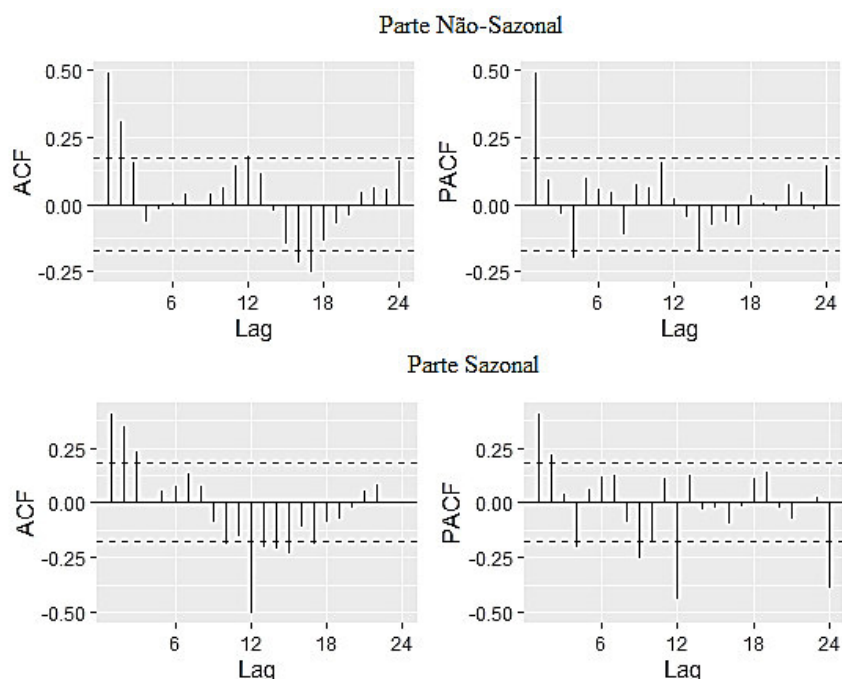


Figura 4.16 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar SP

Entre os trinta e seis modelos candidatos, com base nos valores de $p \pm 1$, $q \pm 1$ e $P \pm 1$, $Q \pm 1$, foram selecionados os cinco modelos, dispostos na Tabela 4.22, conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(1,0,0)(1,0,1)_[12].

Tabela 4.22. Patamar SP: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
(1,0,0) (3,0,0) _[12]	782,00	0,235	6,940
(1,0,1) (3,0,0) _[12]	786,07	0,219	6,850
(1,0,0) (3,0,1) _[12]	786,80	0,233	6,943
(1,0,0) (1,0,1)_[12]	786,83	0,509	5,463
(1,0,0) (1,0,0) _[12]	788,73	0,477	6,029

Fonte: Elaborado pela autora, 2023.

Por sua vez, o modelo SARIMA(1,0,0)(1,0,1)_[12] foi utilizado na sequência de modelagens com a inserção do argumento *Xreg* contendo as séries de variáveis meteorológicas do patamar correspondente e aptas à aplicação na função *Arima* sob a ordem apresentada na Tabela 4.23.

Tabela 4.23. Patamar SP: sequência de Teste do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA		RMSE	MAPE	MAE
5,463	8,351	4,698	(1,0,0) (1,0,1)_[12]		5,463	8,351	4,698
			Xreg				
5,332	8,378	4,688	Tmáx UR Press	Temp UR Press	5,342	8,406	4,703
5,347	8,397	4,701	Tmáx UR	Temp UR	5,367	8,460	4,734
5,402	8,329	4,680	Tmáx	Temp	5,442	8,479	4,755
5,353	8,424	4,712	Orv UR Press	Tmín UR Press	5,326	8,362	4,678
5,377	8,464	4,736	Orv UR	Tmín UR	5,370	8,429	4,720
5,459	8,347	4,696	Orv	Tmín	5,482	8,426	4,737

Fonte: Elaborado pela autora, 2023.

A partir dos resultados da sequência de modelagens, dispostos na Tabela 4.23, foi possível observar que, exceto os regressores da temperatura mínima, [Tmín] todas as variáveis meteorológicas e conjuntos de variáveis que estiveram aptas a ser adicionadas ao *Xreg* melhoraram o desempenho do modelo de entrada SARIMA(1,0,0)(1,0,1)_[12], em razão da diminuição de, ao menos, um dos indicadores de erros para validação. Entretanto, somente os regressores da temperatura máxima [Tmáx] e da temperatura do ponto de orvalho, [Orv], adicionadas de forma individual ao *Xreg*, melhoraram todos os indicadores: RMSE, MAPE e MAE, e os menores índices pertenceram à temperatura máxima [Tmáx], porém, dentre todas as modelagens, o menor RMSE foi pertencente ao conjunto da temperatura mínima com a umidade relativa e a pressão atmosférica, [Tmín, UR, Press].

Mas nenhum modelo gerado com regressores de séries meteorológicas conseguiu inserir os dados reais no intervalo de confiança IC, ao nível 95%. Especificamente, a observação de abril de 2021 manteve-se extrapolando o limite superior do IC ao longo da sequência de modelagens.

Na Figura 4.17 acima estão expostos os dados reais da validação e também os dados gerados pelo modelo SARIMA(1,0,0)(1,0,1)_[12], centralmente estão os dados deste modelo em conjunto com os regressores da temperatura máxima, o qual obteve o melhor desempenho segundo os três indicadores utilizados: RMSE, MAPE e MAE, e, por fim, abaixo está o modelo associado ao conjunto de regressores da temperatura mínima, umidade relativa e pressão atmosférica, [Tmín, UR, Press], o qual apresentou o menor índice RMSE. Nota-se que, além da informação de abril de 2021, no modelo sem adição de regressores, as observações de fevereiro e março de 2021 extrapolaram o limite superior do IC, mas com a inserção da temperatura máxima, somente março de 2020 superou o limite do IC e, no caso do modelo com os regressores da temperatura mínima, umidade relativa e pressão atmosférica, todas as observações estiveram dentro do IC, com a contínua exceção da observação de abril de 2021.

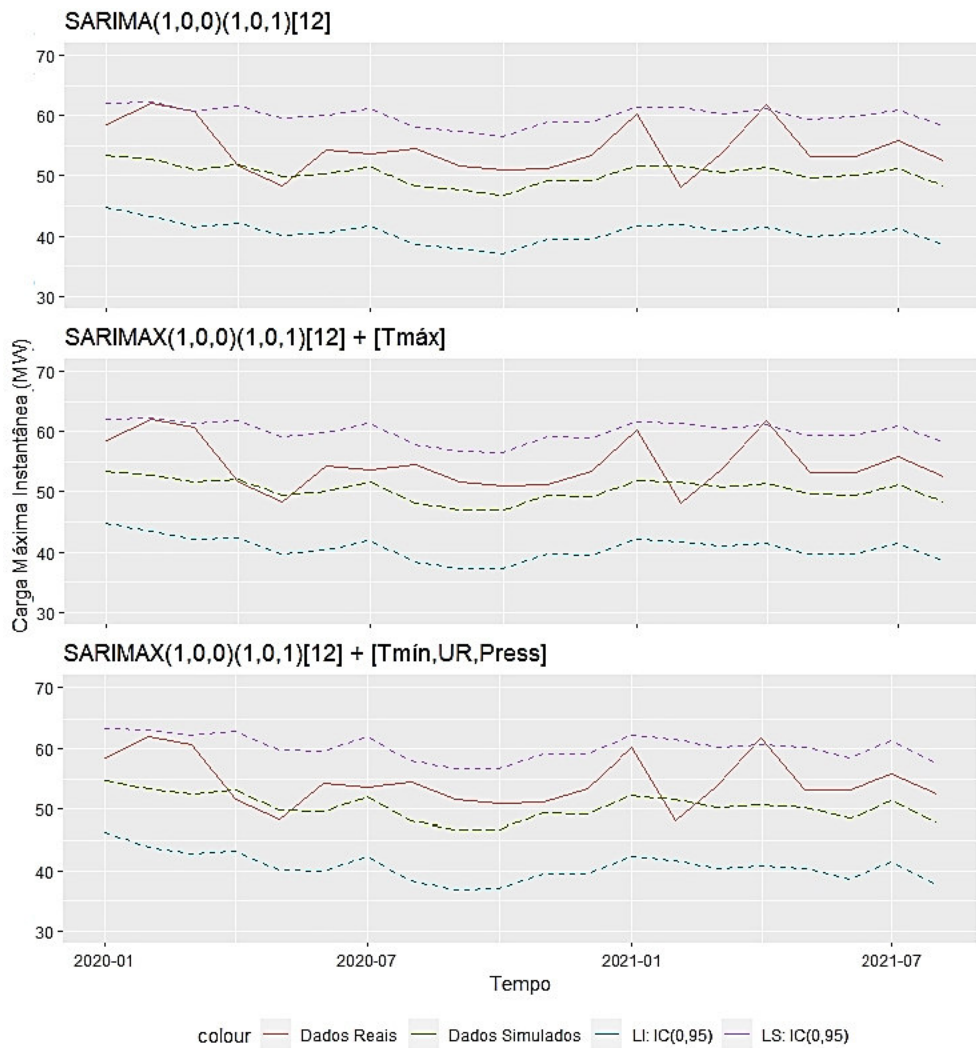


Figura 4.17 - Validação do patamar de carga pesada dos sábados: dados reais e simulados

Em relação ao ajuste dos modelos $SARIMA(1,0,0)(1,0,1)_{[12]}$ e dos modelos $SARIMAX(1,0,0)(1,0,1)_{[12]} + [T_{máx}]$ e $SARIMAX(1,0,0)(1,0,1)_{[12]} + [T_{mín}, UR, Press]$ foi possível observar, por meio dos resultados expostos na Tabela 4.24, que o BIC aumentou com acréscimo dos regressores da temperatura máxima e aumentou consideravelmente com o conjunto de regressores da temperatura mínima, umidade relativa e pressão atmosférica. Na análise residual, ao nível de 5% de significância, os modelos com a inserção das variáveis meteorológicas estiveram de acordo com os critérios de independência, mas somente o modelo com o conjunto de regressores da temperatura mínima, umidade e pressão, apresentou resíduos homoscedásticos, e em nenhum caso foi identificada normalidade nas distribuições residuais.

Tabela 4.24. Patamar SP: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)

Modelos SARIMA/ SARIMAX	BIC	Resultados Residuais		
		Ljung-Box: valor-p (Lag 24)	Jarque-Bera: valor-p (df 2)	Teste ARCH: valor-p (Lag 24)
(1,0,0) (1,0,1) _[12]	786,83	0,508	8,88E-16	0,041
(1,0,0) (1,0,1) _[12] + [Tmáx]	791,06	0,466	2,2E-16	0,040
(1,0,0) (1,0,1) _[12] + [Tmín, UR, Press]	798,56	0,360	2,2E-16	0,111

Fonte: Elaborado pela autora, 2023.

4.6.6. Patamar de Carga Média dos Domingos (DM)

Os gráficos da FAC e FACP referente à parte não-sazonal e sazonal do período de ajuste dos dados do patamar de carga média para os domingos estão disponíveis na Figura 4.18. A partir destes gráficos não foi possível identificar processos autoregressivos nem de médias móveis para as partes não-sazonal e sazonal do modelo, desta forma não foi definido nenhum modelo inicial e os modelos candidatos foram definidos com base nos valores $0 \leq p \leq 3$, $0 \leq q \leq 3$ para parte não-sazonal e $0 \leq P \leq 3$, $0 \leq Q \leq 3$ para parte sazonal.

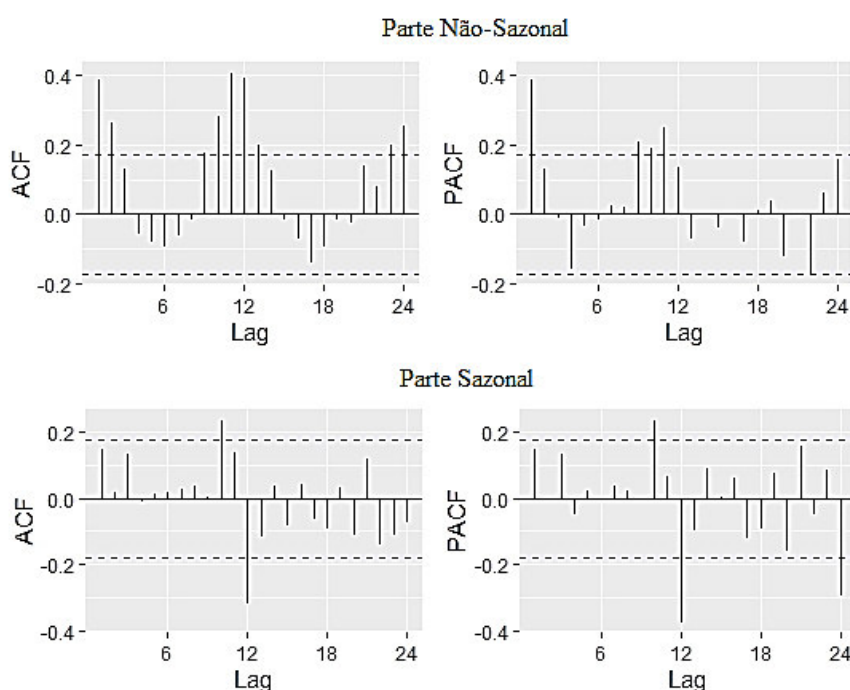


Figura 4.18 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar DM

A partir da determinação dos modelos candidatos, foram selecionados cinco modelos, dispostos na Tabela 4.25, conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(1,0,1)(1,0,1)_[12].

Tabela 4.25. Patamar DM: Modelos SARIMA Candidatos e Selecionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
(1,0,0) (1,0,0) _[12]	818,84	0,113	6,876
(1,0,0) (1,0,1) _[12]	818,98	0,140	6,409
(1,0,1) (1,0,1)_[12]	820,67	0,220	5,779
(2,0,0) (1,0,1) _[12]	821,03	0,348	5,988
(1,0,0) (2,0,0) _[12]	821,47	0,173	6,377

Fonte: Elaborado pela autora, 2023.

Por sua vez, o modelo SARIMA(1,0,1)(1,0,1)_[12] foi utilizado na sequência de modelagens com a inserção do argumento *Xreg* contendo as séries de variáveis meteorológicas do patamar correspondente e aptas à aplicação na função *Arima* sob a ordem apresentada na Tabela 4.26.

Tabela 4.26. Patamar DM: sequência de Teste do Modelo com os regressores meteorológicos

Modelo SARIMA	RMSE	MAPE	MAE
(1,0,1) (1,0,1)_[12]	5,779	9,789	4,423
Xreg			
Tmáx UR Press	5,347	9,712	4,304
Tmáx UR	5,428	9,891	4,385
Tmáx	5,539	9,946	4,445

Fonte: Elaborado pela autora, 2023.

A partir dos resultados da sequência de modelagens, dispostos na Tabela 4.26, foi possível observar que somente as variáveis meteorológicas: temperatura máxima, umidade relativa e pressão atmosférica adicionadas de forma conjunta ao *Xreg* melhoraram o desempenho do modelo de entrada segundo os três indicadores de erros utilizados: RMSE, MAPE e MAE, e também obtiveram os menores índices de erros. Assim, o modelo SARIMAX(1,0,1)(1,0,1)_[12] + [Tmáx, UR, Press] conseguiu maior representatividade dos dados reais da validação em relação às demais configurações.

Em relação ao intervalo de confiança, assim como nos patamares anteriores, os modelos gerados com os regressores das séries meteorológicas não conseguiram inserir os dados reais no IC ao nível de 95%. De forma específica, a informação de março de 2020 extrapolou continuamente o limite superior do IC ao longo da sequência de modelagens.

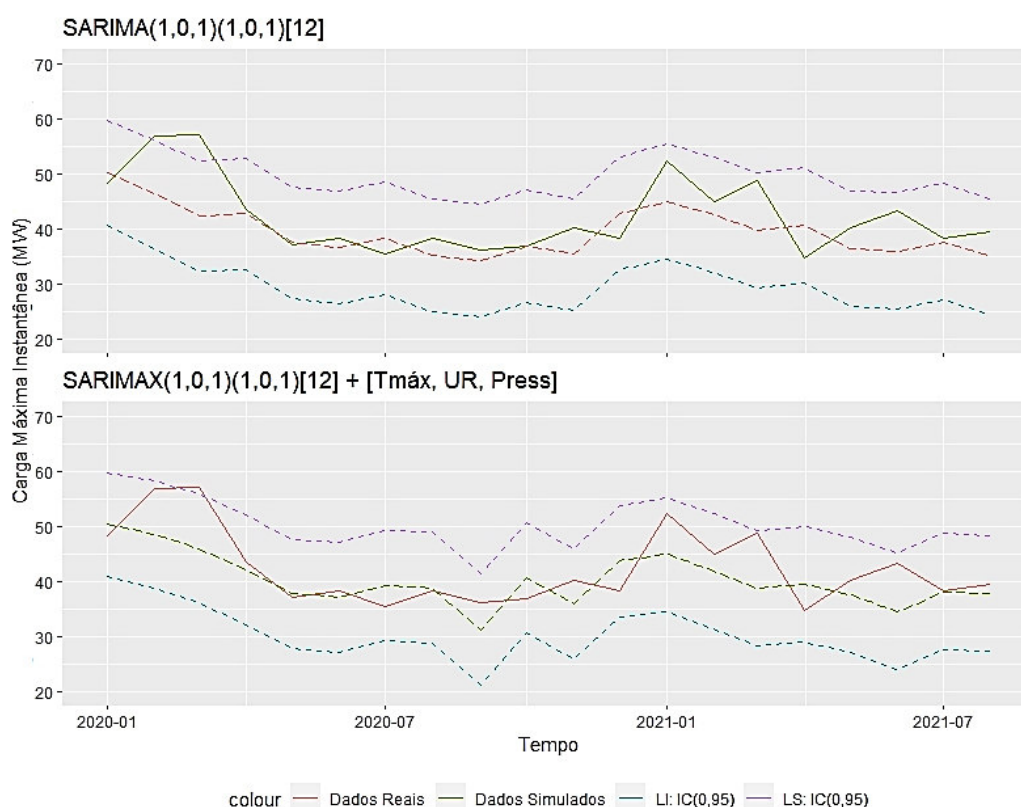


Figura 4.19 - Validação do patamar de carga média dos domingos: dados reais e simulados

Na Figura 4.19, à esquerda estão expostos os dados reais e gerados pelo modelo SARIMA(1,0,1)(1,0,1)_[12], juntamente com o IC de 95%, e abaixo está o modelo com os regressores da temperatura máxima, umidade relativa e pressão atmosférica, que obteve o melhor desempenho segundo os indicadores utilizados: RMSE, MAPE e MAE.

Tabela 4.27. Patamar DM: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)

Modelos SARIMA/ SARIMAX	BIC	Resultados Residuais		
		Ljung-Box: valor-p (Lag 24)	Jarque-Bera: valor-p (df 2)	Teste ARCH: valor-p (Lag 24)
(1,0,1) (1,0,1) _[12]	820,67	0,220	5,07E-11	0,008
(1,0,1) (1,0,1) _[12] + [Tmáx, UR, Press]	824,81	0,314	2,33E-14	0,002

Fonte: Elaborado pela autora, 2023.

Em relação ao ajuste dos modelos SARIMA(1,0,1)(1,0,1)_[12] e SARIMAX(1,0,1)(1,0,1)_[12] + [T_{máx}, UR, Press] foi possível observar, por meio dos resultados expostos na Tabela 4.27, que o BIC aumentou com a inserção do conjunto de regressores das variáveis meteorológicas. Na análise residual, ambos os modelos estiveram de acordo com os critérios de independência, mas não apresentaram homoscedasticidade e nem normalidade nas distribuições de seus resíduos.

4.6.7. Patamar de Carga Pesada dos Domingos (DP)

Os gráficos da FAC e FACP referente à parte não-sazonal e sazonal do período de ajuste dos dados do patamar de carga pesada para os domingos estão disponíveis na Figura 4.20. A partir destes gráficos não foi possível identificar processos autoregressivos nem de médias móveis para as partes não-sazonal e sazonal do modelo, desta forma não foi definido nenhum modelo inicial e os modelos candidatos foram definidos com base nos valores $0 \leq p \leq 3$, $0 \leq q \leq 3$ para parte não-sazonal e $0 \leq P \leq 3$, $0 \leq Q \leq 3$ para parte sazonal.

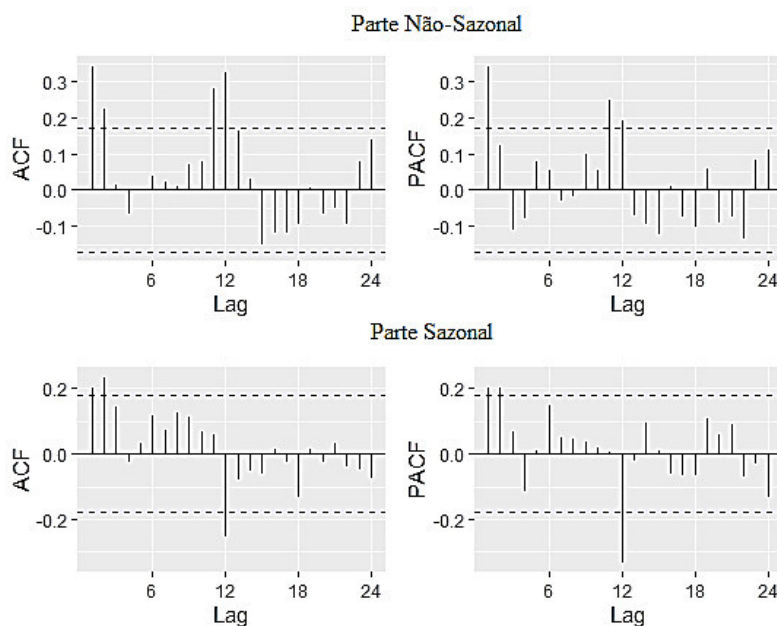


Figura 4.20 - Gráficos da FAC e FACP da parte não-sazonal e sazonal do Ajuste do Patamar DP

A partir da determinação dos modelos candidatos, foram selecionados cinco modelos, dispostos na Tabela 4.28, conforme os critérios de menor BIC e independência identificada no teste Ljung-Box. A partir destes cinco modelos candidatos, foi selecionado o modelo de melhor desempenho no período da validação, segundo o RMSE, o qual correspondeu ao SARIMA(2,0,0)(1,0,0)_[12].

Tabela 4.28. Patamar DP: Modelos SARIMA Candidatos e Seleccionados ($\alpha = 0,05$)

Modelos SARIMA	BIC	Ljung-Box: valor-p (Lag 24)	RMSE (Validação)
(1,0,0) (1,0,0) _[12]	788,05	0,727	6,861
(0,0,2) (1,0,0) _[12]	788,25	0,914	6,811
(2,0,0) (1,0,0)_[12]	788,55	0,897	6,329
(2,0,0) (1,0,1) _[12]	789,42	0,843	6,640
(1,0,0) (0,0,1) _[12]	789,93	0,656	7,441

Fonte: Elaborado pela autora, 2023.

Por sua vez, o modelo SARIMA(2,0,0)(1,0,0)_[12] foi utilizado na sequência de modelagens com a inserção do argumento *Xreg* contendo as séries de variáveis meteorológicas do patamar correspondente e aptas à aplicação na função *Arima* sob a ordem apresentada na Tabela 4.29.

Tabela 4.29. Patamar DP: sequência de Teste do Modelo com os regressores meteorológicos

RMSE	MAPE	MAE	Modelo SARIMA		RMSE	MAPE	MAE
6,329	8,734	4,938	(2,0,0) (1,0,0)_[12]		6,329	8,734	4,938
			Xreg				
6,002	8,655	4,831	Tmáx UR Rad	Orv UR Rad	6,159	8,658	4,863
6,137	8,759	4,896	Tmáx UR	Orv UR	6,279	8,772	4,932
6,201	8,817	4,944	Tmáx	Orv	6,303	8,682	4,910

Fonte: Elaborado pela autora, 2023.

A partir dos resultados da sequência de modelagens, dispostos na Tabela 4.29, foi possível observar que todas as variáveis meteorológicas e conjunto de variáveis, que estiveram aptas a ser adicionadas ao *Xreg* melhoraram o desempenho do modelo SARIMA(2,0,0)(1,0,0)_[12], em razão da diminuição de, ao menos, um dos indicadores de erros para validação. Por sua vez, o conjunto de regressores da temperatura máxima, umidade e radiação, [Tmáx, UR, Rad], melhorou o desempenho do modelo segundo os três indicadores de erros utilizados: RMSE, MAPE e MAE, e também obteve os menores índices de erros.

Na avaliação do intervalo de confiança, assim como nos patamares anteriores, os modelos gerados com os regressores das séries meteorológicas não conseguiram inserir os dados reais no IC ao nível de 95%. Especificamente, as informações de março de 2020 e janeiro de 2021 estiveram acima do limite superior do IC ao longo da sequência de modelagens.

Na Figura 4.21, acima estão expostos os dados reais e gerados pelo modelo SARIMA(2,0,0)(1,0,0)_[12], juntamente com o IC de 95%, e abaixo está o modelo com os regressores da temperatura máxima, umidade relativa e radiação, que obteve o melhor desempenho segundo os indicadores utilizados: RMSE, MAPE e MAE.

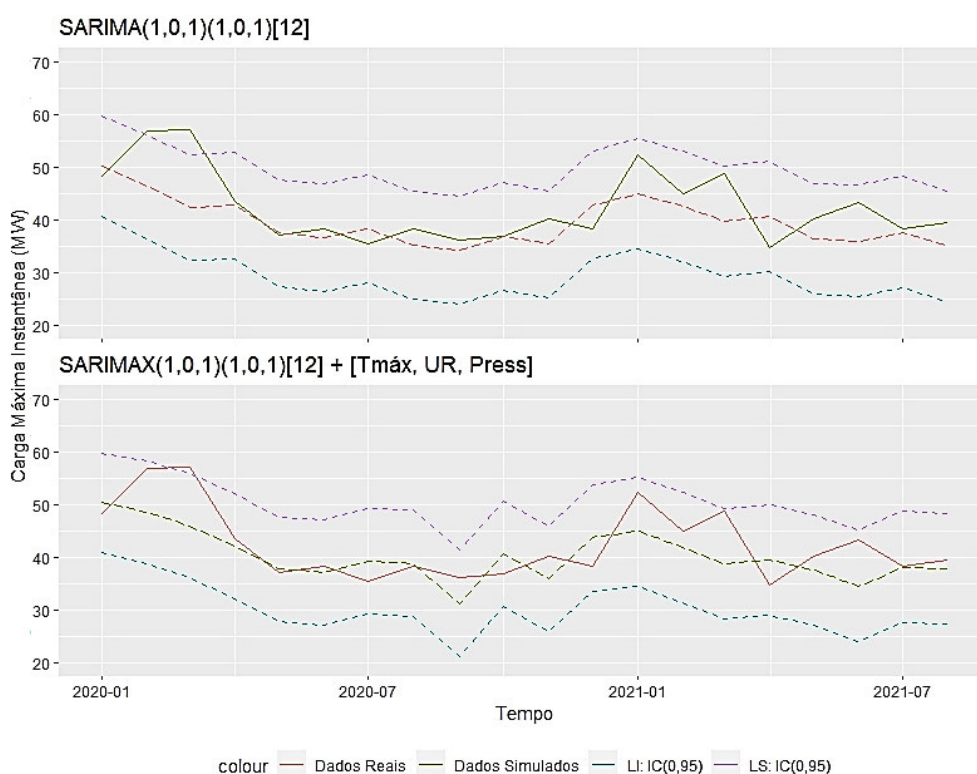


Figura 4.21 - Validação do patamar de carga pesada dos domingos: dados reais e simulados

Em relação ao ajuste do modelo $\text{SARIMA}(2,0,0)(1,0,0)_{[12]}$ e $\text{SARIMAX}(2,0,0)(1,0,0)_{[12]} + [\text{Tmáx}, \text{UR}, \text{Rad}]$ foi possível observar, por meio dos resultados expostos na Tabela 4.30, que o BIC aumentou consideravelmente com a inserção do conjunto de regressores das variáveis meteorológicas. Na análise residual, ao nível de 5% de significância, ambos os modelos estiveram de acordo com os critérios de independência e homoscedasticidade, e em nenhum caso foi identificada normalidade nas distribuições residuais.

Tabela 4.30. Patamar DP: características do período de ajuste dos modelos SARIMA ($\alpha = 0,05$)

Modelos SARIMA/ SARIMAX	BIC	Resultados Residuais		
		Ljung-Box: valor-p (Lag 24)	Jarque-Bera: valor-p (df 2)	Teste ARCH: valor-p (Lag 24)
$(2,0,0) (1,0,0)_{[12]}$	788,55	0,897	5,52E-13	0,071
$(2,0,0) (1,0,0)_{[12]} + [\text{Tmáx}, \text{UR}, \text{Rad}]$	796,66	0,899	3,91E-13	0,236

Fonte: Elaborado pela autora, 2023.

4.6.8. Análise dos Resultados

A análise exploratória dos dados de demanda de eletricidade identificou variações sazonais significativas, ou seja, existiram alterações cíclicas da demanda em períodos inferiores a um ano. Na maioria dos patamares, estas alterações ocorreram de forma semelhante, com picos no período do verão e valores menos acentuados durante os meses de inverno, porém houve exceção nos patamares de carga pesada, onde a sazonalidade esteve presente na variabilidade dos meses em si, independentemente de qual estação do ano pertencessem.

Em todos os patamares, os maiores índices de correlação dos dados da demanda de eletricidade ocorreram com variáveis do tipo temperatura, que foram consideradas como variáveis exógenas prioritárias, ou primárias, à entrada dos modelos SARIMAX.

A análise de correlação também identificou que as variáveis do tipo temperatura se relacionaram de forma diretamente proporcional com a demanda de eletricidade, ou seja, à medida que estas variáveis aumentavam, os índices de demanda de eletricidade também apresentavam aumento. Porém, as demais variáveis: umidade, radiação, pressão e precipitação; revelaram-se de maneira inversamente proporcional ao comportamento da demanda de energia elétrica, com exceção da radiação nos patamares de carga média dos dias úteis e carga pesada dos domingos (UM e DP, respectivamente).

Do ponto de vista da estacionariedade, conforme as informações exibidas no Apêndice 3, algumas séries meteorológicas apresentaram ordens de integração distintas dos dados de demanda de energia elétrica e, em detrimento dos pressupostos da modelagem Box-Tiao, não foram consideradas como variáveis exógenas dos modelos SARIMAX.

Os resultados dos modelos para os dias úteis revelaram que durante o período da madrugada e as primeiras horas da manhã, no intervalo da 01:00h até às 08:00h, exceto às segundas-feiras, houve influência da temperatura mínima na demanda de eletricidade. Ao longo do horário comercial, entre 09:00h e 17:00h, incluindo segundas-feiras, a demanda de eletricidade sofreu leve influência da temperatura mínima e também da temperatura mínima associada à umidade. Já nos intervalos diários de maior demanda, das 19:00h às 22:00h, não houve influência de componentes meteorológicos.

Por outro lado, aos finais de semana, os modelos mostraram mais variáveis meteorológicas influenciando nos índices de demanda de energia elétrica. Partindo da metade do período das manhãs e abrangendo grande parte das tardes dos sábados, mais especificamente das 09:00h às 17:00h, dentre os casos exitosos, foi identificada uma certa influência de todas as variáveis e conjuntos de variáveis meteorológicos testados, principalmente por parte de variáveis do tipo temperatura, consideradas individualmente. Nos intervalos de maior demanda dos sábados, entre 19:00h e 22:00h, com exceção

da temperatura mínima na forma isolada, todas as variáveis e conjunto de variáveis analisados apresentaram influência na demanda de energia elétrica, principalmente das variáveis do tipo temperatura associadas a fatores de umidade e pressão.

Aos domingos, das 10:00h às 17:00h nos meses de janeiro, fevereiro e março e das 09:00h às 17:00h nos demais meses do ano, foi verificada a influência da temperatura máxima na demanda de eletricidade, principalmente quando acompanhada da umidade e da pressão. E, por fim, nos intervalos de maior demanda, das 19:00h às 22:00h, observou-se influência da temperatura máxima e do ponto de orvalho, principalmente nos casos de associação com a umidade e a radiação.

Na comparação entre os resultados da parcimonia (BIC) e da capacidade preditiva (RMSE) dos modelos SARIMA, foi possível observar que, em todos os patamares, exceto a carga média dos sábados (SM), os modelos com ajuste razoável aos dados obtiveram menores erros na predição em relação aos modelos com ajuste ótimo. Em contrapartida, o patamar SM apresentou o modelo SARIMA de ajuste ótimo portando a melhor capacidade preditiva e, além disso, foi o único a melhorar os índices quanto à parcimonia ao ser transformado em SARIMAX, ou seja, ao inserir as respectivas variáveis exógenas, algo que, a princípio, tende a piorar tal característica, uma vez que, mais parâmetros são adicionados ao modelo.

Em relação ao intervalo de confiança (IC) dos modelos, ao nível de 95% de confiança, em todos os patamares de carga, em nenhum momento os dados reais estiveram inseridos entre os limites inferior e superior do IC dos modelos SARIMA e SARIMAX, independentemente dos regressores meteorológicos utilizados. Neste caso, uma explicação pode ser a ausência de normalidade dos dados, observada através dos resultados dos testes Jarque-Bera, visto que, a normalidade consiste em uma premissa da construção de intervalos de confiança a partir da aplicação da teoria assintótica de amostragem. Porém, a estrutura dos intervalos de confiança não foi o foco do presente estudo.

5. CONSIDERAÇÕES FINAIS

A bibliografia indicou que apesar de reconhecida uma certa hierarquia padrão para as variáveis meteorológicas mais influentes na demanda: temperatura, umidade, velocidade do vento, em ordem decrescente de importância, os fatores climáticos locais podem apresentar características muito peculiares, provocando alterações no comportamento padrão. Desta forma, é muito importante avaliar a relação das variáveis meteorológicas com a demanda de energia elétrica de uma região específica de interesse.

Portanto, o presente trabalho teve como objetivo analisar a influência das variáveis meteorológicas na demanda de eletricidade, ao longo de sete patamares de energia, encontrados na região da campanha do Rio Grande do Sul. O fator motivacional deste estudo foi verificar a possibilidade de melhorar a capacidade preditiva dos modelos de previsão de demanda de eletricidade acrescentando parâmetros meteorológicos da região.

A análise exploratória dos dados de demanda de eletricidade identificou sazonalidade em todas as séries. Desta forma, a metodologia de regressão linear aplicada neste estudo foi a modelagem Box-Jenkins para séries temporais com características sazonais, mais especificamente, aplicou-se o modelo Autorregressivo Integrado de Média Móvel Sazonal, ou SARIMA, a fim de prever o comportamento dos dados da demanda de eletricidade por meio das características no comportamento da própria variável de demanda de eletricidade.

A previsão por modelos do tipo SARIMA também foi utilizada na imputação dos dados de demanda de eletricidade e dos dados meteorológicos pertencentes às variáveis aptas à modelagem. O mesmo método foi aplicado em séries de origem divergente, em vista da estimativa dos dados de preenchimento ser realizada com base no comportamento das séries meteorológicas em si, ou seja, sem a necessidade de adquirir informações de estações meteorológicas adjacentes, conforme analisado no estudo de Veiga *et al.* (2021), as quais, também precisariam ser reorganizadas em patamares, algo que demandaria muito tempo da parte do analista. Desta forma, cabe a ressalva de explorar demais técnicas de imputação que não utilizam dados de estações vizinhas, a fim de adquirir e proporcionar mais opções de metodologias neste âmbito.

Por sua vez, a influência das variáveis meteorológicas na demanda de eletricidade foi verificada por meio da modelagem Box-Tiao, com a aplicação de modelos SARIMAX, correspondentes aos modelos SARIMA com fatores exógenos, os quais, neste caso, foram equivalentes às variáveis meteorológicas.

Em conformidade com a bibliografia, os maiores índices de correlação da demanda de eletricidade ocorreram com variáveis do tipo temperatura, as quais foram consideradas como exógenas primárias à entrada dos modelos SARIMAX.

Em relação à estacionariedade, algumas séries meteorológicas apresentaram ordens de integração distintas dos dados de demanda de energia elétrica e não foram consideradas como variáveis exógenas dos modelos. Sendo assim, em conformidade com Pereira (2021), é exequível a utilização de outras dinâmicas de análise do relacionamento destas variáveis com a demanda de eletricidade.

Os modelos revelaram influência da temperatura mínima durante o período da madrugada e as primeiras horas da manhã dos dias úteis, exceto segundas-feiras. Ao longo do horário comercial, a demanda sofreu leve influência da temperatura mínima e também da temperatura mínima associada à umidade. Aos finais de semana, uma maior quantidade de variáveis meteorológicas associadas a variáveis do tipo temperatura apresentaram influência na demanda de energia elétrica.

É importante ressaltar que os resultados adquiridos neste trabalho são referentes ao período de validação dos modelos, com início em janeiro de 2020 e estendendo-se até agosto de 2021, um período atípico em vista do estado de pandemia de COVID-19, declarado no primeiro semestre de 2020 pela Organização Mundial da Saúde (OMS), com regulamentações no Brasil sobre critérios de isolamento de quarentena (*lockdown*), onde diversas pessoas tiveram que aderir ao trabalho remoto, também conhecido por *home office* (PONTES *et al.*, 2021). A declaração do fim da Emergência de Saúde Pública da pandemia de COVID-19 foi feita pela OMS somente no primeiro semestre de 2023. Em relação ao período utilizado, outro fator relevante é sua duração de um ano e meio, no qual as informações de setembro a dezembro foram analisadas somente uma vez e os demais meses do ano tiveram ocorrência duplicada. Desta forma, existe a possibilidade da mesma metodologia aplicada neste estudo gerar diferentes resultados a partir da consideração de outro período na validação.

A partir da comparação entre as características de parcimonialidade e capacidade preditiva dos modelos, foi possível inferir que, muitas vezes, os modelos com ajuste razoável aos dados das séries possui um potencial maior de obter melhores índices de desempenho para predição em relação a um modelo com ajuste ótimo, em vista da possibilidade de ocorrências de alteração no padrão da variabilidade dos dados das séries temporais, fenômeno conhecido como *innovation*. Contudo, algumas séries apresentam um padrão estático na variabilidade, nestes casos, o modelo de ajuste ótimo também possui a melhor capacidade preditiva.

Acerca do intervalo de confiança dos modelos, em todos os casos avaliados, os dados reais não estiveram inseridos no IC. O motivo mais provável corresponde à ausência de normalidade dos

dados, uma vez que, tal propriedade consiste em um requisito fundamental na construção de intervalos de confiança a partir da aplicação da teoria assintótica de amostragem. Desta forma, em trabalhos posteriores, fica uma sugestão para aplicar outras técnicas de construção de intervalos de confiança, como Monte Carlo ou *bootstrap*, e possibilitar a análise e comparação com os resultados obtidos neste trabalho. Porém, a estruturação de IC não foi o foco desta pesquisa e, além disso, conforme Bueno (2011), mesmo que os dados não apresentem uma distribuição normal, é possível a obtenção de boas previsões de séries temporais utilizando modelos SARIMA.

Portanto, todos os objetivos propostos no presente estudo foram atingidos, elaborando-se um trabalho pertencente à área interdisciplinar, envolvendo conceitos de estatística, engenharia e meio ambiente, a fim de melhorar os subsídios para o planejamento energético, de forma a diminuir riscos e minimizar perdas, no intuito de proporcionar o aprimoramento das condições de fornecimento de energia elétrica à população consumidora.

6. REFERÊNCIAS

- AKAIKE, H. A. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v.19, n.6, p.716-723, 1974.
- AKAIKE, H. A. Bayesian analysis of the minimum AIC procedure. **Annals of the Institute of Statistical Mathematics A**, v.30, n.1, p.9-14, 1978.
- ALVES, M. F.; LOTUFO, A. D. P.; LOPES, M. L. M. Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics (SBMAC)**, v. 1, n. 1, 2013.
- ALVES, M. J. **Estudo da influência das variações de temperatura e umidade no comportamento da demanda da energia elétrica e formação do custo marginal de operação no Brasil**, 2008. Dissertação de Mestrado. SP: Escola Politécnica da Universidade de São Paulo, Departamento de Engenharia de Energia Automação Elétricas.
- APADULA, F.; BASSINI, A.; ELLI, A.; SCAPIN, S. Relationships between meteorological variables and monthly electricity demand. **Applied Energy**. v. 98, p. 346-356, 2012.
- ASSIS, E. S. A abordagem do clima urbano e aplicações no planejamento da cidade: reflexões sobre uma trajetória. Encontro Nacional e IV Encontro Latino-Americano sobre Conforto no Ambiente Construído. **Anais...** p. 92-101, 2005.
- BABA, R. K.; VAZ, M. S. M. G.; COSTA, J. DA. Correção de dados agrometeorológicos utilizando métodos estatísticos. **Revista Brasileira de Meteorologia**, v. 29, n. 4, p. 515-526, 2014.
- BARBOSA, M. T. M.; LIMA, A.; KUEHNE, B. T.; BATISTA, B. G.; FILHO, D. M. L.; PEIXOTO, M; L. M. Imputação de dados faltantes no monitoramento de consumo energético residencial em Smart Grids. Workshop de Computação Urbana. **Anais...** 2018. Disponível em: <<https://sol.sbc.org.br/index.php/courb/article/view/2344>>. .
- BDMEP. **Banco de dados meteorológicos do INMET**, 2022. Disponível em: <<https://bdmep.inmet.gov.br/>>.
- BENNETT, C.; STEWART, R. A.; LU, J. Autoregressive with exogenous variables and neural network short-term load forecast models for residential low voltage distribution networks. **Energies**, v. 7, n. 5, p. 2938-2960, 2014. Disponível em: <<https://www.mdpi.com/1996-1073/7/5/2938>>. .
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis: Forecasting and control**. [s.l.] San Francisco: Holden-Day, 1970.
- BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**, ed. 5, New York: John Wiley & Sons, 2015.
- BOX, G. E.; TIAO, G. C. Intervention analysis with application to economic and enviromental problems. **Journal of the American Statistical Association**, v. 70, p. 70-79, 1975.

- BRASIL. Lei nº 10.847 de 15 de Março de 2004. Autoriza a criação da Empresa de Pesquisa Energética – EPE e dá outras providências. **Diário Oficial da União**. seção 1, Brasília, DF, 2004.
- BRUBACHER, J. P.; OLIVEIRA, G. G. DE; GUASSELLI, L. A. Preenchimento de falhas e espacialização de dados pluviométricos: desafios e perspectivas. **Revista Brasileira de Meteorologia**, v. 35, n. 4, p. 615-629, 2020.
- BUENO, R. D. L. S. **Econometria de séries temporais**, ed. 2, São Paulo: Editora Cengage Learning, 2011.
- BUUREN, S. V.; OUDSHOORN, K. G. mice: Multivariate Imputation by Chained Equations in R. **Journal of Statistical Software**, v. 45, n. 3, p. 1-67, 2011.
- CAMELO, H. DO N. **Modelos Híbridos Estocástico - Matemático para Previsão de Velocidade do Vento**, 2018. Tese de Doutorado. RN: Natal, Programa de Pós-Graduação em Ciências Climáticas. Universidade Federal do Rio Grande do Norte.
- CAMPOS, S. F. B.; CUNHA, J. P. A. R. DA; LEMES, E. M. Controle de plantas daninhas avaliado visualmente e por imagens aéreas. **Revista Brasileira Multidisciplinar**, v. 24, n. 3, p. 216-225, 2021.
- CANOVA, F.; HANSEN, B. Are seasonal patterns constant over time? a test for seasonal stability, **Journal of business & economic statistics**, v. 13, p. 237–252, 1995.
- CARDOSO, F. C.; BERRI, R. A.; LUCCA, G.; BORGES, E. N.; MATTOS, V. L. D. de. Testes de Normalidade: estudo dos resíduos obtidos na modelagem da tendência de uma série temporal. **Exacta Engenharia de Produção**, Artigos online - aceitos, 2023.
- CARVALHO, J. V. **Modelagem temporal das medidas de vazão de drenos na Barragem de Funil (RJ) utilizando redes neurais e métodos estatísticos**, 2005. Tese de Doutorado. RJ: Rio de Janeiro, Programa de Pós-Graduação em Engenharia Civil da PUC-Rio, Pontifícia Universidade Católica do Rio de Janeiro.
- CARVALHO, L.; KUPFER, D. A Transição Estrutural da Indústria Brasileira: da Diversificação para a Especialização. XXXV Encontro Nacional de Economia da ANPEC. **Anais...** 2007. Disponível em <<https://www.anpec.org.br/encontro2007/artigos/A07A054.pdf>>. .
- CHATFIELD, C. **Time-Series Forecasting**. New York: Chapman and Hall/CRC, 2000.
- CHAVES, A. **A Revolução Tecnológica**, 2017. Disponível em: < <http://alaorchaves.com.br/wp-content/uploads/2017/02/Revolu%C3%A7%C3%A3o-Tecnol%C3%B3gica.pdf>>. .
- CHEN, P. Y.; POPOVICH, P. M. **Correlation: Parametric and Nonparametric Measures**. Thousand Oaks, CA: Sage Publications, 2002.
- CORREIA W. F. **Inclusão de métodos estatísticos como apoio ao faturamento de energia realizado por medidores inteligentes**. 2018. Dissertação de Mestrado. RJ: Rio de Janeiro, Programa de Pós-Graduação em Metrologia, Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio.

- COX, D. R.; STUART, A. Some Quick Sign Tests for Trend in Location and Dispersion, **Biometrika**, v. 42, n. 1/2, p.80-95, 1955.
- CRAN: COMPREHENSIVE R ARCHIVE NETWORK. **Time Series Missing Value Imputation**. Package 'imputeTS'. Version 3.2. 2021. Disponível em: <<https://cran.r-project.org/web/packages/imputeTS/imputeTS.pdf>>. .
- CURTIS, A. E.; SMITH, T. A.; ZIGANSHIN, B. A.; ELEFTERIADES, J. A. The Mystery of the Z-Score, **AORTA**, v. 4, n. 4, p. 124-130, 2016.
- DICKEY, D. A.; FULLER, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. **Econometrica: Journal of the Econometric Society**, v. 49, n. 4, p. 1057-1072, 1981. Disponível em: <http://www.jstor.org/stable/1912517?seq=1#page_scan_tab_contents>. .
- ANEEL: Agência Nacional de Energia Elétrica. **Resolução ANEEL N° 505 de 26 de Novembro de 2001**. Estabelece de forma atualizada e consolidada, as disposições relativas à conformidade dos níveis de tensão de energia elétrica em regime permanente. Diário Oficial da União (DOU), 2001.
- EMILIANO, P. C., VEIGA, E. P.; VIVANCO, M. J. F.; MENEZES, F. S. De. **Crítérios de informação de Akaike versus Bayesiano: Análise Comparativa**. 19º Simpósio Nacional de Probabilidade e Estatística, 2010.
- FALCÃO, A. J. T. **Detecção de Correlação e Causalidade em Séries Temporais não Categóricas**, 2012. Dissertação de Mestrado, Portugal: Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- FANTE, K. P., SANT'ANNA NETO, J. L. **Técnicas Estatísticas para Homogeneização de Dados de Temperatura em Séries Temporais Climatológicas**. Revista Brasileira de Climatologia, v. 18, n. 12, p. 143-156, 2016.
- FARIAS, D. A. A. de. **Previsão de carga de energia em patamares combinando técnicas estatísticas e de inteligência computacional**, 2009. Dissertação de Mestrado. RJ: Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.
- FARIAS, L. M.; SELLITTO, M. A. Uso da energia ao longo da história: evolução e perspectivas futuras. **Revista Liberato**, v. 12, n. 17, p. 01-106, 2011.
- FELICIANI, A. V.; SOUZA, A. M.; SOUZA, F. M. Métodos Paramétricos e Não Paramétricos na Estimção da Tendência. II Congresso Brasileiro de Engenharia de Produção. **Anais...** 2012. Disponível em: <<http://anteriores.aprepro.org.br/conbrepro/2012/anais/artigos/gestaoproducao/1.pdf>>. .
- FLORES, T.; OLIVEIRA, G. G. DE; HAETINGER, C.; BRESOLIN JÚNIOR, N. A.; GERHARD, F. Método para avaliação de eventos extremos de precipitação a partir de dados sub-horários obtidos por estação meteorológica automática. XXII Simpósio Brasileiro de Recursos Hídricos. **Anais...** 2017. Disponível em: <<https://lume.ufrgs.br/handle/10183/222147?locale-attribute=en>>. .
- FRACARO, N. **Estacionariedade das Séries Temporais do Modelo Matemático ARIMAX de Propulsores Eletromecânicos**. 2018. Dissertação de Mestrado. RS: Ijuí, Programa de Pós-

- Graduação em Modelagem Matemática, Unijuí: Universidade Regional do Noroeste do Estado do Rio Grande do Sul.
- GONÇALVES, L. DE F. **O método da correlação cruzada destendenciada em séries temporais**, 2018. Trabalho de Conclusão de Curso. RS: Instituto de Matemática e Física, Universidade Federal do Rio Grande do Sul.
- GUERRA, R. R. **Modelo Beta Autorregressivo de Médias Móveis: Critérios de Seleção e Aplicações**. 2015. Dissertação de Mestrado. RS: Santa Maria, Programa de Pós-Graduação em Engenharia de Produção, Centro de Tecnologia, Universidade Federal de Santa Maria.
- GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**, ed. 5, São Paulo: McGraw-Hill, 2010.
- GUNAY, M. E. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. **Energy Policy**, v. 90, p. 92-101, 2016.
- HANNAN, E.J.; QUINN, B.G. The determination of the order of an autoregression. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 41, n. 2, p. 190-195, 1979.
- HOOPER, D. U.; CHAPIN, F. S.; EWEL, J. J.; HECTOR, E.; INCHAUSTI, P.; LAVOREL, S.; LAWTON, J. H.; LODGE, D. M.; LOREAU, M.; NAEEM, S.; SCHMID, B.; SETALA, H.; SYMSTAD, A. J.; VANDERMEER, J.; WARDLE, D. A. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. **Ecological Monographs**, v. 75, p 3-35, 2005.
- HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. **Biometrika**. v. 76, n. 2, p. 297-307, 1989.
- HYLLEBERG, S.; ENGLE, R.; GRANGER, C.; YOO, B. Seasonal integration and cointegration, **Journal of econometrics**, v. 44, p. 215–238, 1990.
- HYNDMAN R. J.; KHANDAKAR, Y. Automatic Time Series Forecasting: The Forecast Package for R. **Journal of Statistical Software**, v.27, p.1-22, 2008.
- HYNDMAN, R.J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**, ed. 2, 2018. Disponível em: < <https://otexts.com/fpp2/index.html/>>. .
- HYNDMAN, R.J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**, ed. 3, 2021. Disponível em: <<https://otexts.com/fpp3/>>. .
- JARQUE, C. M.; BERA, A. K. A Test for Normality of Observations and Regression Residuals. **International Statistical Review**, v. 55, n. 2, p. 163-172, 1986.
- JOANES, D. N.; GILL, C. A. **Comparing measures of sample skewness and kurtosis. The Statistician**, v. 47, n. 1, p. 183-189, 1998.
- KENDALL, M. G. A new measure of rank correlation. **Biometrika**, v. 30, n. 1-2, p. 81-89, 1938.
- KHAN, P. W.; BYUN, Y. C.; LEE, S. J.; PARK, N. Machine Learning Based Hybrid System for

- Imputation and Efficient Energy Demand Forecasting. **Energies**, v. 13, n. 11, p. 2681, 2020.
- KIM, Y.; SON, H.; KIM, S. Short term electricity load forecasting for institutional buildings. **Energy Reports**, v. 5, p. 1270 - 1280, 2019.
- KWIATKOWSKI, D. PHILLIPS, P. C.; SCHMIDT, P.; SHIN, Y.; Testing the null hypothesis of stationarity against the alternative of a unit root. **Journal of econometrics**, v.54, n. 1-3, p.159-178, 1992.
- LEITE, M. DE S. **Testando normalidade multivariada**. RPubS by RStudio, 2018. Disponível em: <<https://rpubs.com/melinatarituba/356739>>. .
- LEVENBACH, H. Change & Chance Embraced, Editora: David Coen Freelance copyeditor, 2017.
- LEYBOURNE, S.; McCABE, B. A Consistent Test for a Unit Root. **Journal of Business & Economic Statistics**, v.12, n.2, p.157-66, 1994.
- LIMA, R. T. **Redes neurais artificiais aplicadas no controle de tensão de sistemas elétricos de potência**, 2008. Dissertação de Mestrado. RJ: Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.
- LIPPO, B. R. da S.; ROCHA, G. S. **Análise do Corte de Energia Elétrica sob o Prisma do Mínimo Existencial**. Cadernos de Graduação: Ciências Humanas e Sociais. v. 5, n. 1, p. 104-136. Pernambuco, 2021.
- LJUNG, G.M.; BOX, G. E. P. On a Measure of Lack of Fit in Time Series Models. **Biometrika**, v. 65, n. 2, p. 297-303, 1978.
- MACHADO, A. B.; RAMPINELLI, G. A. Avaliação de método probabilísticos para preenchimento de lacunas de dados solarimétricos e previsão da irradiância solar de curto prazo. VII Congresso Brasileiro de Energia Solar. **Anais...** 2018. Disponível em: <<https://anaiscbens.emnuvens.com.br/cbens/article/view/673>>. .
- MAIA, D. S. A energia elétrica e a relação espaço e tempo. IV Simposio Internacional de la Historia de la electrificación. **Anais...** 2017. Disponível em: < <http://www.ub.edu/geocrit/Electr-y-territorio/Indice-Electr-y-territorio.htm>>. .
- MICROSOFT EXCEL for windows 10. Versão 2203 [S. l.] **Microsoft Corporation**: 2019.
- MIRASGEDIS, S.; SARAFIDIS, Y.; GEORGOPOULOU, E.; LALAS, D. P.; MOSCHOVITS, M.; KARAGIANNIS, F.; PAPAKONSTANTINOU, D. Models for mid-term electricity demand forecasting incorporating weather influences. **Energy**. v. 31, p. 208-227, 2006.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to Linear Regression Analysis**, ed. 5, [S.l.]: Wiley, 2012.
- MORAIS JR. A. A. DE. **Elaboração de um Analisador Virtual utilizando Sistema Híbrido Neuro-Fuzzy para Inferir a Composição num Processo de Destilação**. 2011. Dissertação de Mestrado. AL: Maceió, Programa de Pós-Graduação em Engenharia Química, Unidade Acadêmica Centro de

- Tecnologia, Universidade Federal de Alagoas.
- MORETTIN, P. A.; TOLOI, C. M. **Análise de Séries Temporais**, ed. 2, Edgard Blucher, 2006.
- MORETTIN, P. A.; TOLOI, C. M. **Análise de séries temporais: modelos lineares univariados**, ed.3, Edgard Blucher, 2018.
- MORITZ, S. BEIELSTEIN, B. T. imputeTS: Time Series Missing Value Imputation in R. **The R Journal**, v. 9, n. 1, p. 207 - 2018, 2017.
- NAGUETTINI, M; PINTO, É. J. DE A. **Hidrologia Estatística**: CPRM, 2007.
- ONS. **Acompanhamento e previsão meteorológica e climática**. Procedimentos de Rede: Sub-Módulo 9.6. 2008a. Disponível em: <<https://www.ons.org.br/download/procedimentos/Submodulo9.6.pdf>>.
- OSBORN, D. R.; CHUI, A. P. L. ; SMITH, J.; BIRCHENHALL C.R. Seasonality and the order of integration for consumption, **Oxford Bulletin of Economics and Statistics**, v. 50, n. 4, p. 361-377, 1988.
- PARDO, A.; MENEU, V.; VALOR, E. Temperature and seasonality influences on Spanish electricity load. **Energy Economics**, v. 24, p. 55-70, 2002.
- PEARSON, E. S. **The test of significance for the correlation coefficient: some further results**. Journal of the American Statistical Association. v. 27, p. 424-426, 1932.
- PEREIRA, T. P. **O Uso dos Métodos SARIMA e SARIMAX na Modelagem de Demanda Mensal de Energia Elétrica no Estado do Rio Grande do Sul**, 2021. Dissertação de Mestrado. RS: Rio Grande, Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal do Rio Grande.
- PHILLIPS, P. C.; PERRON, P. Testing for a unit root in time series regression, **Biometrika**, v.75, n.2, p. 335-346, 1988.
- PONTES, A. L.; HIRATA, L. C.; NEVES, D. L. F. O Home Office na Pandemia de COVID-19: uma Perspectiva da Eficiência dos Profissionais da Tecnologia da Informação. **Revista Fatec Zona Sul**, ed. 1, v. 7, n. 5, 2021.
- R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2022. Disponível em: <<https://www.R-project.org/>>.
- RIBEIRO, A. E. **Tecnologia Digital**. Centro de Alfabetização, Leitura e Escrita (CEALE). Faculdade de Educação da UFMG, 2022. Disponível em: <<https://www.ceale.fae.ufmg.br/glossarioceale/verbetes/tecnologia-digital#:~:text=As%20tecnologias%20digitais%20surgiram%20no,seu%20presente%20e%20seu%20futuro>>. .
- RUGGLES, T.H.; FARNHAM, D. J.; TONG, D.; CALDEIRA, K. Developing reliable hourly electricity demand data through screening and imputation. **Sci Data**, v. 7, n. 155, 2020.

- SACRAMENTO, J. A. do; FERREIRA, V. H.; SCHILLING, M. T. Novos resultados para a modelagem probabilística da curva de carga global do SIN. IV Seminário Brasileira de Sistemas Elétricos (SBSE). **Anais...** p. 15-18, 2012.
- SANTOS, C. A. C dos.; BRITO, J. I. B de; RAO, T. V. R.; MENEZES, H. E. A. Tendências dos índices de precipitação no Estado do Ceará. **Revista Brasileira de Meteorologia**. v.24, n.1, p. 39-47, 2009.
- SCHWARZ, G. Estimating the dimension of a model. *Annals of Statistics*, v.6, n.2, p. 461-464, 1978.
- SHAPIRO, S. S.; WILK, M. B. An Analysis of Variance Test for Normality (Complete Samples), **Biometrika**, v. 52, n. 3/4, p. 591-611, 1965.
- SIEBEN, E. **Modelos para Previsão em Séries Temporais: A Metodologia Box e Jenkins na Prática**. 2017. Trabalho de Conclusão de Curso. RO: Ji-Paraná, Departamento de Matemática e Estatística, UNIR: Fundação Universidade Federal de Rondônia.
- SILVA, A. A. F.; JARDIM, C.H. Aplicação de técnicas de preenchimento de falhas de dados de pluviosidade mensal e anual para o noroeste do estado de Minas Gerais - Brasil. **Revista Geografias**. v. 15, n. 3, p. 83-106, 2017.
- SILVA, A. T.; NAGUETTINI, M.; PORTELA, M. M. Sobre a Estimação de Intervalos de Confiança para os Quantis de Variáveis Aleatórias Hidrológicas, **Revista Recursos Hídricos**, v. 32, n. 2, p. 63-76, 2011.
- SILVA, D. S. B. S.; JIMENEZ, J. C. V.; PEREIRA, C. C.; GOIS, G. DE; OLIVEIRA JÚNIOR, J. F. DE. Imputação de dados diários de radiação solar global via ambiente R. **Enciclopédia Biosfera**, v. 16, n. 29, p. 957-969, 2019.
- SILVA, M. V. da. **Um modelo matemático para estudo de otimização do consumo de energia elétrica**, 2007. Dissertação de Mestrado. SP: Universidade Estadual Paulista, Faculdade de Engenharia de Ilha Solteira.
- SILVEIRA, A. G. **Estudo da demanda de energia elétrica no Brasil**, 2017. Dissertação de Mestrado. RS: Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal do Rio Grande.
- SILVEIRA, A. G. **Estratégia para a definição do número de defasagens na aplicação do teste ADF**, 2023. Tese de Doutorado. RS: Rio Grande, Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal do Rio Grande.
- SPEARMAN, C. The Proof and Measurement of Association between Two Things. **The American Journal of Psychology**, v. 15, n. 1, p. 72-101, 1904.
- SUGIURA, N. Further analysts of the data by Akaike's information criterion and the finite corrections - Further analysts of the data by Akaike's. **Communications in Statistics - Theory and Methods**, v.7, n.1, p.13-26, 1978.
- THEIL, H. **Applied economic forecasting**. Amsterdam: North-Holland, 1966.

- USP: UNIVERSIDADE DE SÃO PAULO. **Testes de Tendência Central (média, mediana, proporção)**. E-disciplinas. Disponível em: <https://edisciplinas.usp.br/pluginfile.php/3832949/mod_resource/content/3/Rela%C3%A7%C3%A3o%20de%20testes%20de%20hip%C3%B3teses.pdf>. Acesso em: 5 maio. 2023.
- VALER, L. A. **Modelo Matemático ARIMAX de um Propulsor Eletromecânico utilizado em Naves do tipo Multirrotor**, 2016. Dissertação de Mestrado. RS: Ijuí, Programa de Pós-Graduação em Modelagem Matemática, Unijuí: Universidade Regional do Noroeste do Estado do Rio Grande do Sul.
- VEIGA, L. K. DA; MATTOS, V. L. D. DE; KONRATH, A. C. Modelagem Box-Jenkins aplicada à Imputação de Dados de Demanda de Energia Elétrica. **Revista Mundi Engenharia, Tecnologia e Gestão**, v. 8, n. 1, 2023.
- VEIGA, L. K. DA; MATTOS, V. L. D. DE; KONRATH, A. C.; NAKAMURA, L. R. Preenchimento de Falhas e Inferência Estatística em Variáveis Meteorológicas da Cidade de Pelotas, RS. XXIV Encontro Nacional de Modelagem Computacional (ENMC). **Anais...** 2021. Disponível em <https://www.even3.com.br/anais/xxivenmc_xiectm/420501-preenchimento-de-falhas-e-inferencia-estatistica-em-variaveis-meteorologicas-da-cidade-de-pelotas-rs/>. .
- VENTURA, T.M., SANTANA, L.L.R., Martins, C.A., and de Figueiredo, J.M. Análise da aplicabilidade de métodos estatísticos para preenchimento de falhas em dados meteorológicos. **Revista Brasileira de Climatologia**, v. 19, p. 168-177, 2016.
- VU, D.; MUTTAGI, K. M.; AGALGAONKAR, A. P. Assessing the influence of climatic variables on electricity demand. IEEE Power and Energy Society General Meeting. **Anais...** p. 1-5, 2014.
- WALD, A.; WOLFOWITZ, J. An Exact Test for Randomness in the Nonparametric Case Based on Serial Correlation. **The Annals of Mathematical Statistics**, v. 14, n. 4, p. 378-388, 1943.
- XAVIER, J. M. N. **Análise de Previsão de Séries Temporais com Modelos ARIMA e Análise Espectral Singular**, 2016. Mestrado em Bioestatística e Biometria, Universidade Aberta do Brasil.
- XIE, Y. 2015. **Dynamic Documents with R and Knitr**. 2 ed. Boca Raton, Florida: Chapman; Hall/ CRC, 2015. Disponível em: <<https://yihui.org/knitr/>>. .
- YOGARAJAH, B.; ELANKUMARAN, C.; VIGNESWARAN, R. Application of ARIMAX model for forecasting paddy production in Trincomalee district in Sri Lanka. Proceedings of the Third International Symposium. **Anais...** p. 21-25, 2013. Disponível em: <<http://ir.lib.seu.ac.lk/handle/123456789/370>>. .

APÊNDICE 1 – Medidas resumo das séries

UL	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	981.7992	-3.5050	5.5188	4.3188	7.5000	-2.2000	71.9688	6.4000	0.2644	13.6203
1st Qu.	1.0000	986.2225	-3.2554	11.9805	9.6456	20.5000	3.0750	83.8316	12.3000	2.5891	23.9564
Median	3.8000	988.7628	-3.0668	15.0836	12.6040	23.8000	6.9000	86.0347	14.0000	3.0279	25.9607
Mean	7.1827	988.8505	-3.0121	15.0834	12.6199	23.2387	6.8347	85.7990	14.8701	2.9331	26.9416
3rd Qu.	10.1000	991.2347	-2.8131	18.4844	15.7787	25.7000	10.5500	87.7185	16.7000	3.3816	29.3399
Max.	48.6000	998.2063	-2.0147	22.3486	19.6118	30.1000	16.4000	93.7578	40.0000	4.1993	42.5788
var	83.3949	8.9162	0.1036	14.0054	13.1466	13.1291	21.6690	10.8323	16.2240	0.4438	24.3077
sd	9.1321	2.9860	0.3219	3.7424	3.6258	3.6234	4.6550	3.2912	4.0279	0.6662	4.9303
assimetria	2.0322	0.1696	0.7757	-0.1770	-0.1528	-0.6838	0.0357	-0.6394	2.0499	-1.0820	0.6823
curtose	4.5665	-0.4201	0.2538	-0.9293	-0.8975	1.1548	-0.9930	1.6981	9.0046	2.1947	0.9535
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	5.0000	5.0000	19.0000

UM	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	982.9956	592.8675	8.8136	5.3667	16.4000	-2.7000	50.9365	10.4000	1.0306	25.7150
1st Qu.	2.8500	986.9917	1016.4671	15.0001	10.3504	28.0250	2.5000	69.8351	14.3500	3.2106	43.2326
Median	7.2000	989.2331	1343.4519	18.8000	13.0697	30.7000	6.2500	73.4217	16.2000	3.6203	47.7746
Mean	8.5693	989.5029	1376.9045	18.6439	13.0407	30.4033	6.0493	72.9345	16.4225	3.6165	50.1243
3rd Qu.	12.1500	991.8376	1722.5460	22.4980	15.9438	33.1500	9.5250	77.2128	18.1500	4.0979	55.4440
Max.	32.4000	998.0455	2539.6092	25.7150	20.2687	38.3000	16.4000	85.3492	28.3000	5.4082	79.8865
var	56.30268	8.07741	208361.2	19.06263	12.33522	15.61294	21.73688	40.82005	9.62970	0.54184	104.67490
sd	7.50351	2.84208	456.4660	4.36608	3.51215	3.95132	4.66228	6.38906	3.10318	0.73610	10.23108
assimetria	1.06166	0.14410	0.231392	-0.16397	-0.07579	-0.46865	0.09163	-0.58773	0.77098	-0.63460	0.64583
curtose	0.39584	-0.52973	-0.948151	-1.16801	-0.89501	0.04508	-0.91491	0.51007	0.91669	1.43533	-0.00811
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	5.0000	5.0000	20.0000

UP	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	981.0325	141.7145	11.1417	5.4489	17.2000	2.2000	37.1630	7.6000	0.6738	25.7150
1st Qu.	1.2000	985.1408	445.2744	16.8216	10.2674	27.4000	8.4500	57.7066	13.0000	2.9427	46.7235
Median	3.8000	987.7961	732.6003	20.5573	12.7731	30.8500	12.7000	64.0551	14.3000	3.4417	50.5029
Mean	6.4400	987.8962	810.6449	20.8202	12.7484	30.3660	12.6813	62.9976	14.8279	3.3531	51.9997
3rd Qu.	9.4000	990.4186	1175.6225	24.9720	14.9693	33.5750	17.3750	69.9296	16.4500	3.8211	57.1784
Max.	43.2000	996.7333	1669.5660	29.0830	19.7750	37.6000	22.3000	80.7857	25.3000	5.0905	72.0853
var	56.31463	9.72132	162090.1	23.33629	10.90482	16.01837	28.71347	87.39655	9.30312	0.57698	55.26938
sd	7.50431	3.11790	402.6042	4.83077	3.30225	4.00230	5.35850	9.34861	3.05010	0.75959	7.43434
assimetria	2.11568	0.09757	0.2395557	-0.10502	-0.01072	-0.43910	-0.13528	-0.47452	0.87862	-0.83237	0.10345
curtose	5.70738	-0.69818	-1.150771	-1.21540	-0.74247	-0.30240	-1.24544	-0.19073	1.43757	1.61358	0.45977
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	5.0000	5.0000	17.0000

SM	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	981.8861	516.4348	8.9722	3.7056	18.9000	-2.7000	53.6389	6.3000	0.8167	25.6424
1st Qu.	0.0500	987.5326	936.8007	15.5550	10.3854	25.1250	5.5750	68.8111	10.7500	3.0042	36.8728
Median	0.2000	989.5667	1256.7309	18.7356	13.3267	28.0000	9.0000	73.2222	12.5000	3.5467	39.8451
Mean	3.0333	989.4516	1358.4440	18.8113	13.1460	27.9973	9.1473	72.6405	13.1510	3.6294	41.4738
3rd Qu.	3.3500	991.4986	1783.2028	22.6701	15.9451	31.7000	12.8750	77.8819	15.1500	4.2589	44.6345
Max.	28.8000	995.5472	2499.1893	26.5778	20.4750	39.1000	19.0000	88.5000	23.0000	6.0884	72.7215
var	30.55754	8.81820	251523.6	19.16346	12.67167	19.74013	24.70076	50.57850	10.34060	1.10373	59.90889
sd	5.52789	2.96955	501.5212	4.37761	3.55973	4.44299	4.96999	7.11186	3.21568	1.05058	7.74008
assimetria	2.37355	-0.05392	0.2945546	-0.13213	-0.12027	-0.06051	-0.11050	-0.48358	0.71204	0.03775	1.29931
curtose	5.43006	-0.41621	-0.989744	-1.07345	-0.68671	-0.67591	-0.78510	-0.12388	0.12664	-0.05748	2.57936
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	5.0000	5.0000	20.0000

SP	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	979.8500	96.2652	10.8438	2.8438	17.9000	3.4000	36.6875	0.0000	0.2313	25.7150
1st Qu.	0.0000	985.5812	443.0244	16.9266	10.4453	24.7000	10.8250	57.0719	9.8750	2.7281	45.9983
Median	0.0000	987.5738	713.8024	20.6575	13.1588	27.7000	15.2500	64.1563	11.6000	3.3900	49.3921
Mean	1.8453	987.8125	796.2299	20.9808	12.9664	28.0007	15.1787	63.1485	11.8128	3.4176	49.7699
3rd Qu.	0.7500	990.1363	1164.3304	25.2453	15.2300	31.8750	19.3000	70.0906	12.9500	4.1656	53.6589
Max.	39.2000	994.7000	1783.5413	30.4688	21.2625	38.2000	26.6000	89.0000	25.1000	6.0941	62.8784
var	25.742360	10.370900	180808.2	23.78549	12.20321	21.02638	29.61417	107.87780	10.20562	1.14685	31.73746
sd	5.073693	3.220388	425.2154	4.87704	3.49331	4.58545	5.44189	10.38642	3.19462	1.07091	5.63360
assimetria	4.470193	-0.013189	0.4816485	-0.06429	-0.17155	-0.13342	-0.05162	-0.33307	0.71050	-0.30673	-0.63137
curtose	23.575720	-0.631370	-0.915491	-1.17718	-0.42353	-0.78464	-0.94933	-0.19424	2.52945	0.47410	2.49990
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	4.0000	5.0000	19.0000

DM	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	982.6000	560.8771	9.5889	4.8028	18.5000	-2.4000	54.3438	4.7000	1.2469	18.6607
1st Qu.	0.0000	987.1347	978.4358	15.5706	10.7263	25.4500	5.0000	65.6250	11.2000	3.1358	33.1686
Median	0.2000	989.2352	1350.7007	19.1764	13.2581	28.4000	8.9500	71.1691	12.6000	3.6750	35.4144
Mean	2.1560	989.4074	1464.5880	19.1426	13.1727	28.1187	9.2607	71.3244	13.1633	3.7136	37.1680
3rd Qu.	1.6000	991.3646	1930.8840	22.6569	16.0133	31.1500	13.5750	77.1667	14.9500	4.2889	40.1409
Max.	38.4000	997.1033	2720.8629	27.9469	20.9594	37.3000	19.6000	86.3889	24.2000	6.9889	57.8914
var	24.00369	9.50936	306406.6	21.28756	13.80152	19.01455	27.66965	55.00768	10.62713	0.88699	43.53100
sd	4.89936	3.08373	553.5401	4.61384	3.71504	4.36057	5.26020	7.41672	3.25993	0.94180	6.59780
assimetria	4.31101	0.14475	0.2544628	-0.07042	-0.13941	-0.27300	-0.03219	-0.08341	0.56310	0.22450	0.89944
curtose	23.76309	-0.40363	-1.114063	-1.03111	-0.73548	-0.55877	-0.93126	-0.76754	0.48317	0.50405	1.12123
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	5.0000	5.0000	16.0000

DP	Pc	Press	Rad	Temp	Orv	Tmáx	Tmín	UR	Raj	VV	Carga
Min.	0.0000	980.3500	158.6336	11.8000	4.8100	18.4000	3.7000	37.8750	4.1000	0.0688	25.7150
1st Qu.	0.0000	984.9703	425.9975	16.7266	10.5147	25.1250	10.6250	54.7031	10.0000	2.7969	43.4283
Median	0.0000	987.5125	745.2503	21.5156	13.0638	28.2500	15.0500	62.8750	11.6000	3.4000	46.7173
Mean	1.9133	987.7552	820.7788	21.0913	12.8728	28.1133	15.0187	62.3767	11.8157	3.3601	47.4233
3rd Qu.	1.2000	990.0094	1156.5250	24.7975	15.2694	31.0500	19.4750	70.1000	13.3500	3.9656	50.7451
Max.	49.2000	995.9438	1834.9354	30.8200	20.9125	37.3000	28.4000	82.6875	21.3000	6.4500	63.9280
var	28.70412	11.80169	184988.0	24.12244	12.08109	20.79284	31.46368	102.77800	7.90489	1.11910	34.30216
sd	5.35762	3.43536	430.1023	4.91146	3.47579	4.55992	5.60925	10.13795	2.81156	1.05788	5.85681
assimetria	5.72741	0.25460	0.3603486	-0.07566	-0.10346	-0.21711	-0.01729	-0.12025	0.45503	-0.26023	0.10453
curtose	41.76605	-0.50311	-0.990269	-1.08308	-0.39874	-0.69978	-0.84797	-0.71177	0.95019	0.66468	1.68749
NA's	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	5.0000	5.0000	17.0000

APÊNDICE 2 – Coeficientes de correlação

	Prec	Press	Rad	Temp	Orvalho	Tmax	Tmin	UR	Rajada	VV	Demanda_UL
Prec	1.0000000	-0.1894040	0.2371065	0.1634285	0.1765711	0.1984052	0.1146988	0.0719667	0.2557312	0.0972985	0.0380568
Press	-0.1894040	1.0000000	0.1775641	-0.8157218	-0.7796396	-0.7227127	-0.7535495	0.2423413	-0.2503626	0.0106351	-0.4570061
Rad	0.2371065	0.1775641	1.0000000	-0.2855604	-0.2086096	-0.2300574	-0.2508207	0.4550057	0.1290402	0.2152926	-0.3435242
Temp	0.1634285	-0.8157218	-0.2855604	1.0000000	0.9826552	0.8178693	0.9165407	-0.1595052	0.0755628	-0.0250201	0.6054214
Orvalho	0.1765711	-0.7796396	-0.2086096	0.9826552	1.0000000	0.7653010	0.9191773	-0.0254959	0.0507960	-0.0709244	0.5752114
Tmax	0.1984052	-0.7227127	-0.2300574	0.8178693	0.7653010	1.0000000	0.6759413	-0.3326351	0.1787505	0.1266361	0.4782125
Tmin	0.1146988	-0.7535495	-0.2508207	0.9165407	0.9191773	0.6759413	1.0000000	-0.0469789	0.0096346	-0.1042648	0.5424314
UR	0.0719667	0.2423413	0.4550057	-0.1595052	0.0254959	-0.3326351	-0.0469789	1.0000000	-0.1471026	-0.2385220	-0.2080768
Rajada	0.2557312	-0.2503626	0.1290402	0.0755628	0.0507960	0.1787505	0.0096346	-0.1471026	1.0000000	0.3097406	-0.1215502
VV	0.0972985	0.0106351	0.2152926	-0.0250201	-0.0709244	0.1266361	-0.1042648	-0.2385220	0.3097406	1.0000000	-0.1306962
Demanda_UL	0.0380568	-0.4570061	-0.3435242	0.6054214	0.5752114	0.4782125	0.5424314	-0.2080768	-0.1215502	-0.1306962	1.0000000
	Prec	Press	Rad	Temp	Orvalho	Tmax	Tmin	UR	Rajada	VV	Demanda_UM
Prec	1.0000000	-0.1802273	-0.0948551	0.0887205	0.1989298	0.0769314	0.1158191	0.1857354	0.2158474	-0.0204741	0.0229973
Press	-0.1802273	1.0000000	-0.7043262	-0.8171302	-0.7615910	-0.6583841	-0.7676687	0.4648570	-0.1007636	0.0885153	-0.3271464
Rad	-0.0948551	-0.7043262	1.0000000	0.8097382	0.5765054	0.7601669	0.6703289	-0.8308824	-0.1040133	0.0026036	0.4046870
Temp	0.0887205	-0.8171302	0.8097382	1.0000000	0.9191098	0.8320938	0.9121022	-0.6034752	-0.0553177	-0.0770500	0.5147580
Orvalho	0.1989298	-0.7615910	0.5765054	0.9191098	1.0000000	0.6710910	0.9079895	-0.2432543	-0.0546432	-0.1279574	0.4965266
Tmax	0.0769314	-0.6583841	0.7601669	0.8320938	0.6710910	1.0000000	0.6908590	-0.6888431	-0.0721813	0.0568391	0.5245486
Tmin	0.1158191	-0.7676687	0.6703289	0.9121022	0.9079895	0.6908590	1.0000000	-0.4029671	-0.1021366	-0.1207284	0.5555439
UR	0.1857354	0.4648570	-0.8308824	-0.6034752	-0.2432543	-0.6888431	-0.4029671	1.0000000	0.1399872	-0.0702442	-0.3327587
Rajada	0.2158474	-0.1007636	-0.1040133	-0.1055317	-0.0546432	-0.0721813	-0.1021366	0.1399872	1.0000000	0.3606794	-0.3404622
VV	-0.0204741	0.0885153	0.0026036	-0.0770500	-0.1279574	0.0568391	-0.1207284	-0.0702442	0.3606794	1.0000000	-0.1126668
Demanda_UM	0.0229973	-0.3271464	0.4104874	0.5467580	0.4965262	0.5245486	0.5555439	-0.3327587	-0.3404622	-0.1126668	1.0000000
	Prec	Press	Rad	Temp	Orvalho	Tmax	Tmin	UR	Rajada	VV	Demanda_UP
Prec	1.0000000	-0.2757808	0.2332923	0.2478570	0.3512116	0.2024754	0.2205976	0.0140507	0.2062437	0.1067763	-0.0835622
Press	-0.2757808	1.0000000	-0.8018424	-0.8518059	-0.7245660	-0.7388217	-0.8000767	0.5515081	-0.3319468	-0.2931105	-0.0146098
Rad	0.2332923	-0.8018424	1.0000000	0.8788052	0.5974491	0.7361803	0.8146072	-0.7738408	0.2402119	0.3517743	-0.0396359
Temp	0.2478570	-0.8518059	0.8788052	1.0000000	0.8146900	0.8673206	0.9408189	-0.7140843	0.1637598	0.2708222	-0.1576670
Orvalho	0.3512116	-0.7245660	0.5974491	0.8146900	1.0000000	0.6627006	0.8185392	-0.1827313	0.1062100	0.1195651	0.1597628
Tmax	0.2024754	-0.7388217	0.7361803	0.8673206	0.6627006	1.0000000	0.7513435	-0.6682287	0.1859076	0.2899181	0.2080446
Tmin	0.2205976	-0.8000767	0.8146072	0.9408189	0.8185392	0.7513435	1.0000000	-0.6155448	0.0876346	0.2726445	0.1656349
UR	0.0140507	0.5515081	-0.7738408	-0.7140843	-0.1827313	-0.6682287	-0.6155448	1.0000000	-0.1397764	-0.2920954	-0.0701416
Rajada	0.2062437	-0.3319468	0.2402119	0.1637598	0.1062100	0.1859076	0.0876346	-0.1397764	1.0000000	0.3110423	-0.1870722
VV	0.1067763	-0.2931105	0.3517743	0.2708222	0.1195651	0.2899181	0.2276445	-0.2920954	0.3110423	1.0000000	-0.1406621
Demanda_UP	-0.0835622	-0.0146098	-0.0396359	0.1576670	0.1597628	0.2080446	0.1656349	-0.0701416	-0.1870722	-0.1406621	1.0000000
	Prec	Press	Rad	Temp	Orvalho	Tmax	Tmin	UR	Rajada	VV	Demanda_SM
Prec	1.0000000	-0.1619274	-0.0922832	0.0657848	0.1709925	-0.0416349	0.1834478	0.2112684	0.1930575	0.0443034	-0.0656306
Press	-0.1619274	1.0000000	-0.4778723	-0.7377306	-0.7215784	-0.5720543	-0.7504136	0.2965410	-0.0620432	-0.0243741	-0.3019486
Rad	-0.0922832	-0.4778723	1.0000000	0.7504781	0.5022913	0.6853698	0.5367937	-0.7820692	-0.1163205	-0.0595616	0.4345971
Temp	0.0657848	-0.7377306	0.7504781	1.0000000	0.9038167	0.8612249	0.8759586	-0.5717643	0.0412109	-0.0393487	0.5688155
Orvalho	0.1709925	-0.7215784	0.5022913	0.9038167	1.0000000	0.6862204	0.8931691	-0.1703106	0.0012777	-0.0678882	0.4445322
Tmax	-0.0416349	-0.5720543	0.6853698	0.8612249	0.6862204	1.0000000	0.6746050	-0.6569242	-0.0297866	-0.0266396	0.5952466
Tmin	0.1834478	-0.7504136	0.5367937	0.8759586	0.8931691	0.6746050	1.0000000	-0.2969627	0.0075561	0.0133956	0.4158625
UR	0.2112684	0.2965410	-0.7820692	-0.5717643	-0.1703106	-0.6569242	-0.2969627	1.0000000	0.0743835	-0.0720603	-0.4386720
Rajada	0.1930575	-0.0620432	-0.1163205	-0.0412109	0.0012777	-0.0297866	0.0075561	0.0743835	1.0000000	0.5872066	-0.1703793
VV	0.0443034	-0.0243741	-0.0595616	-0.0393487	-0.0678882	-0.0266396	0.0133956	-0.0720603	0.5872066	1.0000000	-0.1491805
Demanda_SM	-0.0656306	-0.3019486	0.4345971	0.5688155	0.4445322	0.5952466	0.4158625	-0.4386720	-0.1703793	-0.1491805	1.0000000
	Prec	Press	Rad	Temp	Orvalho	Tmax	Tmin	UR	Rajada	VV	Demanda_SP
Prec	1.0000000	-0.2599325	0.0894763	0.1700162	0.2566251	0.2035185	0.1170107	0.0456720	0.4908443	0.1295521	0.0401054
Press	-0.2599325	1.0000000	-0.6393270	-0.7740171	-0.6677493	-0.6609398	-0.7425182	0.4078532	-0.2149074	-0.2951252	-0.0471625
Rad	0.0894763	-0.6393270	1.0000000	0.8254954	0.4990079	0.6706355	0.8006241	-0.7419251	0.1839393	0.2428886	0.0406112
Temp	0.1700162	-0.7740171	0.8254954	1.0000000	0.7873021	0.8622165	0.9462916	-0.6671555	0.1628404	0.2386719	0.2196548
Orvalho	0.2566251	-0.6677493	0.4990079	0.7873021	1.0000000	0.6368804	0.7786418	-0.0758754	0.1434676	0.1727040	0.1768411
Tmax	0.2035185	-0.6609398	0.6706355	0.8622165	0.6368804	1.0000000	0.7373472	-0.6193285	0.2416534	0.2373054	0.2533300
Tmin	0.1170107	-0.7425182	0.8006241	0.9462916	0.7786418	0.7373472	1.0000000	-0.5892769	0.0734620	0.2134779	0.1529908
UR	0.0456720	0.4078532	-0.7419251	-0.6671555	-0.0758754	-0.6193285	-0.5892769	1.0000000	-0.0802368	-0.1688009	-0.1681085
Rajada	0.4908443	-0.2149074	0.1839393	0.1628404	0.1434676	0.2416534	0.2373054	-0.0802368	1.0000000	0.5011472	0.0192299
VV	0.1295521	-0.2951252	0.2428886	0.2386719	0.1727040	0.2373054	0.2134779	-0.1688009	0.5011472	1.0000000	-0.0923647
Demanda_SP	0.0401054	-0.0471625	0.0406112	0.2196548	0.1768411	0.2533300	0.1529908	-0.1681085	0.0192299	-0.0923647	1.0000000
	Prec	Press	Rad	Temp	Orvalho	Tmax	Tmin	UR	Rajada	VV	Demanda_DM
Prec	1.0000000	-0.1728861	-0.1779918	-0.0310447	0.0739071	-0.0254995	0.0396307	0.2503479	0.2931356	0.1015787	-0.0555939
Press	-0.1728861	1.0000000	-0.5465720	-0.7586993	-0.7382415	-0.6643795	-0.7134561	0.3343121	-0.1428485	-0.0477560	-0.1913151
Rad	-0.1779918	-0.5465720	1.0000000	0.7851747	0.5792430	0.6520738	0.5905597	-0.7456085	-0.0654096	-0.1265585	0.3586145
Temp	-0.0310447	-0.7586993	0.7851747	1.0000000	0.9146156	0.8743067	0.8873173	-0.5696288	-0.0584806	-0.0524979	0.4743612
Orvalho	0.0739071	-0.7382415	0.5792430	0.9146156	1.0000000	0.7494048	0.8852107	-0.1952122	-0.1172637	-0.1120491	0.4124006
Tmax	-0.0254995	-0.6643795	0.6520738	0.8743067	0.7494048	1.0000000	0.7157431	-0.5908757	0.0349970	0.0155632	0.4420076
Tmin	0.0396307	-0.7134561	0.5905597	0.8873173	0.8852107	0.7157431	1.0000000	-0.3518996	-0.042.		

APÊNDICE 3 – Resultados dos testes de Estacionariedade

Demanda de Carga Leve - Dias Úteis (UL)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	0	0	0	0	0	0
Temperatura	0	0	0	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	0
Radiação	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Pressão	0	0	0	0	0	0	0	0
Orvalho	0	0	0	0	0	0	0	2
Temp. Máxima	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2
Rajada	0	0	0	0	0	0	0	2

Demanda de Carga Média - Dias Úteis (UM)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	1	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	0
Temperatura	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Radiação	0	0	0	0	0	0	0	0
Pressão	0	0	0	0	0	0	0	0
Orvalho	0	0	0	0	0	0	0	2
Temp. Máxima	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2
Rajada	0	0	1	0	0	0	0	2

Demanda de Carga Pesada - Dias Úteis (UP)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	1	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	0
Temperatura	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Orvalho	0	0	0	0	0	0	0	2
Temp. Máxima	0	0	0	0	0	0	0	2
Precipitação	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2
Rajada	0	0	0	0	0	0	0	2

Demanda de Carga Média - Sábado (SM)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	1	0	0	0	0	0
Temp. Máxima	0	0	0	0	0	0	0	0
Temperatura	0	0	0	0	0	0	0	0
Orvalho	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	0
Pressão	0	0	0	0	0	0	0	0
Precipitação	0	0	0	0	0	0	0	0
Radiação	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2
Rajada	0	0	0	0	0	0	0	2

Demanda de Carga Pesada - Sábado (SP)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	1	0	0	0	0	0
Temp. Máxima	0	0	0	0	0	0	0	0
Temperatura	0	0	0	0	0	0	0	0
Orvalho	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	0
Pressão	0	0	0	0	0	0	0	0
Rajada	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2

Demanda de Carga Média - Domingo (DM)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	1	0	0	0	0	0
Temp. Máxima	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Pressão	0	0	0	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	2
Orvalho	0	0	0	0	0	0	0	1
Precipitação	0	0	0	1	0	0	0	2
Radiação	0	0	0	0	0	0	0	2
Temperatura	0	0	0	0	0	0	0	2
Rajada	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2

Demanda de Carga Pesada - Domingo (DP)								
Teste →	Parte Não-Sazonal						Parte Sazonal	
	ADF		KPSS		PP		OCSB	HEGY
Variável ↓	Level	Trend	Level	Trend	Level	Trend		
Demanda	0	0	1	1	0	0	0	0
Temp. Máxima	0	0	0	0	0	0	0	0
Orvalho	0	0	0	0	0	0	0	0
Umidade	0	0	0	0	0	0	0	0
Radiação	0	0	0	0	0	0	0	0
Temp. Mínima	0	0	0	0	0	0	0	2
Temperatura	0	0	0	0	0	0	0	2
Velocidade Vento	0	0	0	1	0	0	0	2
Rajada	0	0	0	0	0	0	0	2