

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

**O USO DOS MÉTODOS ARIMA E VAR-VEC NO ESTUDO DA DEMANDA DE
ENERGIA ELÉTRICA NO RIO GRANDE DO SUL**

por

Mestrando: Gérson dos Santos Nunes
Orientadora: Viviane Leite Dias de Mattos
Coorientadora: Andrea Cristina Konrath

Projeto de Dissertação para obtenção do Título de
Mestre em Modelagem Computacional

Rio Grande, maio de 2019.

O uso dos métodos ARIMA e VAR-VEC no estudo da demanda de energia elétrica no
Rio Grande do Sul

Gérson dos Santos Nunes

BANCA EXAMINADORA

Prof.^a. Dra. Viviane Leite de Dias Mattos
Instituto de Matemática, Estatística e Física - FURG

Prof.^a. Dra. Andrea Cristina Konrath
Departamento de Informática e Estatística - UFSC

Prof.^a. Dr. Luiz Ricardo Nakamura
Departamento de Informática e Estatística - UFSC

Prof. Dr. André Andrade Longaray
Departamento de Ciências Econômicas, Administrativas e Contábeis - FURG

Rio Grande, maio de 2019

AGRADECIMENTOS

Agradeço aos meus pais, Adão Botelho Nunes e Elizaete dos Santos Nunes, pela dedicação e incentivo durante toda minha trajetória acadêmica.

A minha orientadora, Viviane Leite Dias de Mattos, por seus ensinamentos, sua dedicação, compreensão e disponibilidade durante toda a execução do trabalho.

A minha coorientadora, Andrea Cristina Konrath, por dedicar parte de seu tempo para contribuir em aprimorar este trabalho.

A minha esposa, Silvana Rodrigues Nunes, e ao meu filho, Pedro Henrique Rodrigues Nunes, pela compreensão nos meus momentos de ausência e seus incentivos durante todo o período de execução do trabalho.

Ao meu irmão, Sérgio Adão dos Santos Nunes, e minha irmã, Celina Nunes Silva, por sempre me apoiarem e pelos momentos de descontração que tornaram muitos momentos de dificuldade mais leves.

Aos membros da banca Luiz Ricardo Nakamura e André Andrade Longaray, por terem aceitado o convite e pelas importantes considerações realizadas.

Aos professores do Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal do Rio Grande – FURG, que contribuíram para a construção do meu conhecimento.

À Fundação Getúlio Vargas, pelo fornecimento sem custo dos dados, sem os quais a realização deste trabalho não seria possível.

Ao Instituto Federal Sul Rio Grandense, por permitir minha redução de carga horária para que pudesse me dedicar aos estudos.

Aos meus familiares e amigos, que muitas vezes tiveram que compreender meus momentos de ausência.

RESUMO

A demanda de energia elétrica mostra-se como um fator importante para o crescimento da economia de uma região. Visando ajudar no planejamento do setor energético do Rio Grande do Sul, este trabalho abordou a demanda nas três principais classes consumidoras de energia elétrica deste estado: residencial, comercial e industrial, utilizando as metodologias Box-Jenkins, Vetores Auto-regressivos e Vetores Autoregressivos em conjunto com Correção de Erros Vetoriais, comparando os resultados obtidos para a previsão pelos diferentes modelos, assim como com os resultados da aplicação conjunta de Vetores Autoregressivos e Correção de Erros Vetoriais obtidos em recente estudo desenvolvido a nível nacional. Após a realização da transformação logarítmica, são utilizados correlogramas, correlogramas parciais, testes de raiz unitária ADF e KPSS, critérios de informação de Akaike, Schwarz e Hannan-Quinn, avaliação da significância dos coeficientes e análise dos resíduos. Foram utilizados os testes Ljung-Box, Jarque-Bera e ARCH para avaliar, respectivamente, autocorrelação, normalidade e homocedasticidades dos resíduos, além da determinação do erro absoluto médio, o erro absoluto percentual médio, a raiz do erro quadrático médio e a estatística U de Theil para avaliar a qualidade do modelo encontrado. Também foram construídos intervalos de confiança para a previsões. Os resultados obtidos para o setor residencial indicam o modelo de Vetores Auto-regressivos, que também considerou o PIB/RS, o preço da tarifa e o preço dos eletrodomésticos, como o de melhores previsões. Já no setor comercial, o modelo univariado de Box-Jenkins foi o que apresentou o melhor resultado para a previsão, o mesmo acontecendo no setor industrial. Na comparação entre os modelos estaduais e nacionais, encontrados pela combinação dos métodos de Vetores Autoregressivos e Correção de Erros Vetoriais, os modelos estaduais forneceram melhores previsões. Além disso, a relação entre as variáveis independentes consideradas nas duas regiões com a demanda sempre apresentaram o mesmo sentido.

Palavras-chave: Séries Temporais, Modelo Box-Jenkins, Modelo de Vetores Auto-regressivos, Modelo de Correção de Erros Vetoriais, Demanda de Energia Elétrica.

ABSTRACT

The electricity energy demand is shown as an important factor to the growth of a region economy. Aiming to assist in the planning of the energy sector of Rio Grande do Sul, this work addressed the demand in the three main consumer classes of electric power of this state: residential, commercial and industrial, using the Box-Jenkins methodologies, Vector Autoregressive and Vector Autoregressive in conjunction with Vector Error Correction, comparing the results obtained for the prediction by the different models, as well as the results of the joint application of vector autoregressive and vector error correction obtained in a recent study developed at national level. After the logarithmic transformation, we used correlograms, partial correlograms, ADF and KPSS unit root tests, information criteria of Akaike, Schwarz e Hannan-Quinn, evaluation of coefficient significance and residue analysis. The Ljung-Box, Jarque-Bera and ARCH tests were used to evaluate, respectively, autocorrelation, normality and homoscedasticity of the residuals, besides the determination of the mean absolute error, mean absolute percentage error, the root mean square error and statistical U the Theil to evaluate the quality of the model found. Confidence intervals were also built for predictions. The results obtained for the residential sector indicate the model of the vector autoregression, which also considered the PIB/RS, the price of the tariff and the price of home appliances, as the best forecasts. In the commercial sector, the univariate model of Box-Jenkins was the one that presented the best result for the forecast, as well as in the industrial sector. In the comparison between the state and national models, found by combining the methods of Vector Autoregression and Vector Error Correction, the state models provided better predictions. Moreover, the relation between the independent variables considered in the two regions with the demand always had the same meaning.

Keywords: Time Series, Model Box-Jenkins, Vector Autoregressive Model, Vector Error Correction Model, Electric Energy Demand.

Índice

1. Introdução	16
1.1. Tema e sua importância	16
1.2. Justificativa	19
1.3. Objetivos	20
1.3.1. Objetivos gerais	20
1.3.2. Objetivos específicos	20
2. Revisão bibliográfica	21
2.1. A produção de energia elétrica no Rio Grande do Sul: um breve histórico	21
2.2. Metodologia Box-Jenkins	24
2.2.1. Introdução	24
2.2.2. Identificação do modelo	25
2.3. Metodologia VAR-VEC	26
2.3.1. Modelo VAR	26
2.3.2. Modelo VAR-VEC	27
2.4. Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP)	28
2.5. Teste de Raiz Unitária	30
2.5.1. Dickey-Fuller (DF)	30
2.5.2. Dickey-Fuller Aumentado (ADF)	31
2.5.3. Teste de KPSS	32
2.6. Critérios de informação	33
2.6.1. Critério de informação de Akaike (CIA)	33
2.6.2. Critério de informação de Schwarz (CIS)	33
2.6.3. Critério de Hannan-Quinn (HQ)	34
2.7. Testes das suposições do modelo teórico	34
2.7.1. Teste de Ljung-Box	35
2.7.2. Teste de Jarque Bera (JB)	35

2.7.3. Teste ARCH	36
2.8. Análise dos erros	37
2.8.1. Erro absoluto médio (MAE)	38
2.8.2. Erro quadrático médio (RMSE)	38
2.8.3. Erro absoluto percentual médio (MAPE)	38
2.8.4. Estatística U de Theil	38
2.9. Literatura relacionada	38
3. Metodologia	45
3.1. Metodologia de Box-Jenkins	45
3.2. Metodologia VAR-VEC	48
4. Resultados	55
4.1. Residencial	55
4.1.1. Metodologia Box-Jenkins	55
4.1.2. Metodologia VAR-VEC	62
4.1.3. Comparações para o setor Residencial	72
4.2. Comercial	75
4.2.1. Metodologia Box-Jenkins	75
4.2.2. Metodologia VAR-VEC	82
4.2.3. Comparações para o setor Comercial	94
4.3. Industrial	96
4.3.1. Metodologia Box-Jenkins	96
4.3.2. Metodologia VAR-VEC	104
4.3.3. Comparações para o setor Industrial	119
5. Considerações Finais	121
Referências	123
APÊNDICE 01 – Código utilizado na verificação dos modelos e da geração da estatística U de Theil	128

LISTA DE FIGURAS

Figura 1.1. Relação entre taxas de crescimento anual do PIB e da demanda de eletricidade com projeção para 2026	16
Figura 1.2. Relação entre demanda de energia elétrica e PIB per capita com projeção para 2026	17
Figura 1.3. Evolução da demanda de energia elétrica no Brasil (a) e no Rio Grande do Sul (b)	18
Figura 3.1 -Fluxograma da metodologia Box Jenkins.....	47
Figura 3.2 - Fluxograma da metodologia VAR-VEC	54
Figura 4.1-Demanda residencial de energia elétrica no Rio Grande do Sul no período 1971 - 2010.....	55
Figura 4.2 - Gráfico do logaritmo da demanda residencial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)	56
Figura 4.3 - Gráfico da primeira diferença do logaritmo da demanda residencial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)	57
Figura 4.4 – Gráfico do logaritmo da demanda residencial de energia elétrica/RS com modelo de previsão ARIMA (1,1,1).....	61
Figura 4.5 – Gráfico da série em nível da demanda residencial de energia elétrica/RS com modelo de previsão ARIMA (1,1,1)	61
Figura 4.6 – Comportamento das séries temporais deflacionadas ligadas à demanda residencial.....	62
Figura 4.7 – Gráfico do modelo de previsão VAR para demanda residencial de energia elétrica/RS	71
Figura 4.8 – Gráfico do modelo de previsão VAR-VEC para demanda residencial de energia elétrica/RS.....	72
Figura 4.9 – Demanda comercial de energia elétrica no Rio Grande do Sul no período 1971 - 2010.....	75
Figura 4.10 – Gráfico do logaritmo da demanda comercial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)	76

Figura 4.11 – Gráfico da primeira diferença do logaritmo da demanda comercial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)	77
Figura 4.12 – Gráfico do logaritmo da demanda comercial de energia elétrica/RS com modelo de previsão ARIMA (2,1,0)	81
Figura 4.13 – Gráfico da demanda comercial de energia elétrica/RS com modelo de previsão ARIMA (2,1,0)	82
Figura 4.14 – Comportamento das séries temporais deflacionadas ligadas à demanda comercial em relação ao primeiro valor observado em escala para 100	83
Figura 4.15 – Gráfico do modelo de previsão VAR para demanda comercial de energia elétrica/RS	93
Figura 4.16 – Gráfico do modelo de previsão VAR para demanda comercial de energia elétrica/RS	93
Figura 4.17 – Demanda industrial de energia elétrica no Rio Grande do Sul no período 1961 - 2010	96
Figura 4.18 – Gráfico do logaritmo da demanda industrial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)	97
Figura 4.19 – Gráfico da primeira diferença do logaritmo da demanda industrial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)	98
Figura 4.20 – Gráfico do logaritmo da demanda industrial de energia elétrica/RS com modelo de previsão ARIMA (1,1,2)	103
Figura 4.21 – Gráfico da demanda industrial de energia elétrica/RS com modelo de previsão ARIMA (1,1,2)	103
Figura 4.22 – Comportamento das séries temporais deflacionadas ligadas à demanda industrial em relação ao primeiro valor observado em escala para 100	104
Figura 4.23 – Gráfico do modelo de previsão VAR para demanda industrial de energia elétrica/RS	118
Figura 4.24 – Gráfico do modelo de previsão VAR-VEC para demanda industrial de energia elétrica/RS	118

LISTA DE TABELAS

Tabela 3.1 – Dados e fontes.....	49
Tabela 4.1 – Significância dos coeficientes dos modelos selecionados	58
Tabela 4.2 – Resultados dos critérios de informação nos modelos candidatos ARIMA (p,1,q)	58
Tabela 4.3 – Resultados dos testes de Ljung-Box, Jarque-Bera e ARCH para os resíduos dos modelos candidatos ARIMA (p,1,q)	59
Tabela 4.4 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão.....	59
Tabela 4.5 – Intervalos de confiança para o modelo ARIMA (1,1,1) de previsão pela métrica logaritmica e original.....	60
Tabela 4.6 – Teste ADF no setor residencial	63
Tabela 4.7 – Teste KPSS no setor residencial	63
Tabela 4.8 – Seleção de defasagens do modelo VAR	64
Tabela 4.9 – Teste do traço de Johansen para o modelo com Dr, Rr, Tr, Er e constante sem restrição	66
Tabela 4.10 – Teste do máximo autovalor de Johansen para o modelo Dr, Rr, Tr, Er com constante sem restrição	67
Tabela 4.11 – Vetor de cointegração normalizado - Residencial	67
Tabela 4.12 – Testes de qualidade do modelo VAR para o setor residencial	68
Tabela 4.13 – Testes das suposições do modelo VEC para o setor residencial	69
Tabela 4.14 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos VAR e VARVEC de previsão.....	70
Tabela 4.15 – Intervalos de confiança para o modelo VAR de previsão pela métrica logaritmica e original.....	70
Tabela 4.16 – Intervalos de confiança para o modelo VAR-VEC de previsão pela métrica logaritmica e original.....	71
Tabela 4.17 – Coeficientes do modelo VAR Residencial/RS e BR	73
Tabela 4.18 – Coeficientes do modelo VAR-VEC Residencial/RS e BR.....	73
Tabela 4.19 – Significância dos coeficientes dos modelos selecionados	78

Tabela 4.20 – Resultados dos critérios de informação nos modelos candidatos para modelos ARIMA (p,1,q).....	78
Tabela 4.21 – Resultados dos testes de Ljung-Box, Jarque-Bera e ARCH para os resíduos dos modelos candidatos ARIMA (p,1,q).....	79
Tabela 4.22 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão.....	80
Tabela 4.23 – Intervalos de confiança para o modelo ARIMA (2,1,0) de previsão pela métrica logaritmica e original.....	80
Tabela 4.24 – Teste ADF no setor comercial.....	84
Tabela 4.25 – Teste KPSS no setor comercial.....	84
Tabela 4.26 – Seleção de defasagens do modelo VAR.....	85
Tabela 4.27 – Teste do traço de Johansen para o modelo com Dc, Rc, Tc, Ec, DYc3 e DYc6 com tendência sem restrição.....	87
Tabela 4.28 – Teste do máximo autovalor de Johansen para o modelo Dc, Rc, Tc, Ec, DYc3 e DYc6 com tendência sem restrição.....	88
Tabela 4.29 – Vetor de cointegração normalizado para o setor comercial.....	88
Tabela 4.30 – Testes de qualidade do modelo VAR para o setor comercial.....	89
Tabela 4.31 – Testes das suposições teóricas modelo VAR-VEC para o setor comercial.....	90
Tabela 4.32 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão.....	91
Tabela 4.33 – Intervalos de confiança para o modelo VAR de previsão pela métrica logaritmica e original.....	91
Tabela 4.34 – Intervalos de confiança para o modelo VAR-VEC de previsão pela métrica logaritmica e original.....	92
Tabela 4.35 – Coeficientes do modelo VAR Comercial/RS e BR.....	95
Tabela 4.36 – Coeficientes do modelo VAR-VEC Comercial/RS e BR.....	95
Tabela 4.37 – Significância dos coeficientes dos modelos selecionados.....	99
Tabela 4.38 – Resultados dos critérios de informação nos modelos candidatos para modelos ARIMA (p,1,q).....	100

Tabela 4.39 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão.....	101
Tabela 4.40 – Intervalos de confiança para o modelo ARIMA (1,1,2) de previsão pela métrica logaritmica e original.....	102
Tabela 4.41 – Teste ADF no setor industrial.....	105
Tabela 4.42 – Teste KPSS no setor industrial	105
Tabela 4.43 – Seleção de defasagens do modelo VAR	106
Tabela 4.44 – Teste do traço de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3 e tendência restrita.....	108
Tabela 4.45 – Teste do máximo autovalor de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3 e tendência restrita.....	109
Tabela 4.46 – Vetor de cointegração normalizado para o setor Industrial.....	109
Tabela 4.47 – Seleção de defasagens do modelo VAR	111
Tabela 4.48 – Teste do traço de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3 e sem constante	113
Tabela 4.49 – Teste do máximo autovalor de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3, sem constante	113
Tabela 4.50 – Vetor de cointegração normalizado para o setor Industrial	114
Tabela 4.51 – Resultados da avaliação das suposições do modelo VAR para o setor Industrial.....	114
Tabela 4.52 – Testes de qualidade do modelo VAR-VEC para o setor industrial	115
Tabela 4.53 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão.....	116
Tabela 4.54 – Intervalos de confiança para o modelo VAR de previsão pela métrica logaritmica e original.....	116
Tabela 4.55 – Intervalos de confiança para o modelo VAR-VEC de previsão pela métrica logaritmica e original.....	117
Tabela 4.56 – Coeficientes do modelo VAR Industrial/RS e BR	120
Tabela 4.57 – Coeficientes do modelo VAR-VEC Industrial/RS e BR	120

LISTA DE SÍMBOLOS

D	demanda de energia elétrica
P	produção total de um bem
L	quantidade de trabalho
Z	quantidade de capital investido
b	intercepto da função de produção
k	constante positiva
T	Tarifa
R	Renda
E	preço de eletrodomésticos/eletrointensivos
S	preço de bens substitutos à energia elétrica
I	ordem de integração
ut	Variável explicativa do termo de correção de erros
r	número de vetores de cointegração
y	série a ser modelada
t	instante de tempo
d	ordem de integração
DY	variável <i>dummy</i>
n	número de variáveis
p	Coefficiente auto-regressivo
q	Coefficiente de médias móveis

Símbolos gregos

Δ	variação/diferença
ν	elasticidade trabalho-produção
Γ	elasticidade-preço em relação à tarifa
θ	elasticidade-renda

A	elasticidade-preço em relação aos eletrodomésticos
B	elasticidade-preço em relação ao bem substituto à energia elétrica
ε	vetor de erros
$\hat{\lambda}$	valores estimados das raízes características
ρ	coeficiente Y
τ	estatística de teste

Sub índices

t	Tempo
r	Residencial
c	Comercial
i	Industrial

LISTA DE ABREVIATURAS

VECM	Modelo vetor de correção de erros
VAR	vetores auto-regressivos
ADF	Dickey-Fuller Aumentado
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
IPA	Índice de Preços por Atacado
IGP	Índice Geral de Preços
FGV	Fundação Getúlio Vargas
EPE	Empresa de Pesquisa Energética
MQO	Mínimos Quadrados Ordinários
CEEE	Companhia Estadual de Energia Elétrica
RGE	Rio Grande Energia S/A
AES	Applied Energy Services
CGTEE	Companhia de Geração Térmica de Energia Elétrica
ANEEL	Agência Nacional de Energia Elétrica
PIB	Produto Interno Bruto
IPEA	Instituto de Pesquisa Econômica Aplicada
II PND	II Plano Nacional de Desenvolvimento
VBC	Votorantim, Bradesco e Camargo Correa
AR	auto-regressivos
MA	médias móveis
ARMA	auto-regressivos e de médias móveis
ARIMA	auto-regressivos integrados e de médias móveis
SARIMA	auto-regressivos integrados e de médias móveis sazonal
FAC	função de autocorrelação
FACP	função de autocorrelação parcial

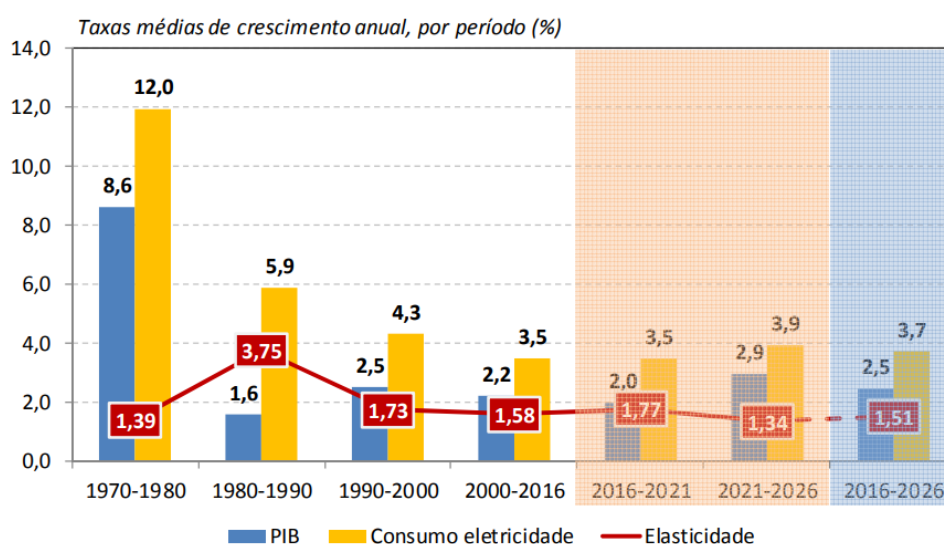
MAE	erro absoluto médio
MAPE	erro absoluto percentual médio
RMSE	erro quadrático médio
ARCH	Autoregressive Conditional Heteroskedasticity
Gretl	Gnu Regression, Econometrics and Times-Séries Library
BR	Brasil
RS	Rio Grande do Sul
CIS	Critério de Informação Schwarz
CIA	Critério de Informação Akaike
HQ	Critério de Informação Hannan Quinn

1. INTRODUÇÃO

1.1. Tema e sua importância

De acordo com ANEEL (2008), na questão do desenvolvimento econômico, a energia elétrica apresenta-se como uma das mais versáteis e apropriadas fontes de energia, estando diretamente ligada aos índices de desenvolvimento socioeconômico de diversas nações. Considera-se que ela é imprescindível para a sobrevivência da humanidade, assim como para a qualidade de vida do homem moderno, fazendo-se presente nas mais variadas áreas da sociedade, tais como saúde, educação, lazer, cultura e economia, entre outras. Desta forma, é difícil imaginar a humanidade sem a presença da energia elétrica.

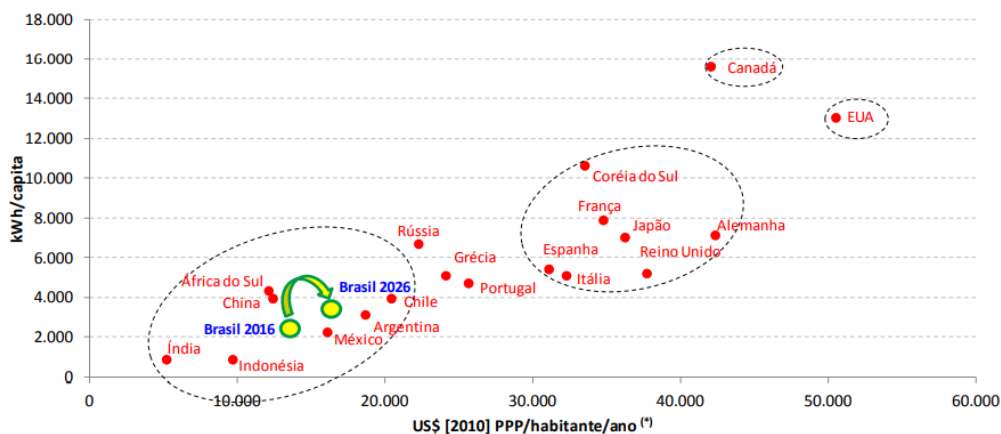
Dados da Empresa de Pesquisa em Energética (EPE, 2017b) mostram que as taxas de crescimento do produto interno bruto (PIB) estão relacionadas com o crescimento da demanda, conforme pode ser observado na Fig. 1.1. Na década 1970-1980, as taxas médias de crescimento anual de ambas as variáveis assumiram os valores máximos (8,6% e 12%), embora o mesmo não acontecendo com os valores mínimos: a menor taxa média de crescimento do PIB ocorreu na década 1980-1990, enquanto a menor taxa média de demanda de energia elétrica ocorreu no período 2000-2021.



Fonte: (EPE, 2017b).

Figura 1.1. Relação entre taxas de crescimento anual do PIB e da demanda de eletricidade com projeção para 2026

A relação entre o PIB e a demanda de energia elétrica per capita, para diversos países pode ser observada na Fig. 1.2, em que é possível constatar que países mais desenvolvidos apresentam maior demanda de energia elétrica per capita, se comparados a países em desenvolvimento. O Brasil, por exemplo, apresenta demanda de energia per capita muito baixo se comparado aos países desenvolvidos, como Canadá e EUA, e um pouco mais baixo que alguns países vizinhos, como Chile e Argentina.



(*) PIB per capita referenciado a US\$ [2010] PPP (*Power Purchase Parity*). Os dados são relativos ao ano de 2014 para todos os países com exceção do Brasil.

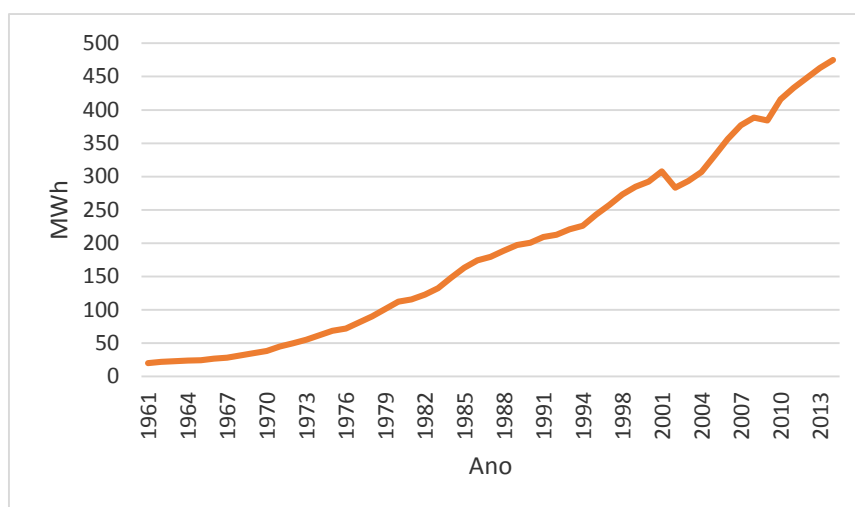
Nota: considera o consumo total de eletricidade, incluindo a autoprodução.

Fonte: (EPE, 2017b).

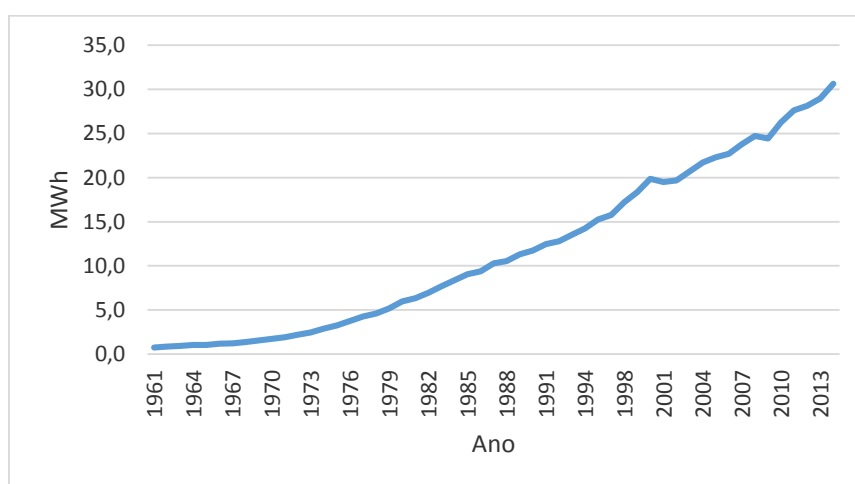
Figura 1.2. Relação entre demanda de energia elétrica e PIB per capita com projeção para 2026

Diante dessas considerações, fica evidente que o país necessita de investimento em sua matriz energética para seguir em sua busca por desenvolvimento econômico. No Brasil, a crise energética ocorrida no início dos anos 2000 foi motivada pelas mudanças realizadas na década anterior e, juntamente com fatores climáticos pouco favoráveis, fizeram com que o país se visse em situação de emergência. Um período de baixa precipitação de chuvas resultou na baixa do nível dos reservatórios das hidrelétricas, principal fonte geradora de eletricidade da matriz energética nacional. Visando evitar um colapso na oferta de energia, o governo teve que adotar medidas duras, o que resultou em um período de racionamento. Devido a essa crise energética, o governo aumentou os investimentos na diversificação da matriz energética nacional, preferindo formas de geração que não fossem diretamente dependentes do clima. Dessa maneira, as termoelétricas passaram a receber maior destaque, principalmente aquelas que operam com biomassa e gás natural como combustíveis (LIMA, 2011).

As evoluções das demandas de energia elétrica no Brasil e no Rio Grande do Sul podem ser observadas na Fig. 1.3.a e na Fig. 1.3.b, respectivamente, em que pode-se notar, em ambos os casos, que entre os anos de 2001 e 2002 há uma alteração na série do demanda. Essa alteração pode ser justificada pelo racionamento ocorrido nesses anos. Além disso, também se observa na Fig. 1.3.a que, no início dos anos noventa, o crescimento foi lento devido à crise econômica vivida pelo país, o que não se evidencia na Fig. 1.3.b. Durante o período que compreendeu a crise econômica mundial de 2008 e 2009, é possível notar outra oscilação nas Fig. 1.3a e na Fig 1.3b, respectivamente.



(a)



(b)

Fonte: Elaborado a partir de Eletrobrás (2018).

Figura 1.3. Evolução da demanda de energia elétrica no Brasil (a) e no Rio Grande do Sul (b)

1.2. Justificativa

A economia brasileira, assim como a gaúcha, está passando por um momento difícil, em que há desconfiança dos agentes de investimento, com rebaixamentos consecutivos de sua nota de crédito, além de inflação e tentativas de ajustes fiscais. Segundo EPE (2017a), a esperança de retomada na economia vem sendo postergada em virtude da divulgação de indicadores conjunturais com resultados inferiores ao esperado. A demanda familiar, que foi um dos balizadores do crescimento nos últimos anos, no curto prazo, não deve cooperar da mesma forma para o crescimento econômico, dado o fraco desempenho do mercado de trabalho e de crédito. Pelo lado da oferta, há um elevado grau de ociosidade dificultando a retomada dos investimentos. Desta maneira, para que o Rio Grande do Sul retorne à trajetória de um crescimento, assim como o Brasil, serão indispensáveis esforços, como por exemplo, investimentos em infraestrutura, inovação, educação e qualificação profissional, bem como avanços no mercado de negócios.

Salienta-se que o crescimento econômico tem um relevante papel no crescimento da demanda de energia e que o aumento do PIB brasileiro nos últimos anos tem sido acompanhado pelo aumento dos níveis de demanda, assim como de petróleo e gás natural em todo o país, incluindo o Rio Grande do Sul. Considerando que a energia elétrica constitui um importante segmento da infraestrutura de um país e que esse produto não pode ser estocado, sua produção adequada à demanda representa um fator relevante para a conservação do seu crescimento econômico. Além disso, o aumento de sua produção não pode ser feito rapidamente, pois suas bases geradoras levam alguns anos para serem construídas e colocadas em funcionamento.

Por outro lado, a crise do setor elétrico, ocorrida no país no início dos anos 2000, alertou para problemas de continuidade no fornecimento de energia e para a importância do aumento da capacidade de geração, transmissão e distribuição. De acordo com Viana e Silva (2014), no Brasil, há uma preocupação em aumentar suas bases geradoras de energia, o que é demonstrado pela realização de fóruns nacionais, compostos por agentes ligados à questão energética, em que são debatidas medidas e metas para o setor.

Nesse sentido, a comunidade acadêmica também tem colaborado com o setor, pois já existem estudos sobre o tema no cenário nacional, bem como para algumas regiões, tais como: Modiano (1984), Andrade e Lobão (1997), Schmidt e Lima (2004), Mattos e Lima (2005), Leite (2006), Viana e Silva (2014) e Silveira (2017), entre outros.

Entretanto, existe uma considerável diferença entre as elasticidades encontradas

nos estudos publicados sobre a demanda de energia elétrica no país. Estas diferenças podem ser explicadas pelas variáveis selecionadas, pelas proxies escolhidas para cada variável, pelo método de estimação dos parâmetros da função, assim como pelo número de elementos das séries históricas utilizadas em cada estudo. Além disso, não foram encontrados estudos específicos desta variável para o Rio Grande do Sul, o que justifica a realização deste estudo.

1.3. Objetivos

1.3.1. Objetivos gerais

Este trabalho tem o objetivo de modelar a demanda de energia elétrica ocorrida no Rio Grande do Sul nas últimas décadas, para cada uma das três principais classes consumidoras: residencial, comercial e industrial.

1.3.2. Objetivos específicos

1. comparar a qualidade do modelo encontrado pelo método VAR-VEC para o Rio Grande do Sul com o obtido para o Brasil por este mesmo método.
2. avaliar a necessidade de regionalização dos modelos de previsão da demanda de energia elétrica, visando um melhor resultado para cada setor;
3. realizar previsões para a demanda de energia elétrica no estado do Rio Grande do Sul para os próximos anos em cada setor;
4. avaliar o desempenho dos métodos utilizados na modelagem da demanda de energia elétrica neste estudo.

2. REVISÃO BIBLIOGRÁFICA

2.1. A produção de energia elétrica no Rio Grande do Sul: um breve

histórico

No século XIX, a principal fonte de renda do Brasil provinha majoritariamente da cultura de café para a exportação. A valorização deste produto no mercado externo e as condições internas favoráveis estimularam o aumento da sua produção, passando a aumentar o interesse em novos investimentos. Desta forma, o Brasil seguiu em direção à modernização e, com o crescimento das cidades, a indústria da construção civil expandiu, melhorando a infraestrutura urbana. Foi justamente neste período que começaram as primeiras iniciativas do uso da energia elétrica no país. Entretanto, somente em 1879 foi inaugurado o primeiro serviço permanente de iluminação elétrica do Brasil (GOMES et al., 2002).

Em 1887, a energia elétrica começou a ser utilizada no Rio Grande do Sul, por meio da usina elétrica da "Sociedade Fiat Lux", após seis meses de experimentos. A usina era constituída por uma máquina a vapor semifixa de 50 cv (cavalo-vapor), de fabricação inglesa. Essa máquina acionava três dínamos (geradores de corrente contínua) de fabricação alemã, podendo fornecer energia para acender 800 lâmpadas de 10 velas (aproximadamente 10,19 candelas no Sistema Internacional) do sistema Edison (similar as lâmpadas de filamentos ou incandescentes). Dessa usina partiam dois circuitos de distribuição. Em 1898, em Santa Maria, foi inaugurada a primeira usina de energia elétrica do interior do estado (ABID, 2007).

Em 1943, por meio do Decreto-Lei Estadual nº 328, foi criada a Comissão Estadual de Energia Elétrica - CEEE, então subordinada à Secretaria de Estado dos Negócios das Obras Públicas, com o objetivo de prever e sistematizar para todo o Estado, em plano geral, o aproveitamento de seus potenciais hidráulicos em conexão com suas reservas carboníferas. Em 1945, foi lançada a programação de obras de eletrificação, constituída de pequenas usinas hidrelétricas. Em 1947, a CEEE passou a ser subordinada diretamente ao governo do estado do Rio Grande do Sul (CEEE, 2017).

No século XX, entre as décadas de 50 e 70, o Brasil investiu massivamente no setor de infraestrutura, o que ocasionou crescimento do setor energético do país. Durante o governo Kubitschek, na segunda metade da década de 50, ocorreu um investimento

estatal nos setores de transporte e energia, o que ligado à facilidade na aquisição de recursos financeiros internacionais e à disponibilidade de recursos hidrelétricos próximos aos mercados consumidores, possibilitou a livre expansão do setor de energia elétrica (IRFFI et al., 2009). Na década de 60, a partir da criação da Eletrobrás, o potencial elétrico brasileiro foi melhor aproveitado com a conexão dos mercados de produção, de distribuição e de demanda. Com o aumento da capacidade de geração e ampliação da distribuição, a demanda de energia elétrica se expandiu ao longo dos anos de 1960, 1970 e 1980 (IRFFI et al., 2009).

Em 1961, na cidade de Bagé, na localidade de Candiota (emancipada em 1992), entra em operação a usina termoelétrica de Candiota 1, que utiliza carvão mineral como combustível. A potência instalada nessa usina era de 20 MW (ABID, 2007). Em setembro do mesmo ano, por meio da Lei Estadual nº 4.136, o Governo do Estado do Rio Grande do Sul é autorizado a promover a organização de uma sociedade por ações, a ser denominada Companhia Estadual de Energia Elétrica e destinada a projetar, construir e explorar sistemas de produção, transmissão e distribuição de energia elétrica. Em 19/12/1963, em conformidade com a Lei Estadual nº 4.136, de 13/09/1961, a antiga Comissão Estadual de Energia Elétrica foi transformada em Sociedade de Economia Mista, com a designação de Companhia Estadual de Energia Elétrica, conservando a sigla CEEE (CEEE, 2017).

Durante a década de 70, com o ambiente econômico favorável, foi lançado o II Plano Nacional de Desenvolvimento (IIPND) que buscava possibilitar a produção dos principais insumos básicos e a geração de encomendas de equipamentos às indústrias locais. Com isso foram elaborados os projetos de Tucuruí, de Itaipu e da Ferrovia do Aço, além do programa nuclear (DE LORENZO, 2002).

Em 1974, novamente na localidade de Candiota/Bagé, entra em operação a Fase A da usina termoelétrica Presidente Médici. Utilizando carvão mineral como combustível, possuía uma capacidade instalada inicial de 126 MW. Em dezembro de 1986, entrou em operação a Fase B, e em janeiro de 2010, a Fase C. Em julho de 1997, a usina passou a fazer parte da Companhia de Geração Térmica de Energia Elétrica (CGTEE), do grupo ELETROBRAS (ABID, 2007).

Nesta década de 90, foram realizadas licitações para novos empreendimentos de geração e regulamentação da produção de energia elétrica por produtos independentes e por autoprodutor, aumentando a liberdade dos grandes consumidores em escolher em que iriam suprir a sua demanda de energia. Em seguida, com a privatização de parte do setor,

as áreas de geração, transmissão e distribuição passaram a funcionar como negócios independentes. Cabe ressaltar que a geração e a comercialização foram desreguladas com a finalidade de estimular a competição, e a distribuição e transmissão continuaram como serviços de utilidade pública regulados devido ao fato de serem monopólios naturais. Diante disso, em 1996 o governo criou a Agência Nacional de Energia Elétrica (ANEEL) para regular o setor (LIMA, 2011).

Em janeiro de 1996, é sancionada a Lei Estadual nº 10.681, que dá nova redação ao artigo 4º da Lei 4.136, de 13/09/1961, permitindo a alienação de até 49% do Capital Social da Companhia a pessoas jurídicas de direito público ou privado e, ainda, a pessoas físicas. Desta maneira, em 21/10/1997, por meio de leilão público, ocorreu a alienação das Companhias Centro-Oeste de Distribuição de Energia Elétrica e Norte-Nordeste de Distribuição de Energia Elétrica. A Centro-Oeste foi vendida à AES (Applied Energy Services) Guaíba Empreendimentos e a Norte-Nordeste foi adquirida pelo consórcio formado pela VBC (Votorantim, Bradesco e Camargo Correa), Previ (Fundo de Pensão dos Funcionários do Banco do Brasil) e Community Energy Alternatives. Em dezembro deste mesmo ano, a Centro-Oeste alterou sua razão social para AES Sul Distribuidora Gaúcha de Energia S/A – AES Sul e a Norte-Nordeste passou à denominação de Rio Grande Energia S/A – RGE. Em dezembro de 1998, o controle da Companhia de Geração Térmica de Energia Elétrica (CGTEE) passou para o Governo Federal, em troca de dívidas do Governo Estadual para com o Governo Federal. A CEEE permaneceu com as concessões para as atividades de geração hidrelétrica, transmissão de energia elétrica no Rio Grande do Sul e a distribuição de eletricidade na região sul e sudeste do Estado (CEEE, 2017).

Até a década de 90, existia uma tarifa única de energia elétrica no Brasil, que garantia a remuneração das concessionárias, independentemente de seu nível de eficiência. Esse sistema não incentivava a busca pela eficiência por parte da distribuidora, uma vez que a integralidade de seu custo era transferida ao consumidor. Em 1993, com a edição da Lei nº 8.631, as tarifas passaram a ser fixadas por empresa, conforme características específicas de cada área de concessão – por exemplo, número de consumidores, quilômetros de rede de transmissão e distribuição, tamanho do mercado (representado pela quantidade de unidades de demanda atendidas por uma determinada infraestrutura), custo da energia comprada e tributos estaduais, entre outros. Portanto, se essa área coincide com a de uma unidade federativa, a tarifa é única naquele estado. Caso contrário, tarifas diferentes coexistem dentro do mesmo estado (ANEEL, 2008).

No Rio Grande do Sul, a capacidade de geração de energia elétrica nos últimos anos cresceu consideravelmente passando de 1.893,8MW em 1998 para 9.565,8MW em 2016. Este aumento permitiu a ampliação dos níveis de demanda, não só no estado, como também no país, pois o parque gerador opera por meio do Sistema Interligado Nacional de Geração e Transmissão. De acordo com a ANEEL, a capacidade instalada do Estado representa cerca de 6% da capacidade instalada total do Brasil (Secretaria de Planejamento, Governança e Gestão, 2017).

Quanto ao perfil da demanda de energia elétrica no RS, a indústria é o setor de maior demanda com 33,0% da energia produzida, o setor residencial consome 27,8%, o setor comercial, 16,8%, o setor agropecuário, 14,2%, e o restante de 8,2% é consumido pelo setor público, demanda própria e de transporte (EPE, 2018).

Analisando o setor residencial dos trabalhos citados, é possível notar que as últimas publicações têm utilizado o PIB per capita, Tarifa média e IPA-OG: eletrodomésticos como proxies para renda, tarifa e preço de eletrodomésticos, respectivamente. No setor comercial os trabalhos publicados nos últimos anos usaram como proxies para renda, tarifa e preço de eletrointensivos respectivamente o PIB serviços, tarifa média e IPA-OG: Material Elétrico. E no setor industrial, foi possível notar que não houve muita variação na escolha das proxies, sendo o PIB: indústria para a renda, tarifa média para a variável tarifa e para o preço as proxies foram o IPA-OG: Maq. e Equipamentos Industriais, e IPA-OG: Combustíveis e Lubrificantes.

2.2. Metodologia Box-Jenkins

2.2.1. Introdução

A metodologia Box-Jenkins (ARIMA) nada mais é do que um processo de análise de séries temporais. Lembrando que quando se fala em séries temporais, entende-se como uma sequência de observações ordenadas no tempo.

Segundo Gujarati e Porter (2011), os modelos ARIMA univariados, diferentes de outras regressões, não são influenciados por outras variáveis, mas sim, pelos valores defasados de suas próprias séries por meio de uma correlação temporal.

Assim, esta metodologia proporciona uma ampla aplicação em modelagens univariadas de séries temporais, apresentando características de modelos auto-regressivos (AR), modelos de médias móveis (MA), auto-regressivos e de médias móveis (ARMA),

auto-regressivos integrados e de médias móveis (ARIMA) e modelos ARIMA com presença de sazonalidade (SARIMA). Esta técnica necessita seguir algumas suposições como: homocedasticidade, ausência de autocorrelação serial e normalidade. Este método pode ser utilizado, assim como várias outras técnicas, em variáveis de natureza econômica, biológica, física, epidemiológica e a diversas outras áreas do conhecimento, como em NUNES (2016).

2.2.2. Identificação do modelo

A identificação dos modelos $ARIMA(p, d, q)$ é baseada na determinação das ordens de três componentes, também denominados “filtros”: a ordem p , associada aos parâmetros do filtro AR; a ordem q , associada aos parâmetros do filtro MA; e ordem d , associada ao número de diferenciações necessárias para tornar a série estacionária (MORETTIN, 2006).

Vale ressaltar que, nem todas as séries apresentam esses três filtros. Há também a possibilidade da existência de apenas parte desses filtros, assim sendo, a metodologia de Box-Jenkins resulta em diferentes modelos passíveis de análise (VASCONCELLOS; ALVES, 2000). Por esse motivo, a identificação de modelos ARIMA deve ser realizada em dois estágios: primeiro deve-se verificar a estacionariedade da série para que seja possível determinar o número de diferenciações da ordem $I(d)$, e posteriormente, as ordens p e q do modelo.

O modelo utilizado será caracterizado como um $ARIMA(p, d, q)$. A ordem de diferenciação (d), é encontrada através da análise do comportamento da série em termos de estacionariedade, verificando em nível, em primeira diferença, segunda diferença e assim por diante, até que a série temporal seja estacionária. Complementa-se a análise com o teste de Dickey-Fulley Aumentado (ADF), que considera como hipótese nula a presença de raiz unitária, sugerindo que a série é não estacionária, e o teste de Kwiatkowski-Phillips-Schmidt-Shin (KPSS), o qual considera como hipótese nula a ausência de raiz unitária, indicando a estacionariedade da série. A representação gráfica da função de autocorrelação e da função de autocorrelação parcial permite estimar possíveis valores para p e q . Aplicam-se várias combinações destes parâmetros, juntamente com o valor da diferenciação, sendo escolhido o modelo mais parcimonioso, ou seja, aquele que envolve o mínimo de parâmetros possíveis a serem estimados, explicando satisfatoriamente o comportamento da variável resposta, o que é feito pelos de Critério

de Akaike (1973), Critério de Schwarz (1978), Critério de Hannan-Quinn (1979). Utiliza-se o modelo que apresentar o menor valor nesses critérios.

A fórmula geral utilizada na definição do modelo ARMA (p, q) não sazonal é a representada na Eq. 2.1 (BUENO, 2011):

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=0}^q \theta_j \varepsilon_{t-j} , \quad (2.1)$$

Onde: y_t é o t -ésimo valor observado ($t = 1, 2, \dots, T$); c é uma constante; i é a defasagem do processo auto-regressivo ($i = 1, 2, \dots, p$); j é a defasagem do processo de médias móveis ($j = 1, 2, \dots, q$); ϕ_i é o coeficiente da parcela auto-regressiva na i -ésima defasagem; θ_j é o coeficiente de médias móveis na j -ésima defasagem; p é a ordem do processo auto-regressivo; q é a ordem do processo de médias móveis; ε_t é o erro no t -ésimo tempo.

2.3. Metodologia VAR-VEC

De acordo com Gujarati e Porter (2011), em um modelo de vetores autorregressivos (VAR) o conjunto de variáveis deverá possuir simultaneidade verdadeira, em que todas as variáveis devem ser tratadas em pé de igualdade, não havendo qualquer distinção a priori entre as variáveis endógenas e exógenas. O termo autorregressivo se deve à aparência do valor defasado da variável dependente no lado direito e o termo vetor se deve ao fato de se estar trabalhando com um vetor de duas (ou mais) variáveis.

2.3.1. Modelo VAR

De modo geral, pode-se expressar um modelo autorregressivo de ordem p por um vetor com n variáveis endógenas, X_t , conectadas entre si por meio de uma matriz A , dado pela Eq. 2.2 (BUENO, 2011):

$$AX_t = B_0 + \sum_{i=1}^p B_i X_{t-i} + B \varepsilon_t, \quad (2.2)$$

em que: A é uma matriz $n \times n$ que define as restrições contemporâneas entre as variáveis que constituem o vetor $n \times 1$, X_t ; B_0 é um vetor de constantes $n \times 1$; B_j são matrizes $n \times n$;

B é uma matriz diagonal $n \times n$ de desvios-padrão; $\mathcal{E}t$ é um vetor $n \times 1$ de perturbações aleatórias não correlacionadas entre si contemporânea ou temporalmente, isto é: $\varepsilon_t \sim i.i.d. (0; I_n)$.

A Equação 2.2 expressa as relações entre as variáveis endógenas, frequentemente decorrentes de um modelo econômico teoricamente estruturado, e por esse motivo chama-se forma estrutural. Os choques $\mathcal{E}t$ são denominados choques estruturais porque afetam individualmente cada uma das variáveis endógenas. Os choques estruturais são considerados independentes entre si porque as inter-relações entre um choque e outro são captadas indiretamente pela matriz A . Logo, a independência dos choques se dá sem perda de generalidade (Bueno, 2011).

Este modelo é comumente descrito em sua forma reduzida devido a endogeneidade das variáveis, assumindo a forma da Eq 2.3:

$$X_t = A^{-1}B_0 + \sum_{i=1}^p A^{-1}B_i X_{t-i} + A^{-1}B_{\varepsilon t} = \phi_0 + \sum_{i=1}^p \phi_i X_{t-i} + e_t, \quad (2.3)$$

em que: $\phi_i \equiv A^{-1}B_i$, $i = 0, 1, \dots, p$, e $B_{\varepsilon t} \equiv A e_t$.

2.3.2. Modelo VAR-VEC

Segundo Engle e Granger (1987), com a intenção de guiar as variáveis no curto prazo aplica-se o Modelo de Correção de Erros. Este modelo busca ligar o comportamento de curto prazo das variáveis ao seu comportamento de longo prazo. O modelo de correção de erros para n variáveis pode ser expresso pela Eq. 2.4:

$$\Delta X_t = \pi_0 + \pi X_{t-1} + \pi_1 \Delta X_{t-1} + \pi_2 \Delta X_{t-2} + \pi_3 \Delta X_{t-3} + \dots + \pi_p \Delta X_{t-p} + \varepsilon_t, \quad (2.4)$$

em que: π_0 é um vetor ($n \times 1$) que representa o intercepto com elementos π_{i0} ; π_i é uma matriz ($n \times n$) de coeficientes, com elementos $\pi_{jk}(i)$; π é uma matriz com elementos π_{jk} tal que um ou mais elementos $\pi_{jk} \neq 0$; e ε_t é um vetor ($n \times 1$) com elementos ε_{it} .

Nota-se que na hipótese de todos os elementos de π serem zero, tem-se um modelo VAR em primeira diferença. Deste modo, não há um mecanismo de correção de erros. Porém, se um ou mais valores de π forem diferentes de zero, ΔX responde aos desvios em relação ao equilíbrio de longo prazo, representando o mecanismo de correção de erros.

Desta maneira é possível realizar as previsões da demanda de energia em um curto prazo.

2.4. Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP)

Assim como a correlação mede a extensão de uma relação linear entre duas variáveis, a autocorrelação mede a relação linear entre os valores defasados de uma série temporal. Existem vários coeficientes de autocorrelação, correspondentes a cada painel no gráfico de latência. Por exemplo, ρ_1 mede a relação entre y_t e y_{t-1} , ρ_2 mede a relação entre y_t e y_{t-2} , e assim por diante (Hyndman e Athanasopoulos, 2018).

De acordo com Gujarati e Porter (2011), a função de correlação pode ser usada como base para um teste simples de estacionariedade. A FAC com defasagem k , denotada por ρ_k , é definida pela Eq. 2.5.

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covariância com defasagem } k}{\text{variância}} \quad (2.5)$$

É possível notar que se $k = 0$, $\rho_0 = 1$. Como a covariância e a variância são mensuradas nas mesmas unidades de medida, ρ_k é um número sem unidade ou puro. Ele se situa entre -1 e +1, como qualquer correlação coeficiente. Os coeficientes de autocorrelação são plotados para mostrar a função de autocorrelação, sendo este gráfico também conhecido como um correlograma.

Na prática, possui-se apenas uma amostra de um processo estocástico, sendo possível apenas computar a função de correlação amostral, $\hat{\rho}_k$. Para isso, é necessário primeiro calcular a covariância da amostra com defasagem k , $\hat{\gamma}_k$, e a variância da amostra, $\hat{\gamma}_0$, definidas pela Eq. 2.6 e Eq. 2.7.

$$\hat{\gamma}_k = \frac{\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n} \quad (2.6)$$

$$\hat{\gamma}_0 = \frac{\sum (Y_t - \bar{Y})^2}{n} \quad (2.7)$$

Nestas equações, n é o tamanho da amostra e $\bar{Y}$ é a média da variável na amostra.

Sendo assim, a função de correlação amostral, com defasagem k pode ser representado pela Eq. 2.8.

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (2.8)$$

Na Equação 2.8, $\hat{\rho}_k$ é a razão da covariância da amostra (com defasagem k) e a variância da amostra. O gráfico de $\hat{\rho}_k$ para diferentes defasagens k é conhecido como correlograma amostral.

Segundo Bueno (2011), um processo MA de ordem q é identificado pela função de autocorrelação, como pode ser visualizado na Eq. 2.9.

$$MA(q) \rightarrow \rho_j = \frac{\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_q\theta_{q-j}}{\sum_{j=1}^q \theta_j^2}, j = 1, 2, \dots, q \quad (2.9)$$

A função de autocorrelação parcial (FACP) complementa o conceito da FAC, sendo a correlação calculada entre duas observações seriais y_t e y_{t-k} , eliminando-se a dependência dos termos intermediários. É possível filtrar as correlações, mantendo-se apenas a correlação pura entre duas observações (Bueno, 2011). O autor também diz que a autocorrelação parcial é o gráfico $\varphi_{j,j}$ contra j , estimado na Eq. 2.10, em que a série em nível tem sua média subtraída. Nesta equação e_t representa o erro.

$$y_t = \varphi_{j,1}y_{t-1} + \varphi_{j,2}y_{t-2} + \dots + \varphi_{j,j}y_{t-j} + e_t, j = 1, 2, \dots, \quad (2.10)$$

Sendo assim, uma interpretação gráfica é possível de ser realizada, visando a identificação aproximada dos parâmetros autorregressivos e de médias móveis do modelo.

Segundo Morettin (2006), um modelo $AR(p)$ tem FAC que decai lentamente, com muitas autocorrelações significativas, de acordo com exponenciais e/ou senóides amortecidas, e FACP com uma queda abrupta na defasagem p . Já um processo $MA(q)$ tem FAC finita, apresentando um corte após o lag q , e FACP decaindo lentamente com muitas autocorrelações significativas, de acordo com exponenciais e/ou senóides amortecidas. Já um processo $ARMA(p, q)$ tem FAC infinita em extensão, que decai de acordo com exponenciais e/ou senóides amortecidas após o lag $q - p$. Neste modelo a FACP se comporta como a FACP de um processo MA puro.

2.5. Teste de Raiz Unitária

Nessa seção são abordados testes de Dickey-Fuller Aumentado (ADF), introduzindo pelo teste de Dickey-Fuller (DF), os quais consideram como hipótese nula a presença de raiz unitária nos dados, indicando não estacionariedade e o teste de Kwiatkowski–Phillips–Schmidt–Shin (KPSS), que considera como hipótese nula a ausência de raiz unitária, indicando que a série é estacionária. Estes testes foram utilizados de maneira a se complementarem, pois possuem suas hipóteses opostas.

2.5.1. Dickey-Fuller (DF)

Segundo Gujarati e Porter (2011), o teste de Dickey-Fuller considera o modelo apresentado na Eq. 2.11:

$$Y_t = \rho Y_{t-1} + U_t, \quad -1 \leq \rho \leq 1 \quad t = 1, 2, 3, \dots, N \quad (2.11)$$

em que: Y_t é a série a ser modelada; ρ é o coeficiente de Y_{t-1} ; e U_t é um termo de ruído branco (possui média zero, variância constante e é não-correlacionado).

Para a realização do teste, a Eq. 2.11 é trabalhada na forma de diferenças, conforme o apresentado pela Eq. 2.12:

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + U_t = (\rho - 1)Y_{t-1} + U_t. \quad (2.12)$$

Esta equação pode ser reescrita como a Eq. 2.13:

$$\Delta Y_t = \delta Y_{t-1} + U_t, \quad (2.13)$$

em que: $\delta = \rho - 1$ e Δ é o primeiro operador da diferença.

Dickey e Fuller (1981) demonstra que, sob a hipótese nula de que $\delta = 0$, o valor estimado do coeficiente de Y_{t-1} na Eq. (2.13) não segue uma distribuição de *t-student*, mas a distribuição de uma estatística denominada τ (tau). Através de simulações de Monte Carlo, os autores computaram os valores essenciais dessa estatística.

Assim, testa-se a hipótese de que $\delta = 0$, o que sugere um $\rho = 1$, e, consequentemente, a presença de uma raiz unitária, ou seja: $H_0: \delta = 0$ versus $H_1: \delta < 0$.

Na hipótese nula o modelo possui raiz unitária, sendo a série considerada não estacionária, enquanto na hipótese alternativa, o modelo não possui raiz unitária, sendo a série considerada estacionária.

Na realização deste teste foram considerados os seguintes modelos: sem constante ou tendência, com a inclusão de uma constante (β_1), ou ainda com constante e tendência ($\beta_1 + \beta_2 t$), apresentados em Eq. 2.14, Eq. 2.15 e Eq. 2.16, respectivamente.

$$\Delta Y_t = \delta Y_{t-1} + U_t \quad (2.14)$$

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + U_t \quad (2.15)$$

$$\Delta Y_t = \beta_1 + \beta_2 + \delta Y_{t-1} + U_t \quad (2.16)$$

2.5.2. Dickey-Fuller Aumentado (ADF)

O teste ADF é aplicado nos casos em que U_t são correlacionados, partindo do mesmo princípio apresentado na Eq. 2.14, Eq. 2.15 e Eq. 2.16, considerando, no entanto, valores defasados da variável dependente ΔY_t . Neste teste define-se o número máximo de defasagens das diferenças pelo critério de Schwert (1989), normalmente um valor elevado, e analisa-se a significância do último termo. Diminui-se o número de defasagens até que a última seja estatisticamente significativa. Sendo assim, para aplicação do teste ADF, nas equações 2.14, 2.15 e 2.16 são incluídas as defasagens, de tal modo que estas equações são transformadas nas equações 2.17, 2.18 e 2.19, respectivamente.

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i=1}^m \varphi_i \Delta Y_{t-i} + U_t \quad (2.17)$$

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \sum_{i=1}^m \varphi_i \Delta Y_{t-i} + U_t \quad (2.18)$$

$$\Delta Y_t = \beta_1 + \beta_2 + \delta Y_{t-1} + \sum_{i=1}^m \varphi_i \Delta Y_{t-i} + U_t \quad (2.19)$$

Ressalta-se que o teste ADF segue a mesma estatística τ que o teste DF.

2.5.3. Teste de KPSS

De acordo com Kwiatkowski et al. (1992), o teste KPSS avalia a estacionariedade considerando um modelo com tendência, passeio aleatório e erro, dada pela Eq. 2.20:

$$Y_t = \xi t - r_t + \varepsilon_t, \quad (2.20)$$

em que: ξ é uma constante que representa o coeficiente de t ; ε_t é um erro estacionário e um passeio aleatório; e $r_t = r_{t-1} + U_t$.

Considerando que e_t representa os resíduos de uma regressão em Y_t , explicado pelas componentes de tendência, passeio aleatório e intercepto, um estimador para a variância dos erros dessa regressão, é dado pela Eq. 2.21 (Kwiatkowski et al, 1992):

$$\hat{\sigma}_t^2 = \frac{SQE}{N}, \quad (2.21)$$

em que SQE é a soma dos quadrados dos erros.

A soma parcial dos resíduos, S_t , pode ser representado pela Eq. 2.22, sendo a estatística LM dada pela Eq. 2.23.

$$S_t = \sum_{i=1}^t e_i, \quad t = 1, 2, 3, \dots, N \quad (2.22)$$

$$LM = \sum_{t=1}^N \frac{S_t^2}{N^2 \hat{\sigma}_t^2}, \quad (2.23)$$

em que $\hat{\sigma}_\varepsilon^2$ é a estimativa da variância de longo prazo de Y_t .

A estatística de teste calculada na Eq. 2.23 é empregada para testar a hipótese nula de que o modelo não possui raiz unitária e a série pode ser considerada estacionária ($H_0: \hat{\sigma}_u^2 = 0$) versus a hipótese alternativa de que o modelo possui raiz unitária, a série não é estacionária ($H_1: \hat{\sigma}_u^2 > 0$). Os valores da estatística de teste LM são comparados com valores tabelados para dar suporte à tomada de decisão.

2.6. Critérios de informação

Nesta seção será descrito os critérios de informação utilizados na identificação de uma hierarquia entre os modelos candidatos pela parcimonialidade, o que pode ser feito pelos critérios de Akaike, Schwarz, Hannan-Quinn. Segundo Gujarati e Porter (2011), os critérios de informação impõe uma medida corretiva pelo acréscimo de regressores ao modelo, escolhendo os de menor valor.

2.6.1. Critério de informação de Akaike (CIA)

De acordo com Gujarati e Porter (2011), o critério CIA pode ser definido pela Eq. 2.24:

$$CIA = e^{2k/n} \frac{\sum \hat{u}_i^2}{n} = e^{2k/n} \frac{SQR}{n}, \quad (2.24)$$

em que: k é o número de regressores (incluindo o intercepto), u representa o erro estocástico, SQR é a soma dos quadrados dos resíduos e n é o número de observações. Por conveniência matemática, a Eq. 2.24 pode ser reescrita na Eq. 2.25.

$$\ln CIA = \left(\frac{2k}{n} \right) + \ln \left(\frac{SQR}{n} \right) \quad (2.25)$$

Alguns textos e programas definem CIA apenas em termos de sua transformação logarítmica, de modo que não há necessidade de colocar o termo $\ln$ antes do CIA. Na comparação de dois ou mais modelos, o que possuir o CIA mais baixo é o preferido. Este critério é útil não só dentro da amostra, mas também fora dela, prevendo o desempenho de um modelo de regressão. Outra utilização deste critério é na determinação da extensão da defasagem em um modelo $AR(p)$ (Gujarati e Porter, 2011).

2.6.2. Critério de informação de Schwarz (CIS)

Segundo Gujarati e Porter (2011), semelhante ao CIA, o critério CIS é definido pela Eq. 2.26:

$$CIS = n^{k/n} \frac{\sum \hat{u}_i^2}{n} = n^{k/n} \frac{SQR}{n}, \quad (2.26)$$

em que: k é o número de regressores (incluindo o intercepto), u representa o erro estocástico, SQR é a soma dos quadrados dos resíduos e n é o número de observações. Por conveniência matemática, a Eq. 2.26 pode ser reescrita na Eq. 2.27.

$$\ln CIS = \frac{k}{n} \ln n + \ln \frac{SQR}{n} \quad (2.27)$$

O critério de informação de Schwarz impõe edidas corretivas mais duras que o critério de informação de Akaike, como fica evidente na comparação entre a Eq. 2.25 e a Eq. 2.27. Assim como CIA, quanto mais baixo o valor do CIS, melhor será o modelo, que também pode ser usado para comparar o desempenho do modelo quando as previsões são feitas dentro e fora da amostra.

2.6.3. Critério de Hannan-Quinn (HQ)

Este critério assim como CIS e o CIA, pode ser definido pelo seu logaritmo dado pela Eq. 2.28:

$$\ln HQ = \left(\frac{2k}{n} \ln \ln n \right) + \ln \left(\frac{SQR}{n} \right), \quad (2.28)$$

em que: k é o número de regressores (incluindo o intercepto), u representa o erro estocástico, SQR é a soma dos quadrados dos resíduos e n é o número de observações.

De acordo com Bueno (2011), enquanto o critério CIS é consistente assintoticamente, e tende a escolher um modelo mais parcimonioso do que o CIA, em pequenas amostras o CIA apresenta melhores resultados. O critério de informação de Hannan-Quinn também é assintoticamente consistente, no entanto possui um poder de teste menor que o CIS.

2.7. Testes das suposições do modelo teórico

Nessa seção serão abordados os testes aplicados para validação das suposições do modelo teórico: autocorrelação nos resíduos, através do teste Ljung-Box; normalidade dos resíduos, executado pelo teste Jarque-Bera; e homegeneidade entre variâncias, o por meio do teste ARCH.

2.7.1. Teste de Ljung-Box

Box e Pierce (1970, *apud* Morettin, 2006), sugeriram um teste para as autocorrelações dos resíduos estimados, que, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma modificação deste teste foi proposta por Ljung e Box (1978, *apud* Morettin, 2006, p. 102), onde o teste é aplicado aos resíduos de uma série temporal após o ajustamento de um modelo, examinando m autocorrelações dos resíduos. Se as autocorrelações forem muito pequenas, conclui-se que o modelo não exibe falha significativa. Sendo assim, o teste de Ljung-Box é uma ferramenta diagnóstica usada para testar a falha do ajuste de um modelo de série temporal.

Em geral, o teste de Box-Ljung testa as seguintes hipóteses:

H_0 : o modelo não possui falha no ajuste por não apresentar resíduos autocorrelacionados;

H_1 : o modelo possui falha no ajuste por apresentar resíduos autocorrelacionados.

Dada uma série temporal Y de tamanho n , e a estatística de teste é definida pela Eq.2.29:

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{r}_k^2}{n-k}, \quad (2.29)$$

em que: $\hat{r}_k$ é a autocorrelação estimada da série no lag k , e m é o número de lags que está sendo testado. Em geral, basta utilizar as 15 ou 20 primeiras $\hat{r}_k$ (Morettin, 2006).

2.7.2. Teste de Jarque-Bera (JB)

De acordo com Gujarati e Porter (2011), o teste de Jarque-Bera é um teste assintótico ou de amostra grande e se baseia nos resíduos de MQO. Ele calcula a estatística de teste, a partir da assimetria e da curtose dos resíduos, pela Eq. 2.30:

$$JB = n \left(\frac{S^2}{6} + \frac{(K-3)^2}{24} \right), \quad (2.30)$$

em que: n é o tamanho da amostra, S é o coeficiente de assimetria e K é o coeficiente de curtose.

Ainda, segundo Gujarati e Porter (2011), para uma variável normalmente distribuída, $S = 0$ e $K = 3$. Portanto, o teste Jarque-Bera de normalidade é um teste da

hipótese conjunta de que S e K são iguais a 0 e 3, respectivamente. Nesse caso, espera-se que o valor da estatística JB seja igual a 0. Sob a hipótese nula de que os resíduos são normalmente distribuídos, Jarque e Bera demonstraram que, assintoticamente (isto é, em amostras grandes), a estatística JB dada na Eq. 2.30 segue a distribuição de qui-quadrado com 2 graus de liberdade. Se o valor p calculado para a estatística JB em uma aplicação for suficientemente baixo, o que acontece quando o valor da estatística é muito diferente de zero, podemos rejeitar a hipótese de que a distribuição dos resíduos é normal. Mas se o valor p for razoavelmente alto, o que acontece quando o valor da estatística está próximo de zero, não rejeitaremos a hipótese de normalidade.

2.7.3. Teste ARCH

De acordo com Gujarati e Porter (2011), o termo de erro u no tempo t pode estar correlacionado com o termo de erro no tempo $(t - 1)$ em um processo AR(1) ou com os vários termos de erro defasados em um processo geral AR(p), podendo haver autocorrelação na variância σ^2 no tempo t com seus valores defasados em um ou mais períodos. Essa autocorrelação recebe nomes como heterocedasticidade condicional autorregressiva (ARCH - autoregressive conditional heteroscedasticity) e ocorre quando a variância do erro está relacionada com o termo de erro elevado ao quadrado no período anterior.

Um modelo ARCH pode ser definido pela Eq. 2.31:

$$X_t = \sqrt{h_t} \varepsilon_t, \quad (2.31)$$

onde

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2, \quad (2.32)$$

em que ε_t i.i.d. (0,1), $\alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, \dots, r$.

Segundo Gujarati e Porter (2011), uma condição que garante que um processo ARCH(q) é estacionário pode ser visualizado na Eq. 2.33.

$$\sum_{i=1}^q \alpha_i < 1 \quad (2.33)$$

Um processo ARCH estacionário, possui as estatísticas apresentadas na Eq. 2.34, Eq. 2.35 e Eq. 2.36.

$$E(X_t) = 0; \quad (2.34)$$

$$Var(X_t) = \frac{\alpha_0}{(1 - \sum_{i=1}^q \alpha_i)} \quad (2.35)$$

$$Cov(X_t, X_{t-k}) = E(X_t, X_{t-k}) = 0, \quad k \geq 1 \quad (2.36)$$

Para melhorar a compreensão de um modelo ARCH, consideremos um modelo ARCH(1), em que a sua variância condicional depende apenas do instante anterior. Assim as Eq. 2.31 e Eq. 2.32 podem ficar na forma da Eq. 2.37 e Eq. 2.38.

$$X_t = \sqrt{h_t} \varepsilon_t \quad (2.37)$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 \quad (2.38)$$

em que $\alpha_0 > 0$ e $\alpha_1 \geq 0$.

Pode-se demonstrar que o modelo possui as seguintes propriedades: a média não condicionada de X_t é zero e a variância não condicionada de X_t é dada pela Eq. 2.39.

$$Var(X_t) = \frac{\alpha_0}{1 - \alpha_1} \quad (2.39)$$

Na construção de modelos ARCH, um primeiro passo é tentar ajustar um modelo ARIMA, para remover a correlação serial na série, se esta existir.

2.8. Análise dos erros

Nesta seção são apresentados os indicadores utilizados para verificar a qualidade de um modelo a partir dos erros ou resíduos. São eles: erro absoluto médio (MAE), raiz do erro quadrático médio (RMSE) e erro absoluto percentual médio (MAPE), além da estatística U de Theil.

2.8.1. Erro absoluto médio (MAE)

É a média dos erros absolutos, onde o erro absoluto é a diferença absoluta entre o valor real e o estimado.

2.8.2. Erro quadrático médio (RMSE)

É a medida de erro mais comumente usada para aferir a qualidade do ajuste de um modelo. Nada mais é do que a raiz do erro médio quadrático da diferença entre o valor previsto e o valor real. Pode-se pensar nela como sendo uma medida análoga ao desvio padrão. RMSE é uma boa medida, porque geralmente ela representa explicitamente o que vários métodos tendem a minimizar.

2.8.3. Erro absoluto percentual médio (MAPE)

Esta medida mede o erro em percentagem, sendo uma das mais utilizadas em função da sua facilidade de interpretação. É expressa em porcentagem.

2.8.4. Estatística U de Theil

Ela associa o desempenho do modelo predictor ao de um modelo de passeio aleatório. Caso U de Theil seja igual a 1, o modelo utilizado possui o mesmo desempenho que o modelo de passeio aleatório. Sendo maior que 1, o desempenho é considerado inferior ao de um modelo de passeio aleatório; caso contrário, com U de Theil menor que 1, o modelo utilizado possui desempenho superior ao modelo de passeio aleatório. Modelos de melhor desempenho têm essa medida tendendo a 0.

2.9. Literatura relacionada

Inúmeros trabalhos já abordaram o tema no cenário mundial. Westley (1992) mostrou literaturas que abordavam o estudo da demanda em países latinos americanos e comparou com as elasticidades renda e preço da demanda de energia elétrica nos Estados Unidos. Após alguns anos, Silk e Joutz (1997) estimaram a demanda de energia elétrica nos Estados Unidos por meio de um modelo de ajustamento parcial. Poucos anos depois, Halvorsen e Larsen (2001) publicam um trabalho sobre a demanda residencial de energia elétrica na Noruega com a utilização de um Modelo de Correção de Erros. Empregando

a mesma técnica, Ettestol (2002) igualmente analisou a demanda residencial na Noruega.

No trabalho de Chui et al. (2009) é evidenciado a importância de planejamento otimizado da produção de energia elétrica em função do produto não poder ser estocado. Soma-se o fato de que o aumento de produção não pode ser imediato, pois a construção de usinas geradoras leva alguns anos. Esses autores também consideram muito pertinente a importância da incerteza na determinação dos parâmetros de um modelo de previsão da demanda de energia elétrica, além de manifestarem a importância das variáveis climáticas, socioeconômicas e demográficas na previsão desta demanda.

Kavanagh (2017) analisou o mercado irlandês, avaliando a necessidade de previsões para cumprir as metas de segurança no fornecimento, promoção de energias renováveis, promoção da concorrência entre fornecedores, maximização da eficiência dos fornecedores e distribuidores, além de propiciar o aparecimento de um mercado estável e previsível.

Debnath e Mourshed (2018) estudaram os modelos de planejamento energético, analisando 483 estudos, nos quais a maioria dos modelos foram criados visando a previsão de demanda de energia elétrica, sendo os modelos ARIMA o terceiro mais utilizado, atrás dos modelos de redes neurais e dos de suporte vetoriais. No final do artigo fizeram um levantamento do tipo de método mais utilizado em cada um dos países estudados, identificando que os métodos estatísticos são os mais utilizados no Brasil.

Nacionalmente, um dos primeiros trabalhos foi de Modiano (1984), que determinou a nível nacional as elasticidades da demanda de energia elétrica das três principais classes de demanda: residencial, comercial e industrial, além da classe denominada “outros”. No seu trabalho, considerou a periodicidade anual, entre os anos de 1963 e 1981, e utilizou o método dos mínimos quadrados como método de estimação dos coeficientes das equações. O trabalho se ateve a encontrar a relação de longo prazo entre a demanda de energia e as variáveis renda e tarifa, porém, não se estendeu a previsões. Ele utilizou como proxies para o setor residencial apenas o PIB per capita para a variável renda e a tarifa média para a variável tarifa. Nos setores comercial e industrial foram utilizados a tarifa média e o PIB como proxies.

Andrade e Lobão (1997) encontraram as elasticidades ligadas a demanda de energia elétrica do setor residencial, com periodicidade anual, no intervalo dos anos de 1963 à 1997, e realizaram previsões de demanda até o ano de 2005. Empregaram três técnicas distintas: Mínimos Quadrados Ordinários (MQO), Variáveis Instrumentais, e Cointegração e VAR-VECM, comparando os resultados das estimações. Neste trabalho,

as variáveis demanda da energia e tarifa podem ser consideradas dependentes entre si, ou seja, uma pode afetar a outra, existindo uma provável relação de simultaneidade entre as variáveis. Os proxies utilizados foram a tarifa média, o PIB per capita e o índice real de preços dos eletrodomésticos, respectivamente para as variáveis tarifa, renda e preço.

Depois da famigerada crise energética no início dos anos 2000, conhecida nacionalmente como a “Crise do apagão”, em que o país passou por racionamento de energia elétrica, os temas pertinentes à demanda de energia começaram a receber mais atenção. Por esse motivo, apareceram diversos trabalhos relacionados ao tema energético. Em 2004, Schmidt e Lima (2004) avaliaram as elasticidades para os setores residencial, comercial e industrial, no período entre 1969 e 2000, realizando previsões até o ano de 2005. Os autores utilizam a análise de cointegração proposta por Johansen e modelagem VAR-VEC para estimar as elasticidades e previsões, utilizando como proxies a tarifa média, a renda (rendimento médio dos assalariados no caso residencial), produto interno bruto a preços de mercado (PIBpm, para os setores industrial e comercial), IPA-DI - Bens de consumo duráveis - utilidade como preço dos eletrodomésticos no setor residencial, IPA-OG - Mat. Elétrico como preço dos eletrointensivos ligados ao comércio, e IPA-OG - máquinas e equipamentos elétricos como preço dos eletrointensivos ligados à indústria, além disso o índice IPA-OG - combustível e lubrificantes que representa o preço do bem substituto à energia elétrica.

Já o trabalho de Mattos (2004) estima a demanda de energia elétrica para os três principais setores de consumidores do estado de Minas Gerais. O autor utiliza dados com periodicidade anual entre os anos de 1970 e 2002 e realiza projeções por meio de modelagem VAR-VEC para os anos de 2004 a 2008, utilizando como proxies a tarifa média, o PIB de Minas Gerais (MG) como renda e o índice real de preços de eletrodomésticos para o setor residencial. Para o setor comercial foi utilizado o PIB-MG Comercial, a tarifa média e o índice real de preços de material elétrico. No setor industrial foi utilizado o PIB-MG indústria, a tarifa média, IPA-OG - máquinas e equipamentos elétricos como preço dos eletrointensivos ligados à indústria e o índice IPA-OG - combustível e lubrificantes representando o preço do bem que substitui a energia elétrica.

Logo após, Mattos e Lima (2005) estimam as elasticidades da demanda apenas para setor residencial de energia para o estado de Minas Gerais. Os autores utilizam dados com periodicidade anual, entre os anos de 1970 e 2002, além de aplicar uma metodologia semelhante ao utilizado por Mattos (2004). Os autores utilizam como proxies a tarifa média, o PIB de Minas Gerais per capita como renda e o índice real de preços de

eletrodomésticos para o setor residencial.

Em Leite (2006) é projetada a demanda de energia elétrica em nível nacional, para o período de 2006 a 2015. O autor também utiliza a modelagem VAR-VEC, além de um modelo ARIMA, e lida com todos os setores em conjunto e dados anuais de 1966 a 2005. No setor residencial trabalhou com a tarifa média, com o PIB (representando a renda) e o Índice Geral de Preços - Disponibilidade Interna (IGP-DI) para o preço. Nos setores comercial e industrial foram utilizados a tarifa média, o PIB e o IGP-DI como proxies.

Também em 2006, Siqueira et al. (2006) apresentam, para a região nordeste do Brasil, modelos para as classes residencial, comercial e industrial usando MQO, MQO em dois estágios e VAR-VEC. Com dados em períodos anuais entre 1970 e 2003, por meio da técnica VAR-VEC, são realizadas previsões para os anos de 2004 a 2010. Os proxies para o setor residencial utilizados foram a tarifa média, o PIB per capita do Nordeste (representando a renda) e o Índice de Preços por Atacado – Oferta Global (IPA-OG) dos Eletrodomésticos (representando o preço). No setor comercial, os proxies empregados foram a tarifa média, o PIB do Nordeste (NE) de serviços e o IPA-OG de material elétrico, enquanto no âmbito industrial foram utilizados a tarifa média, o PIB do Nordeste (NE) de indústria, o IPA-OG de máquinas e equipamentos industriais e o IPA-OG de combustíveis e lubrificantes.

Do mesmo modo, focado na região Nordeste, o trabalho de Irffí et al. (2009) estima a demanda por energia elétrica das três grandes classes. Os autores utilizam dados anuais de 1970 a 2003 e efetuam previsões para o período de 2004 a 2010, utilizando a técnica de Mínimos Quadrados Dinâmicos. As previsões foram efetuadas pelo modelo VEC com as equações encontradas pelo Método dos Mínimos Quadrados Dinâmicos. Com os seguintes proxies para o setor residencial: a tarifa média para a variável tarifa, com o PIB do Nordeste, representando a renda, e o IPA-OG dos Eletrodomésticos para o preço. Para os setores comercial e industrial, os proxies usados foram a tarifa média, o PIB do Nordeste (NE) de serviços e indústria, respectivamente, o IPA-OG de material elétrico para representar o preço no campo comercial e o IPA-OG de máquinas e equipamentos industriais e o IPA-OG de combustíveis e lubrificantes para a esfera industrial.

No ano seguinte, Viana (2010) realizou projeções para os setores residencial, industrial e comercial, por meio da metodologia VAR-VEC, empregando dados anuais entre 1975 e 2006. Os proxies utilizados para o setor residencial foram a tarifa média, o PIB per capita e o IPA-OG dos Eletrodomésticos para o preço. Nos setores comercial e

industrial, os proxies usados foram a tarifa média para representar a variável tarifa, o PIBpm para simular a renda e o IPA-OG de material elétrico no campo comercial para o preço no setor comercial e o IPA-OG de máquinas e equipamentos industriais e o IPA-OG de combustíveis e lubrificantes para o setor industrial.

Lima (2011) obteve para o estado de Pernambuco, as elasticidades dos setores residencial, industrial e comercial, considerando dados anuais de 1970 a 2009. O autor, assim como vários outros, utilizou a modelagem VAR-VEC e fez projeções de demanda até o ano de 2014. Empregou como proxies para o setor residencial, a tarifa média, o PIB de Pernambuco e o IPA-OG dos Eletrodomésticos. Para os setores comercial e industrial, os proxies usados foram a tarifa média, o PIB de Pernambuco (PE) de serviços e indústria, respectivamente e o IPA-OG de material elétrico para representar o preço no setor comercial e o IPA-OG de máquinas e equipamentos industriais e o IPA-OG de combustíveis e lubrificantes para representar o preço no setor industrial.

Já Viana e Silva (2014) realizaram a projeção da demanda por energia elétrica para o setor residencial no Brasil para o período de 2009 até 2017, por meio da metodologia VAR-VEC. Os autores utilizaram dados anuais de 1975 a 2006. Utilizaram a tarifa média residencial, o PIB per capita como renda e o IPA-OG do eletrodomésticos como a variável preço.

Silveira (2017) estudou alguns modelos já existentes para representar a evolução da demanda de energia elétrica no Brasil, elaborando novos modelos para os três principais setores consumidores: residencial, comercial e industrial. Seus modelos poderiam ser considerados uma atualização dos modelos já existentes para caracterizar esta variável, como Schmidt e Lima (2004) e Viana (2010), com adição de algumas variáveis dummy. Sua modelagem utilizou o período de 1985 a 2014, sendo os dados de 2015 utilizados para sua validação. Previsões foram feitas para os anos de 2016 a 2019 para as três principais classes consumidoras.

O *proxy* utilizada para a variável renda foi o PIB, sendo o PIB per capita para o setor residencial, o PIB - serviços - comércio para o setor comercial e o PIB - indústria para o setor industrial. Já a tarifa média de cada setor no Brasil representa o preço da energia elétrica. No setor residencial, para o preço dos eletrodomésticos, foi usado o índice IPAOG- eletrodomésticos até o ano de 2007, quando foi descontinuado, e o índice IPA- Origem - OG-DI - Produtos Industriais - Indústria de Transformação - Máquinas e Equipamentos - Item Eletrodomésticos, a partir de 2008, tendo como fonte a Fundação Getúlio Vargas (FGV, 2018). O preço dos eletrointensivos empregados no setor comercial

tem como *proxy* o índice IPA-OG Material Elétrico até o ano de 2007, quando foi descontinuado, e o IPA- Origem - OG-DI - Produtos Industriais - Indústria de Transformação - Máquinas, Aparelhos e Materiais Elétricos, a partir do ano de 2008, tendo como fonte a FGV. Já no setor industrial são utilizados como *proxy* o índice IPA-OG- Máquinas e Equipamentos Industriais e o índice IPA- Origem - OG-DI - Produtos Industriais – Indústria de Transformação - Máquinas e Equipamentos, até o ano de 2007 e a partir de 2008, respectivamente. O preço do bem substituto à energia elétrica usado no modelo do setor industrial é representado pelo índice IPA-OG Combustível e Lubrificantes (até o ano de 2007) e IPA-EP - Bens Intermediários - Combustíveis e Lubrificantes para a Produção (a partir do ano de 2008), possuindo como fonte a FGV (2018).

Outra informação importante, é que devido à quebra na estrutura dos dados do modelo durante o período do racionamento de energia elétrica, ocorrida entre os anos de 2001 e 2002 e da crise econômica mundial no ano de 2009, o autor introduziu uma variável dummy para cada um dos três setores estudados. No setor residencial foi notada uma quebra estrutural após o racionamento energético, logo entre os anos de 1985 e 2000, o valor de uma variável dummy foi zero, e a partir o ano de 2001, esta variável dummy assumiu o valor um. No setor comercial, foi notada uma alteração no período de racionamento, sendo assumido que não houve uma quebra na estrutura da série, logo, a variável dummy inserida teve um valor igual a um, somente entre os anos de 2001 e 2002, sendo zero no restante do período analisado. Já no setor industrial foi necessário inserir duas variáveis dummy: uma para o período do racionamento, entre os anos de 2001 e 2002, e outra para o período da crise internacional de 2009. Ambas só possuíram valor diferente de zero nestes anos mencionados, pois assim como no setor comercial, foi assumido que não ocorreu uma quebra na estrutura da série estudada.

Com os resultados obtidos por Silveira (2017), foi possível notar que o setor residencial foi o que apresentou os melhores resultados, dentre os três principais setores de consumo de energia elétrica no Brasil. Este apresentou o menor erro percentual para a validação do modelo, e proporcionou as previsões mais próximas às realizadas pela EPE. Também foi mencionado que sua elasticidade-renda foi superior à de trabalhos anteriores ao período crítico entre os anos 2001 e 2002, e próximas a resultados encontrados por trabalhos posteriores a este período. No que se refere à elasticidade-preço dos eletrodomésticos, nota-se um aumento nesta em comparação aos trabalhos realizados após a alteração da maneira de calcular os índices de IPA feitas pela FGV.

No setor comercial, o erro de previsão foi inferior a 5% na validação do modelo, e diferença percentual absoluta média ficou pouco superior a 6% em comparação com as projeções realizadas pela EPE, mostrando que é possível melhorá-lo, por meio da identificação de variáveis capazes de representar com mais exatidão os eletrointensivos usados pelo comércio e agregá-las ao modelo, desta maneira, foi considerado satisfatório pelo autor. A elasticidade-renda foi semelhante aos resultados dos estudos realizados antes do período da crise do apagão, sugerindo que não ocorreu mudança significativa posteriormente. No que se trata da variável tarifa, notou-se que os valores encontrados foram superiores aos dos trabalhos anteriores ao período de racionamento e próximos aos de trabalhos realizados depois deste período, estando de acordo com outros estudos realizados sobre o tema e indicando que o período de racionamento influenciou esta elasticidade-preço da tarifa. Já a elasticidade-preço dos equipamentos eletrointensivos sofreu uma alteração em seus valores, que foram superiores aos dos trabalhos anteriores à alteração dos índices IPA pela FGV, indicando a sua interferência na elasticidade-preço.

O setor industrial foi o que apresentou mais variáveis com possibilidade de interferência na demanda de energia elétrica, o que, de acordo com Silveira (2017), torna a modelagem difícil. O modelo encontrado apresentou os maiores erros na sua validação, 6,8%, e produziu previsões superiores às realizadas pela EPE, o que, de acordo com o autor, indica que o modelo pode ser melhorado com a inclusão de novas variáveis. A elasticidade-renda e a elasticidade-preço da tarifa mostraram-se próximas aos trabalhos anteriores ao período de racionamento, sem indicar a interferência deste período. Já a elasticidade-preço dos bens eletrointensivos, apresentou um valor inferior à elasticidade da tarifa e de trabalhos realizados antes do ano de reformulação dos índices IPA pela FGV. Este setor apresentou a maior elasticidade-renda, sendo o único com uma demanda elástica à renda.

3. METODOLOGIA

Neste estudo, em acordo com a quase totalidade dos já realizados, considera-se que toda a demanda de energia é efetivamente fornecida e, desta forma, a quantidade demandada foi aproximada à quantidade de energia elétrica consumida.

O estudo da série temporal da demanda de energia elétrica no estado do Rio Grande do Sul foi realizado separadamente em cada um dos três principais setores de demanda: residencial (Dr), comercial (Dc) e industrial (Di), considerando períodos anuais entre os anos de 1971 e 2010. A validação dos modelos encontrados para previsão foi feita para os próximos anos: 2011 a 2017.

Os dados da demanda de energia elétrica no Rio Grande do Sul entre os anos de 1971 e 2011 foram obtidos na base de dados do Instituto de Pesquisa Econômica Aplicada (IPEA, 2017). Já os dados entre os anos de 2012 e 2014 foram coletados da Empresa de Pesquisa Energética (EPE, 2015), vinculada ao Ministério de Minas e Energia. Os dados para os anos que compreendem o período de 2015 até 2017 foram obtidos em Eletrobrás (2018).

Os procedimentos descritos a seguir foram realizados separadamente para os setores residencial, comercial e industrial.

3.1. Metodologia de Box-Jenkins

O ajuste do modelo autoregressivo integrado de médias móveis (ARIMA), apresentado na Eq. 2.1 foi realizado conforme o descrito a seguir.

Inicialmente foi realizada uma análise exploratória de dados (AED), avaliando a possibilidade de alteração de sua métrica pela aplicação da transformação logarítmica, para diminuir a variabilidade. Segue a construção do gráfico em linhas para uma primeira avaliação do comportamento dos dados em nível da série estudada, verificando a presença ou não de tendência. Foram então construídos os gráficos da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) a partir dos dados em nível para complementar a análise anterior. Aplica-se o teste de Dickey-Fuller Aumentado (ADF) na série em nível para verificar a hipótese nula de que a série não é estacionária. Se forem encontradas evidências para rejeitar esta hipótese nula (valor $p < 0,01$), considera-se a série como estacionária. Este procedimento será executado de forma iterativa,

considerando o critério de Schwert (1989) para determinar o número máximo de defasagens a ser considerado. Além disso, foram consideradas a inclusão de tendência e constante, apenas constante e nenhuma delas. Depois de definir o número máximo de defasagens (10), o teste ADF foi aplicado, sendo analisada a significância estatística da última defasagem de maneira iterativa até a identificação de valor significativo. Para a inclusão ou não de constante e tendência no teste ADF, com constante e tendência. Se a tendência não foi significativa, ela foi retirada e a equação foi estimada novamente para avaliar se a constante era significativa. Em caso negativo, foi retirada do modelo. Para complementar o teste ADF, foi aplicado o teste de Kwiatkowski, Phillips, Schmidt e Shin (KPSS) para verificar a hipótese nula de que a série é estacionária. Se não fossem encontradas evidências para rejeitar esta hipótese nula (valor $p > 0,01$), a série era considerada estacionária. Na operacionalização deste teste, a quantidade máxima de diferenças foi definida da mesma forma que no teste ADF, o mesmo acontecendo em relação à tendência. Na constatação de não estacionaridade, foram realizadas diferenças sucessivas com repetição dos procedimentos descritos anteriormente, até consegui-la.

Após, são identificadas as ordens de modelos candidatos pela análise da FAC e da FACP da série identificada como estacionária. A confirmação dos modelos candidatos é feita por meio da utilização do software *software R* (R Core Team, 2018) e sua função *auto.arima*. Código se encontra no Apêndice 01.

Nos modelos candidatos são determinados os coeficientes com avaliação de sua significância, além de ser definida uma hierarquia entre os mesmos por sua parcimonialidade, o que foi feito pelos critérios de informação de Akaike, Schwarz, Hannan-Quinn. A seguir são avaliadas as suposições do modelo teórico sobre propriedades dos dados: existência de autocorrelação nos resíduos, o que foi executado pelo teste Ljung-Box; normalidade, o que foi feito pelo teste Jarque-Bera, e homogeneidade entre variâncias, o que foi feito pelo teste ARCH.

A qualidade do ajuste é determinada pelos indicadores da qualidade dos modelos encontrados: erro absoluto médio (MAE), erro quadrático médio (RMSE) e erro absoluto percentual médio (MAPE), além da estatística U de Theil com identificação dos melhores modelos para ajuste. Segue-se a validação dos modelos encontrados para previsão com a determinação dos indicadores: erro absoluto médio (MAE), erro quadrático médio (RMSE) e erro absoluto percentual médio (MAPE), além da estatística U de Theil para os período 2011-2017, com identificação dos melhores modelos para previsão. Nesta análise também foram considerados os intervalos de confiança para previsão.

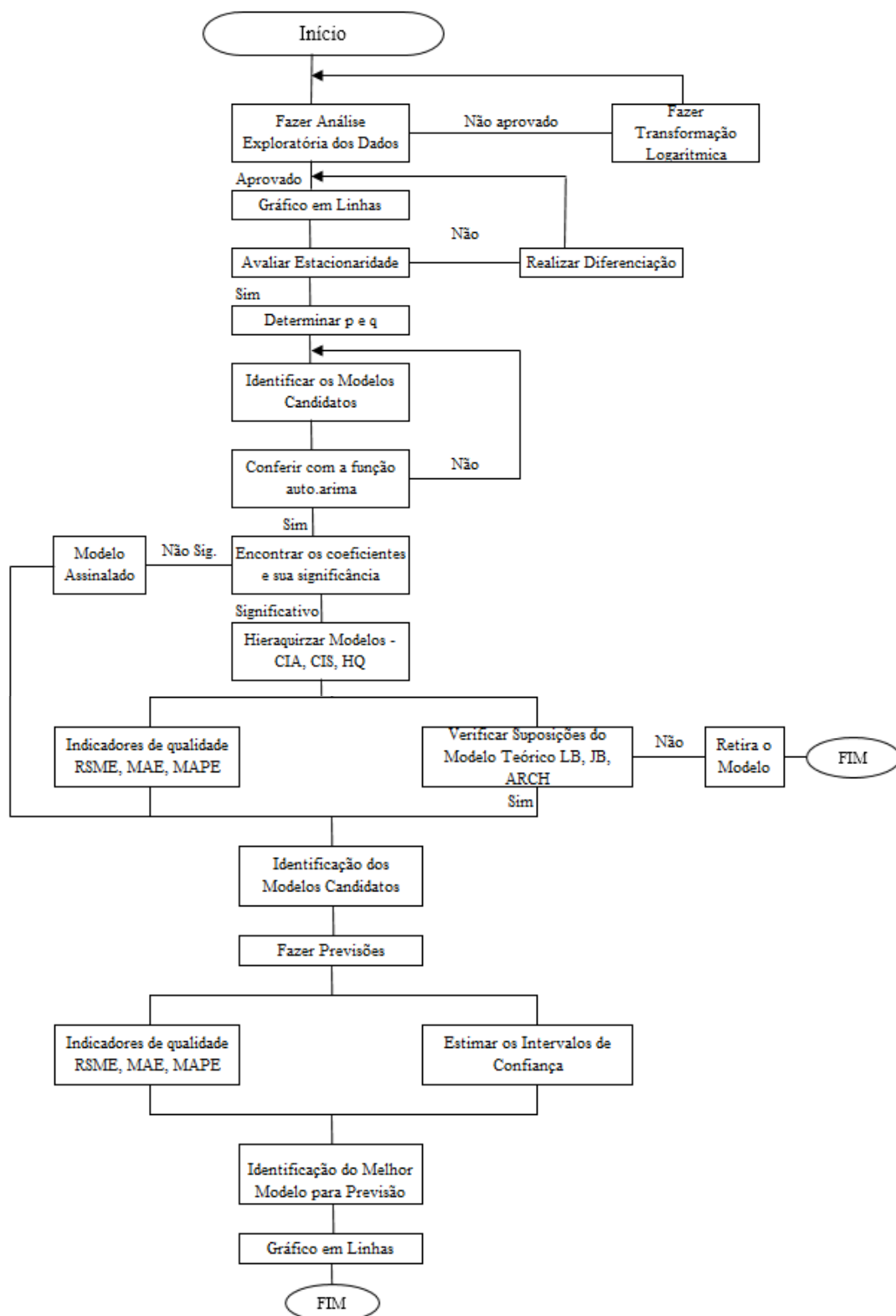


Figura 3.1 -Fluxograma da metodologia Box Jenkins

3.2. Metodologia VAR-VEC

Nesta etapa foi ajustado um modelo de vetores autoregressivos (VAR), complementado com modelo de correção dos erros vetoriais (VEC). Partiu-se de um modelo misto multiplicativo fundamentado na função de Cobb-Douglas, conforme Modiano (1984), Andrade e Lobão (1997), Schmidt e Lima (2004), Mattos e Lima (2005), Leite (2006), Lima (2011), Viana e Silva (2014) e Silveira (2017). Segundo Ferreira e Silva (2015), a função de Cobb-Douglas foi proposta em 1928 partindo do princípio de que a produção é determinada por um modelo multiplicativo determinado a partir da quantidade de capital investido e da mão de obra empregada, dado pela Eq. 3.1:

$$P(L, Z) = b \cdot L^{\psi} \cdot Z^{1-\psi} \quad (3.1)$$

Em que: P é a produção total; L é a quantidade de trabalho; Z é a quantidade de capital investido; ψ é a elasticidade trabalho-produção; $1 - \psi$ é a elasticidade capital-produção; e b é o intercepto da função de produção.

No caso da modelagem da demanda de energia elétrica, esta função assume a forma da Eq. 3.2, que é transformada pela função logarítmica em uma função aditiva, de acordo com a Eq. 3.3, facilitando a determinação de seus parâmetros.

$$D_t = k \cdot T_t^{\gamma} \cdot R_t^{\theta} \cdot E_t^{\alpha} \cdot S_t^{\beta} \quad (3.2)$$

Aplicando logaritmo natural em ambos os lados da equação não linear, tem-se:

$$\ln D_t = \ln k + \gamma \ln T_t + \theta \ln R_t + \alpha \ln E_t + \beta \ln S_t \quad (3.3)$$

Em que, no tempo t :

D_t é o t -ésimo demanda (residencial, comercial ou industrial);

k é uma constante positiva, cujo logaritmo representa o intercepto da função de demanda;

T_t é a t -ésima tarifa (residencial, comercial ou industrial);

R_t é a t -ésima renda;

E_t é o t -ésimo preço dos eletrodomésticos no caso residencial ou dos eletrointensivos nos casos comercial e industrial;

S_t é o t-ésimo preço de um bem substituto à energia elétrica, aplicado apenas no caso industrial;

γ é a elasticidade-preço em relação à tarifa;

θ é a elasticidade-renda;

α é a elasticidade-preço em relação aos eletrodomésticos;

β é a elasticidade-preço em relação ao bem substituto à energia elétrica;

Para a modelagem desta etapa, além dos dados sobre a demanda de energia elétrica no Rio Grande do Sul (Dr , Dc e Di), já identificados anteriormente, foram utilizadas informações sobre as variáveis identificadas na Tabela 3.1.

Tabela 3.1 - Dados e fontes

Variável	Descrição	Fonte
Dr	Demanda total residencial	IPEA e EPE
Tr	Tarifa média residencial ¹	IPEA e ANEEL
Rr	PIB-RS per capita ²	IBGE in FEE
Er	IPA-OG-eletrodomésticos ³ e IPA- Origem - OG-DI - Produtos Industriais - Indústria de Transformação - Máquinas e Equipamentos - Item Eletrodomésticos ³	FGV in IPEA
Dc	Demanda total comercial	IPEA e EPE
Tc	Tarifa média comercial ¹	IPEA e ANEEL
Rc	PIB-RS de serviços – comércio ²	IBGE in FEE
Ec	IPA-OG Mat. Elétrico ³ e IPA- Origem - OG-DI – Produtos Industriais - Indústria de Transformação - Máquinas, Aparelhos e Materiais Elétricos ³	FGV in IPEA
Di	Demanda total industrial	IPEA e EPE
Ti	Tarifa média industrial ¹	IPEA e ANEEL
Ri	PIB-RS da indústria ²	IBGE in FEE
Ei	IPA-OG- Máquinas ³ e Equipamentos Industriais e IPA- Origem - OG-DI - Produtos Industriais – Indústria de Transformação - Máquinas e Equipamentos ³	FGV in IPEA
Si	IPA-OG Combustível ³ e Lubrificantes. e IPA-EP - Bens Intermediários - Combustíveis e Lubrificantes para a Produção ³	FGV in IPEA

Fonte: Elaborada pelo Autor com base em Silveira (2017).

A tarifa média (T_r , T_c , T_i) de cada setor no Rio Grande do Sul, representa o preço da energia elétrica, e foram disponibilizadas pelo Instituto de Pesquisa Econômica Aplicada (IPEA) de 1971 até 1993, quando a tarifa elétrica era nacional. A partir deste ano, a tarifa foi separada por concessionária e os valores se referem ao Rio Grande do Sul (ANEEL, 2017).

O PIB-RS é usado como *proxy* para a variável renda, tendo como fonte a Fundação de Economia e Estatística (FEE) em parceria com o IBGE, sendo o PIB-RS per capita para o setor residencial (R_r), o PIB-RS de serviços - comércio para o setor comercial (R_c) e o PIB-RS indústria para o setor industrial (R_i).

As proxies utilizadas para as variáveis preço neste trabalho foram as mesmas utilizadas por Silveira (2017), pois foram fornecidos a nível nacional. O índice de preços por atacado, origem geral (IPA-OG) é utilizado como *proxy* para a variável preço dos eletrointensivos, tendo como fonte a Fundação Getúlio Vargas (FGV), sendo o IPA-OG dos eletrodomésticos para o setor residencial (E_r), o IPA-OG do Material Elétrico para o setor comercial (E_c), o IPA-OG de Maq. e Equipamentos Industriais (E_i) e IPA-OG de Combustíveis e Lubrificantes (S_i) para o setor industrial.

Devido à quebra ocorrida em algumas das séries estudadas, foi necessária a avaliação da introdução de algumas variáveis dummy, de acordo com as descrições a seguir.

No ano de 1993, houve a mudança na tarifação da energia elétrica no país, deixando de ser um valor nacional e passando a ser calculada por concessionária. Desta forma foi criada a variável dummy 1 ($DYr1$, $DYc1$ e $DYi1$), representada pela Tab. 3.2.

Tabela 3.2 Caracterização das variáveis que representam mudança no cálculo da tarifa
 $DYr1$, $DYc1$ e $DYi1$

Período	$DYr1$, $DYc1$, $DYi1$
1971-1992	0
1993-2010	1

Fonte: Elaborada pelos autores.

Entre os anos 2001 e 2002 houve uma crise energética no país, conhecida como a “crise do apagão”. Desta forma foi criada a variável dummy 2 ($DYr2$, $DYc2$ e $DYi2$), visível na Tab. 3.3.

Tabela 3.3 Caracterização das variáveis que representam ocorrência de crise energética $DYr2$, $DYc2$ e $DYi2$

Período	$DYr2$, $DYc2$, $DYi2$
1971-2000	0
2001-2002	1
2003-2010	0

Fonte: Elaborada pelos autores.

Outras variáveis dummy analisadas foram as que representam troca de moeda no país e implementação de planos econômicos entre os anos de 1986 à 1994. Desta forma foram criadas outras três variáveis dummy: $DYr3$, $DYc3$ e $DYi3$ para o período em que a moeda nacional era o Cruzeiro, representadas pela Tab. 3.4; $DYr4$, $DYc4$ e $DYi4$ para o período de trocas frequentes de moedas, representadas pela Tab. 3.5; e $DYr5$, $DYc5$ e $DYi5$ para o período da moeda Real que vigora atualmente, representadas pela Tab. 3.6.

Tabela 3.4 Caracterização das variáveis que identificam período pré planos econômicos $DYr3$, $DYc3$ e $DYi3$

Período	$DYr3$, $DYc3$, $DYi3$
1971-1985	1
1986-2010	0

Fonte: Elaborada pelos autores.

Tabela 3.5 Caracterização das variáveis que identificam o período dos planos econômicos $DYr4$, $DYc4$ e $DYi4$

Período	$DYr4$, $DYc4$, $DYi4$
1971-1985	0
1986-1993	1
1994-2010	0

Fonte: Elaborada pelos autores.

Tabela 3.6 Caracterização das variáveis que identificam período pós planos econômicos $DYr5$, $DYc5$ e $DYi5$

Período	$DYr5$, $DYc5$, $DYi5$
1971-1993	0
1994-2010	1

Fonte: Elaborada pelos autores.

A crise econômica mundial que, segundo Pereira (2009), “originou-se em meados de 2007 no mercado norte-americano de hipotecas de alto risco (subprime), transformou-se, após a falência do banco de investimentos Lehman Brothers, numa crise sistêmica”, chegando ao Brasil no final do ano de 2008 e se estendendo no ano de 2009. Sendo assim, foi criada a variável dummy 6, que pode ser observada na Tab. 3.7.

Tabela 3.7 Caracterização das variáveis que representam período de crise econômica mundial: $DYr6$, $DYc6$ e $DYi6$

Período	$DYr6, DYc6, DYi6$
1971-2008	0
2009	1
2010	0

Fonte: Elaborada pelos autores.

Assim, o equacionamento da demanda de energia elétrica no Rio Grande do Sul partiu das funções representadas pela Eq. 3.1 para o setor residencial, Eq. 3.2 para o setor comercial e Eq. 3.3 para o setor industrial.

$$DY_r = f(T_r; R_r; E_r; dummies) \quad (3.1)$$

$$DY_c = f(T_c; R_c; E_c; dummies) \quad (3.2)$$

$$DY_i = f(T_i; R_i; E_i; S_i; dummies) \quad (3.3)$$

Inicialmente os dados das variáveis independentes do modelo foram deflacionamento, por serem séries econômicas. Depois da alteração da métrica dos dados de todas as variáveis pela transformação logarítmica, com o objetivo de transformar o modelo multiplicativo em modelo aditivo, foi realizada uma análise exploratória com a construção do gráfico em linhas para analisar e comparar o comportamento das séries estudadas. Para traçá-lo foi necessário adotar o ano de 1971 como ano de referência, atribuindo-lhe o valor 100 para todas as variáveis. Os demais valores observados foram determinados em relação a este.

Após foram aplicados os testes de Dickey-Fuller Aumentado (ADF) e de Kwiatkowski, Phillips, Schmidt e Shin (KPSS) para avaliar a sua estacionariedade,

seguindo o mesmo procedimento descrito anteriormente. Na constatação de não estacionaridade foram realizadas diferenças sucessivas com repetição dos procedimentos descritos na etapa anterior, até consegui-la em cada uma das variáveis.

Na sequência ocorreu a especificação do modelo VAR com identificação do número de defasagens, a presença de termos determinísticos e a presença de variáveis dummy, identificando modelos candidatos. Nesta etapa foram utilizados os critérios de informação de Akaike, Schwarz, Hannan-Quinn. Seguindo com a aplicação do modelo de Johansen executando os testes do traço e os testes do máximo autovalor para verificar se existe vetor de cointegração entre as variáveis consideradas no modelo, sugerindo existência de relação de longo prazo entre elas. No teste traço iniciou-se testando a hipótese nula de que não há vetores de cointegração contra a alternativa de que existe pelo menos um vetor de cointegração; em sequência, testou-se a hipótese nula de que existe no máximo um vetor de cointegração contra a alternativa de que existe mais de um, e assim sucessivamente. No teste do máximo autovalor iniciou-se testando a hipótese nula de que não há vetores de cointegração contra a alternativa de que há um vetor de cointegração; em sequência, testou-se a hipótese nula de que existe um vetor de cointegração contra a alternativa de que existem dois, e assim sucessivamente. Deste forma, foi possível realizar a identificação da equação de cointegração, avaliando a significância estatística dos coeficientes. Em seguida ocorreu a identificação de um novo modelo com o método de correção de erros vetoriais (VEC) para descrever a relação de curto prazo entre as variáveis, permitindo fazer previsões.

Após veio a avaliação das suposições de existência de autocorrelação nos resíduos (teste Ljung-Box), normalidade (teste Jarque-Bera) e homogeneidade entre variâncias (teste ARCH), como executado no protocolo anterior. Isto foi realizado para os modelos VAR e VAR-VEC, respectivamente. Por último foi realizada a validação dos modelos VAR e VAR-VEC para previsão com a determinação dos indicadores: erro absoluto médio (MAE), erro quadrático médio (RMSE) e erro absoluto percentual médio (MAPE), além da estatística U de Theil para os período 2011-2017, com identificação dos melhores modelos para previsão. Neste momento foi desconsiderado o efeito da inflação.

Os procedimentos descritos nos dois protocolos foram realizados com o auxílio dos softwares Gnu Regression, Econometrics and Times-Séries Library (Gretl, 2018), R (R CORE TEAM, 2018) e da planilha Excel. Nas análises inferenciais foram considerados os níveis de significância de 0,01 e 0,05.

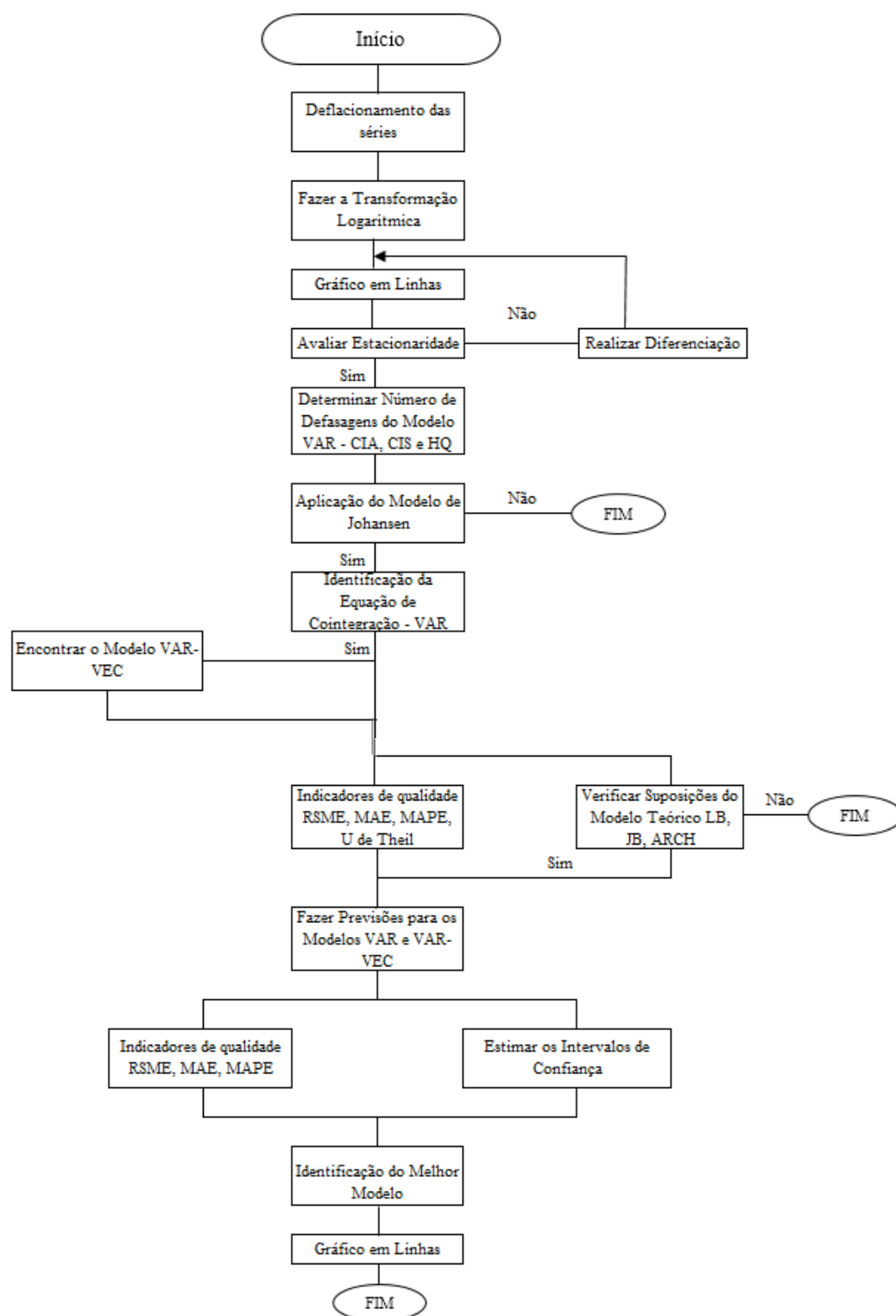


Figura 3.2 - Fluxograma da metodologia VAR-VEC

4. RESULTADOS

4.1. Residencial

4.1.1. Metodologia Box-Jenkins

O gráfico dos dados em nível para a demanda residencial de energia elétrica no Rio Grande do Sul pode ser visto na Fig. 4.1, em que é possível identificar uma tendência exponencial com uma oscilação entre os anos 1990 e 1994, período da renegociação da dívida do estado, com altos juros, e entre os anos 2000 e 2005, período da uma crise energética no país. Estas alterações no comportamento da série podem influenciar negativamente sua análise, sendo suas consequências reduzidas com o uso da transformação logarítmica que estabiliza a variância. Este procedimento também foi adotado por diversos autores, como: Andrade e Lobão (1997), Schmidt e Lima (2004), Mattos (2004), Leite (2006), Siqueira et al. (2006), Irff et al. (2009), Viana (2010), Viana, Lima (2011) e Silva (2014) e Silveira (2017).

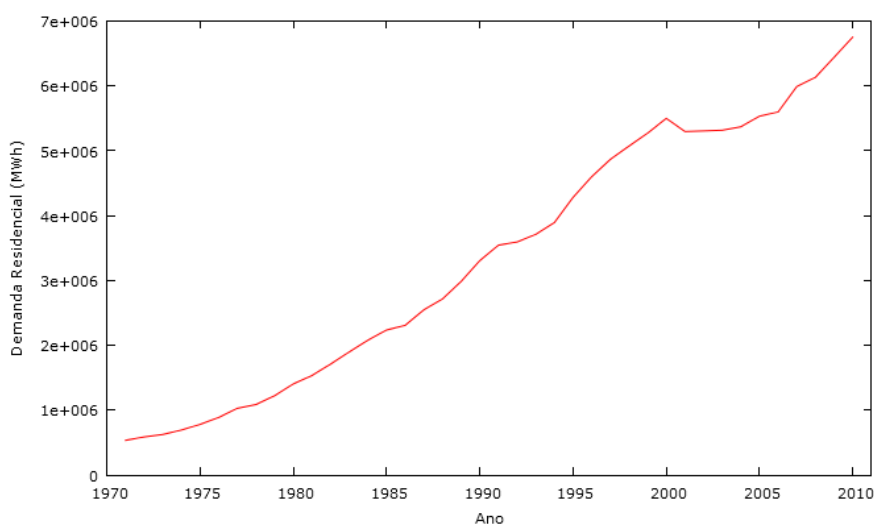


Figura 4.1-Demanda residencial de energia elétrica no Rio Grande do Sul no período 1971 - 2010

Na Figura 4.2.a é possível visualizar o gráfico do logaritmo da demanda residencial, enquanto na Fig. 4.2.b, os gráficos FAC (superior) e FACP (inferior) destas observações, tomando limites com confiança de 95%. A lenta queda exponencial na FAC com muitas correlações significativas caracteriza a não estacionariedade.

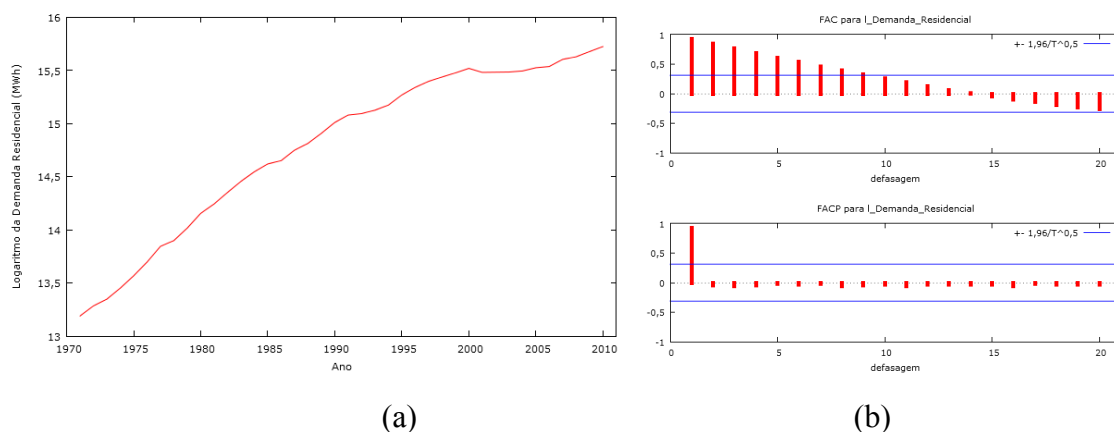


Figura 4.2 - Gráfico do logaritmo da demanda residencial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)

No teste ADF, aplicado para um número máximo de dez defasagens, não foram encontradas evidências para rejeitar a hipótese nula de que a série não é estacionária (*valor $p > 0,01$*) para modelo com tendência e constante. Já no teste KPSS, foram encontradas evidências para rejeitar a hipótese nula de que a série é estacionária (*valor $p < 0,01$*), o que fortalece a ideia de que a série deve ser considerada como não estacionária.

Diante do exposto, parte-se para análise do comportamento da primeira diferença. O gráfico em linhas (a) e os gráficos FAC e FACP (b) da primeira diferença do logaritmo das observações são mostrados na Fig. 4.3, tomando limites de confiança de 95%. No gráfico (a) desta figura, é possível identificar uma leve tendência decrescente entre os anos de 1980 e 2001 e crescente a partir deste último ano. No gráfico (b) superior da Fig. 4.3, observa-se que a FAC apresenta uma queda lenta com muitas correlações significativas e um comportamento senoidal, enquanto no gráfico (b) inferior desta mesma figura, a FACP decai abruptamente na primeira defasagem com oscilações, o que, segundo Morettin (2006), caracteriza um modelo AR(1).

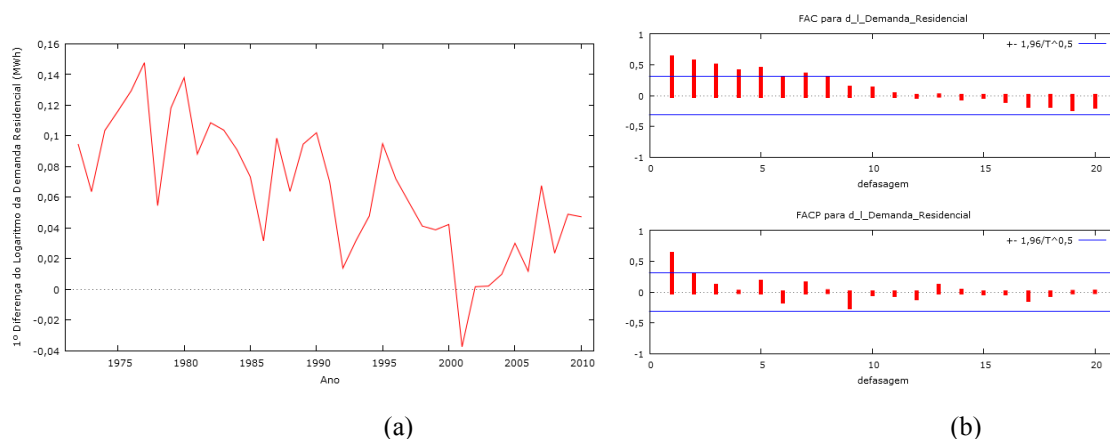


Figura 4.3 - Gráfico da primeira diferença do logaritmo da demanda residencial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)

No teste ADF, aplicado para um número máximo de dez defasagens, foram encontradas evidências para rejeitar a hipótese nula de que a série não é estacionária (*valor $p < 0,01$*) para modelo com tendência e constante, enquanto no teste KPSS não foram encontradas evidências para rejeitar a hipótese nula de que a série é estacionária (*valor $p > 0,10$*). Logo, ela pode ser considerada estacionária.

Para se modelar a primeira diferença, sugere-se um modelo ARIMA (1, 1, 0). Assim, este modelo foi avaliado, admitindo-se variações de ± 1 na ordem do componente auto regressivos e de $+1$ na ordem do componente de médias móveis. Ao parâmetro que assumiu valor zero, apenas foi acrescentado 1. Neste caso, os modelos considerados como candidatos foram: ARIMA (0,1,0), ARIMA (1,1,0), ARIMA (2,1,0), ARIMA (0,1,1), ARIMA (1,1,1), ARIMA (2,1,1). Esses modelos foram estimados com e sem constante e a significância de seus coeficientes está apresentada na Tab. 4.1, em que é possível constatar que os modelos ARIMA (0,1,0), ARIMA (0,1,1), ARIMA (1,1,1), ARIMA (1,1,0) e ARIMA (0,1,2) apresentaram todos os seus coeficientes significativos a um nível de significância de 0,01, não havendo interferência da constante. Com a inserção de uma constante, os modelos ARIMA (0,1,0), ARIMA (0,1,1) e ARIMA (1,1,0) apresentaram todos os seus coeficientes significativos a um nível de significância de 0,01. Estendendo o nível de significância para 0,05, em modelo sem constante, o modelo ARIMA (2,1,0) apresentou seus coeficientes significativos. Já com a inserção de uma constante, o modelo ARIMA (1,1,1) apresenta coeficientes significativos.

Tabela 4.1 – Significância dos coeficientes dos modelos selecionados

ARIMA (p,1,q)	Coeficiente	valor p	
		Sem constante	Com constante
(0,1,0)	C	-	< 0,01*
(1,1,0)	C	-	< 0,01*
	ϕ_1	< 0,01*	< 0,01*
(2,1,0)	C	-	< 0,01*
	ϕ_1	< 0,01*	< 0,01*
	ϕ_2	0,0108**	0,0694
(0,1,1)	-	-	< 0,01*
	θ_1	< 0,01*	< 0,01*
(1,1,1)	C	-	< 0,01*
	ϕ_1	< 0,01*	< 0,01*
	θ_1	< 0,01*	0,0126**
(2,1,1)	C	-	< 0,01*
	ϕ_1	< 0,01*	0,0215
	ϕ_2	0,6026	0,8684
	θ_1	< 0,01*	0,1364

Obs: * indica significativo para nível de significância 0,01. ** indica significativo para nível de significância 0,05.

Os resultados obtidos para avaliação da parcimônia dos modelos candidatos pelos critérios de informação de Akaike, Schwarz e Hannan-Quinn estão apresentados na Tab. 4.2. De acordo com estes critérios de informação, o melhor modelo foi o ARIMA (1,1,1) (CIS = -145,227, CIA = -151,882, HQ = -149,494), seguido pelos modelos ARIMA (2,1,0), ARIMA (1,1,0), ARIMA (2,1,1), ARIMA (0,1,1) e ARIMA (0,1,0), todos com o acréscimo de constante.

Tabela 4.2 - Resultados dos critérios de informação nos modelos candidatos ARIMA (p,1,q)

Modelo	Critério de informação		
	Schwarz	Akaike	Hannan-Quinn
(1,1,1)	-145,227	-151,882	-149,494
(2,1,0)	-144,298	-150,953	-148,565
(1,1,0)	-144,824	-149,814	-148,024
(2,1,1)	-141,593	-149,911	-146,927
(0,1,1)	-137,113	-142,104	-140,313
(0,1,0)	-129,637	-132,964	-131,770

Em sequência, foi realizada a averiguação da presença de autocorrelação, normalidade e homoscedasticidade entre variâncias nos resíduos. Os resultados,

apresentados na Tab. 4.3, para um nível de significância de 0,01, mostram que todos os modelos cumprem todas as suposições do modelo teórico, com exceção dos modelos ARIMA (0,1,1) e ARIMA (0,1,0) que apresentam autocorrelação em seus resíduos.

Tabela 4.3 - Resultados dos testes de Ljung-Box, Jarque-Bera e ARCH para os resíduos dos modelos candidatos ARIMA (p,1,q)

Modelo	Ljung Box		Jarque-Bera		ARCH	
	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor
(1,1,1)	10,71	0,9062	2,11	0,3479	12,75	0,3870
(2,1,0)	11,95	0,8498	1,67	0,4326	10,91	0,5364
(1,1,0)	18,59	0,4835	0,71	0,6997	8,68	0,7300
(2,1,1)	10,93	0,8604	2,08	0,3532	12,90	0,3762
(0,1,1)	36,22	0,0099	0,33	0,8481	4,69	0,9676
(0,1,0)	82,89	< 0,01	0,54	0,7647	6,34	0,8978

Assim, os modelos de melhor desempenho pelos critérios de informação, significância dos coeficientes e cumprimento aos pressupostos do modelo teórico foram ARIMA (1,1,1), ARIMA (2,1,0) e ARIMA (1,1,0).

A validação destes modelos, foi realizada por meio das medidas: raiz do erro quadrático médio (RMSE), o erro absoluto médio (MAE), o erro absoluto percentual médio (MAPE) e o U de Theil, calculados com a métrica original e com a transformação logarítmica, cujos resultados estão apresentadas na Tab. 4.4.

Tabela 4.4 - Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão

Métrica	ARIMA (p,1,q)	RMSE	MAE	MAPE (%)	U de Theil
Logarítmica	(1,1,1)	0,0814	0,0613	0,3851	0,0051
	(2,1,0)	0,1023	0,0760	0,4779	0,0064
	(1,1,0)	0,1154	0,0861	0,5413	0,0073
Original	(1,1,1)	719.274,30	527.826,08	6,4639	0,0910
	(2,1,0)	917.584,79	663.296,41	8,1554	0,1161
	(1,1,0)	1.044.827,86	757.466,29	9,3194	0,1322

O modelo que apresentou os menores erros de previsão foi o modelo ARIMA (1,1,1), seguido pelos modelos ARIMA (2,1,0) e ARIMA (1,1,0).

Complementando a análise, foram construídos intervalos de confiança para os valores previstos, com grau de confiança de 95%. A Tab. 4.5 apresenta os intervalos de confiança e os valores observados, pela métrica original e logarítmica, evidenciando que todos os intervalos de confiança contém os valores observados.

Tabela 4.5 – Intervalos de confiança para o modelo ARIMA (1,1,1) de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC -95%
Logarítmica	2011	15,755	15,774	15,713; 15,834
	2012	15,808	15,824	15,721; 15,927
	2013	15,863	15,876	15,729; 16,023
	2014	15,958	15,929	15,738; 16,120
	2015	15,902	15,984	15,747; 16,220
	2016	15,973	16,039	15,757; 16,322
	2017	15,935	16,096	15,768; 16,424
Original	2011	6.956.044,0	7.119.912,5	6.671.751,7; 7.530.289,2
	2012	7.336.379,0	7.521.574,1	6.722.266,9; 8.262.741,6
	2013	7.750.414,0	7.965.171,2	6.778.544,8; 9.088.414,2
	2014	8.516.881,0	8.448.563,6	6.837.597,6; 10.022.057,8
	2015	8.054.982,7	8.973.622,5	6.900.607,4; 11.074.680,2
	2016	8.275.110,0	9.541.306,7	6.969.186,1; 12.257.456,5
	2017	8.327.298,0	10.153.398,1	7.044.748,5; 13.582.584,8

As Figuras 4.4 e 4.5 ilustram o modelo de previsão selecionado, com as métricas logarítmica e original, respectivamente.

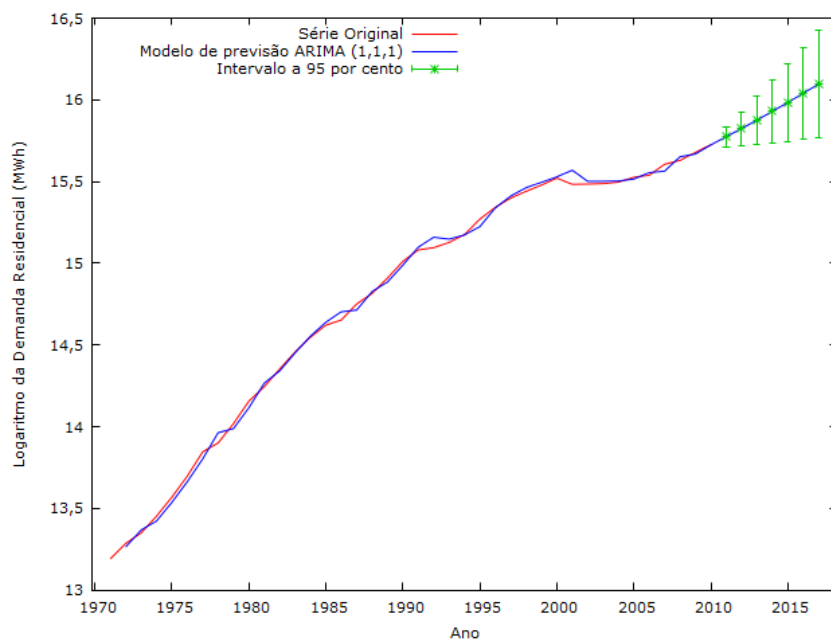


Figura 4.4 – Gráfico do logaritmo da demanda residencial de energia elétrica/RS com modelo de previsão ARIMA (1,1,1)

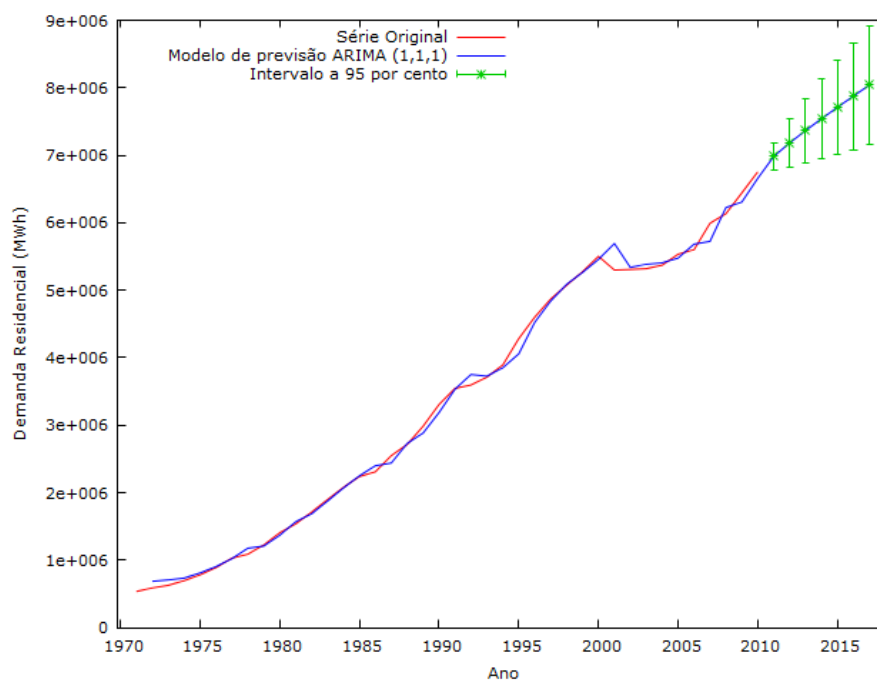
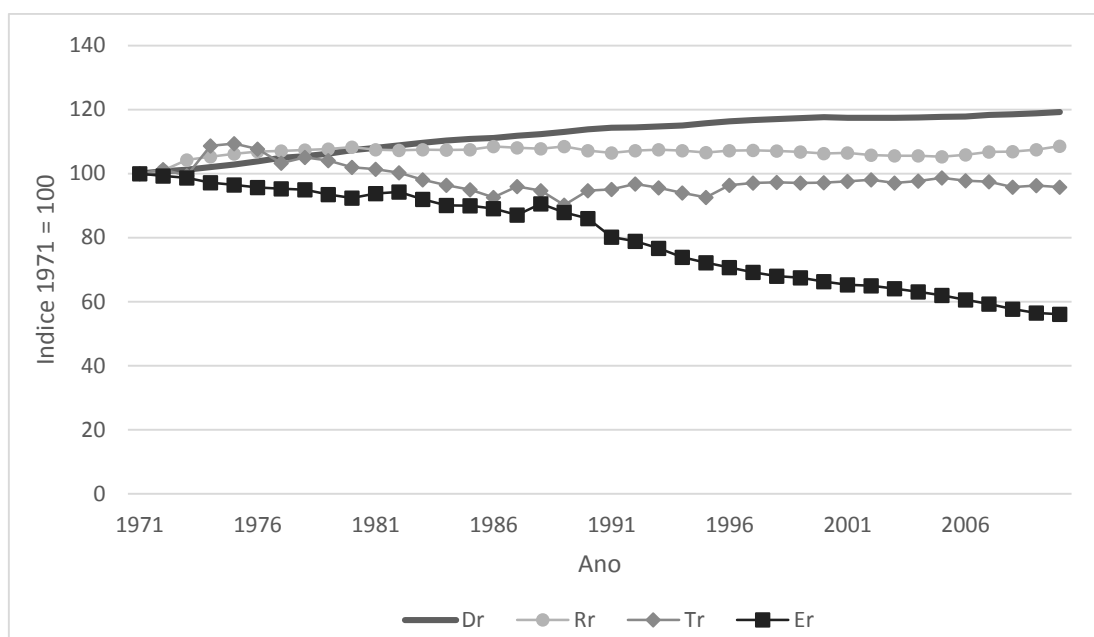


Figura 4.5 – Gráfico da série em nível da demanda residencial de energia elétrica/RS com modelo de previsão ARIMA (1,1,1)

4.1.2. Metodologia VAR-VEC

Gráfico em linhas

Após o deflacionamento das variáveis independentes e a aplicação da transformação logarítmica em todas as variáveis, foi elaborado o gráfico em linhas da Fig. 4.6. Para que seja possível observar o comportamento das séries ao longo do tempo em um mesmo gráfico, optou-se por fixar o primeiro valor observado no valor 100.



Fonte: Dados de Eletrobras (2018).

Obs: Dr = demanda; Rr = renda; Tr = tarifa; Er = preço dos eletrodomésticos.

Figura 4.6 – Comportamento das séries temporais deflacionadas ligadas à demanda residencial

Na Figura 4.6 é possível observar que no período de 1986 à 1994, quando ocorreram mudanças na moeda brasileira, saindo de Cruzeiro para Cruzado, para Cruzado Novo, para o Cruzeiro, para o Cruzeiro Real e, finalmente, para o Real, houve uma certa oscilação econômica, com variação no valor da tarifa e no preço dos eletrodomésticos. Outro fato observado no período analisado é a redução no valor do preço dos eletrodomésticos, o mesmo acontecendo com o valor da tarifa, só que de maneira menos acentuada. Por outro lado, a demanda de energia elétrica cresceu no período estudado, assim como a renda.

Teste de normalidade

Após análise gráfica foram realizados os testes de raiz unitária de Dickey-Fuller Aumentado (ADF) e Kwiatkowski–Phillips–Schmidt–Shin (KPSS) em cada uma das variáveis em nível e em primeira diferença, sendo avaliada a estacionariedade de cada série, além da determinação de sua ordem de integração, condição necessária para a realização dos testes de cointegração. Os resultados dos testes ADF são mostrados na Tab. 4.6, enquanto os resultados dos testes KPSS são mostrados na Tab. 4.7.

Tabela 4.6 – Teste ADF no setor residencial

Variável	Termos na equação	Nº de defasagens	Valor-p
Dr – Demanda	Constante e tendência	1	0,9322
Δ Dr	Constante e tendência	0	0,0034
Rr – Renda	Sem constante	8	0,3602
Δ Rr	Sem constante	7	0,0182
Tr – Tarifa	Sem constante	0	0,5520
Δ Tr	Sem constante	0	< 0,01
Er – Eletrodomésticos	Constante e tendência	0	0,6350
Δ Er	Constante e tendência	0	< 0,01

Tabela 4.7 – Teste KPSS no setor residencial

Variável	Termos da equação	Estatística de teste	Valores críticos (0,05)	Valores críticos (0,01)
Dr	Constante e tendência	0,28	0,149	0,212
Δ Dr	Constante e tendência	0,07	0,149	0,212
Rr	Constante e tendência	0,24	0,149	0,212
Δ Rr	Constante	0,25	0,462	0,719
Tr	Constante	0,19	0,149	0,212
Δ Tr	Constante	0,09	0,462	0,719
Er	Constante e tendência	0,17	0,149	0,212
Δ Er	Constante	0, 17	0,462	0,719

Obs: número de defasagens = 3.

A partir dos resultados apresentados na Tab. 4.6, constata-se que, pelo teste ADF,

todas as variáveis podem ser consideradas não estacionárias em nível para um nível de significância de 0,01 e podem ser consideradas estacionárias em primeira diferença se for flexibilizado o nível de significância para 0,05. Com base nos resultados do teste KPSS, apresentados na Tabela 4.7, pode-se concluir que, considerando um nível de significância de 0,01, os dados estudados apresentaram indícios de não estacionários apenas em nível, por se rejeitar a hipótese nula de que a série é estacionária. Assim, as séries estudadas se tornaram estacionárias em primeira diferença, por não se rejeitar esta hipótese nula. Individualmente as séries podem ser consideradas estacionárias em primeira diferença, tanto pelo teste ADF como pelo teste KPSS.

Teste de Cointegração

Para a construção do modelo VAR é necessário determinar o número de defasagens a serem consideradas, a presença de variáveis dummy e de termos determinísticos (constante e tendência), o que é feito pelos critérios de informação de Schwartz, Hannan-Quinn e Akaike. As variáveis dummy analisadas foram as apresentadas no capítulo 3, identificadas por *DYr1*, *DYr2*, *DYr3*, *DYr4*, *DYr5*, *DYr6*. Os resultados desta análise estão apresentados na Tab. 4.8.

Tabela 4.8 – Seleção de defasagens do modelo VAR (continua)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Dr; Rr; Tr; Er	Constante e	1	-3,8182*	-2,7733*	-3,4498*
	Tendência**	2	-3,7434	-2,0019	-3,1294
	Constante	1	-3,2621*	-2,4002*	-2,9555*
		2	-3,1082	-1,5568	-2,5562
	Sem	1	-3,0484*	-2,3589*	-2,8031*
	Constante	2	-2,8455	-1,4665	-2,3549
	Constante e	1	-3,1820*	-1,9754*	-2,7527*
	Tendência	2	-2,9784	-1,0822	-2,3038
	Constante**	1	-3,3115*	-2,2773*	-2,9436*
		2	-3,1511	-1,4273	-2,5378
DYr2	Sem	1	-3,0651*	-2,2033*	-2,7585*
	Constante	2	-2,9012	-1,3498	-2,3493

Tabela 4.9 – Seleção de defasagens do modelo VAR (continua)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Dr; Rr; Tr; Er; DYr1	Constante e	1	-3,1927*	-1,9860*	-2,7634*
	Tendência**	2	-3,0393	-1,1431	-2,3646
	Constante	1	-3,0893*	-2,0551*	-2,7213*
		2	-2,9204	-1,1966	-2,3071
	Sem	1	-2,8757*	-2,0138*	-2,5690*
	Constante	2	-2,6622	-1,1108	-2,1102
Dr; Rr; Tr; Er; DYr3	Constante e	1	-3,0117*	-1,8051*	-2,5824*
	Tendência	2	-2,8155	-0,9193	-2,1409
	Constante**	1	-3,1059*	-2,0716*	-2,7379*
		2	-2,9512	-1,2275	-2,3379
	Sem	1	-2,8772*	-2,0154*	-2,5706*
	Constante	2	-2,6727	-1,1213	-2,1207
Dr; Rr; Tr; Er; DYr4	Constante e	1	-3,0472*	-1,8406*	-2,6179*
	Tendência	2	-2,8283	-0,9322	-2,1537
	Constante**	1	-3,0980*	-2,0637*	-2,7300*
		2	-2,9345	-1,2107	-2,3212
	Sem	1	-2,8849*	-2,0230*	-2,5783*
	Constante	2	-2,6757	-1,1243	-2,1237
Dr; Rr; Tr; Er; DYr6	Constante e	1	-2,9797*	-1,7731*	-2,5504*
	Tendência	2	-2,7824	-0,8862	-2,1077
	Constante**	1	-3,0947*	-2,0604*	-2,7267*
		2	-2,9372	-1,2134	-2,3239
	Sem	1	-2,8737*	-2,0119*	-2,5670*
	Constante	2	-2,6863	-1,1349	-2,1343
Dr; Rr; Tr; Er; DYr1; DYr2	Constante e	1	-3,2057*	-1,8267*	-2,7150*
	Tendência**	2	-2,9999	-0,9314	-2,2639
	Constante	1	-3,1416*	-1,9350*	-2,7123*
		2	-2,9645	-1,0684	-2,2899
	Sem	1	-2,8924*	-1,8581*	-2,5244*
	Constante	2	-2,7174	-0,9936	-2,1041

Tabela 4.10 – Seleção de defasagens do modelo VAR (conclusão)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Dr; Rr; Tr; Er; DYr5	Constante e	1	-3,1418*	-1,9352*	-2,7125*
	Tendência**	2	-2,9536	-1,0574	-2,2789
	Constante	1	-3,0789*	-2,0446*	-2,7109*
		2	-2,9199	-1,1961	-2,3066
	Sem	1	-2,8873*	-2,0254*	-2,5807*
	Constante	2	-2,6995	-1,1481	-2,1475

Obs: 1 - * indica o menor valor obtido pelo critério; 2 - ** indica o termo da equação que apresentou os menores valores dos critério de informação.

De acordo com os resultados dos critérios de informação, apresentados na Tabela 4.8, o modelo mais parcimonioso não utiliza variáveis dummy, apresenta constante e tendência determinística, além de utilizar apenas uma defasagem.

Para determinar o posto da matriz de cointegração pelos testes de Johansen, foram consideradas as situações: sem constante, com constante apenas no vetor de cointegração (restrita) ou fora do vetor de cointegração (irrestrita) e com tendência restrita ou irrestrita. Os resultados obtidos para o teste do traço e o teste do máximo autovalor para o modelo escolhido que utilizou constante sem restrições podem ser visualizados na Tab. 4.9 e Tab. 4.10.

Tabela 4.11 – Teste do traço de Johansen para o modelo com Dr, Rr, Tr, Er e constante sem restrição

Teste	Hipótese nula (H_0)	Hipótese alternativa (H_1)	Valor-p
1º	$r = 0$	$r > 0$	0,0022
2º	$r \leq 1$	$r > 1$	0,0441
3º	$r \leq 2$	$r > 2$	0,0628
4º	$r \leq 3$	$r > 3$	0,0867

De acordo com a Tabela 4.9, o primeiro teste do traço indica a rejeição da hipótese nula de que não existe relação de longo prazo entre as variáveis, encontrando evidências de que exista pelo menos um vetor de cointegração entre elas. Por outro lado, no segundo

teste do traço, a hipótese de que $r \leq 1$ não pode ser rejeitada para um nível de significância de 0,01, indicando a existência de no máximo um vetor de cointegração entre as variáveis.

Tabela 4.12 – Teste do máximo autovalor de Johansen para o modelo Dr, Rr, Tr, Er com constante sem restrição

Teste	Hipótese nula (H0)	Hipótese alternativa (H1)	Valor-p
1º	$r = 0$	$r = 1$	0,0082
2º	$r = 1$	$r = 2$	0,2026
3º	$r = 2$	$r = 3$	0,1060
4º	$r = 3$	$r = 4$	0,0748

O primeiro teste do máximo autovalor, cujos os resultados estão apresentados na Tab. 4.10, indica rejeição da hipótese de que não existe vetor de cointegração entre as variáveis, corroborando com o resultado do teste do traço. Ao testar a hipótese nula de que $r = 1$, o teste indica que a mesma não pode ser rejeitada. Assim sendo, existem evidências que há apenas um vetor de cointegração entre as séries das variáveis relacionadas a demanda residencial de energia elétrica a um nível de significância de 0,01.

Deste modo, os testes de cointegração indicam que as séries são cointegráveis, sugerindo a existência de relação de longo prazo entre elas. Uma vez que o objetivo deste trabalho é determinar a influência das variáveis estudadas na demanda de energia elétrica, optou-se por trabalhar com o vetor de cointegração normalizado em relação a variável Dr. O vetor normalizado, resultante do procedimento de Johansen é apresentado na Tab. 4.11.

Tabela 4.13 – Vetor de cointegração normalizado - Residencial

Dr	Rr	Tr	Er
1	-0,17141	4,1806	2,9280
Desvio Padrão	(0,25885)	(3,0649)	(0,68065)

A partir da Tabela 4.11 é possível escrever a equação que expressa a relação de longo prazo entre as variáveis, conforme pode ser observado na Eq. 4.1.

$$Dr = 0,17141Rr - 4,1806Tr - 2,9280Er \quad (4.1)$$

Segundo a Equação 4.1, os coeficientes dos parâmetros estão de acordo com o esperado pela teoria econômica, ou seja: a variável renda apresenta sinal positivo, e as variáveis tarifa e preço de eletrodomésticos exibiram sinais negativos. Além do mais, a demanda de energia elétrica residencial sofre mais alterações com a variação da tarifa, seguido pela alteração do preço dos eletrodomésticos, do que com a alteração da renda. Este comportamento talvez possa ser explicado por meio da seguinte análise: com a elevação das tarifas e dos preços dos eletrodomésticos o consumidor tende a consumir menos energia elétrica.

A análise da qualidade do modelo encontrado por meio dos resíduos foi realizada pelo teste de Jarque-Bera para avaliar a normalidade, pelo teste Ljung-Box Q para avaliar a autocorrelação e pelo teste ARCH para avaliar a heterocedasticidade. Os resultados desta análise são mostrados na Tab. 4.12.

Tabela 4.14 – Testes de qualidade do modelo VAR para o setor residencial

Teste	Estatística de Teste	Valor-p
Jarque-Bera	0,53	0,7674
Ljung-Box Q'	21,97	0,2856
ARCH	123,97	0,0524

Os resultados apresentados na Tabela 4.12 indicam que os resíduos do modelo são normalmente distribuídos, pois não é possível rejeitar a hipótese nula de que os dados provêm de uma população com distribuição normal; não apresentam autocorrelação, porque a hipótese de que não existe autocorrelação entre os resíduos não pode ser rejeitada; e não apresentam heterocedasticidade, uma vez que a hipótese da ausência do efeito ARCH não pode ser rejeitada.

Todos os cálculos e procedimentos executados até o momento permitiram encontrar a relação de longo prazo entre as variáveis estudadas. Entretanto, o conhecimento dessa relação nem sempre é suficiente para que seja possível realizar previsões de curto prazo.

Modelo de correção de erros

Com a finalidade de realizar os ajustes dos modelos para que se possa determinar a variação de curto prazo entre as variáveis e realizar previsões, é aplicado um modelo de correção de erros vetorial (VEC). Nesse modelo, o vetor de cointegração de longo prazo é utilizado como variável explicativa (u_t) do termo de correção de erros. O resultado do VEC é expresso pela Eq. 4.2.

$$\Delta Dr_t = 0,4025 - 0,0081(u_{t-1}) \quad (4.2)$$

O modelo apresentou coeficiente de determinação $R^2 = 0,5407$ e $R^2_{ajustado} = 0,5283$, sugerindo que alterações nas variáveis consideradas explicam 54,07% das variações da demanda de energia elétrica residencial.

O coeficiente do termo erro na Eq. 4.2 indica que a velocidade de ajustamento em relação ao equilíbrio de longo prazo é 0,0081, ou seja, 0,81% da diferença entre as relações de longo e curto prazo é corrigida a cada ano.

Os resultados da análise das suposições deste modelo são apresentados na Tab. 4.13, constatando-se que os resíduos podem ser considerados como normalmente distribuídos, não apresentando autocorrelação e nem heterocedasticidade.

Tabela 4.15 – Testes das suposições do modelo VEC para o setor residencial

Teste	Estatística de Teste	Valor-p
Jarque-Bera	0,9604	0,6187
Ljung-Box Q'	15,6470	0,6807
ARCH	115,9020	0,1321

Previsões

Para validação do modelo, foi realizada a previsão para a demanda de energia elétrica para os anos de 2011 à 2017, utilizando apenas o modelo VAR e o modelo VAR em conjunto com o modelo VEC. A validação destes modelos foi realizada por meio das medidas: raiz do erro quadrático médio (RMSE), o erro absoluto médio (MAE), o erro absoluto percentual médio (MAPE) e o U de Theil, calculados com a métrica original e com a transformação logarítmica, cujos resultados estão apresentados na Tab. 4.14.

Tabela 4.16 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos VAR e VARVEC de previsão

Métrica	Modelo	RMSE	MAE	MAPE (%)	U de Theil
Logarítmica	VAR	0,0724	0,0616	0,3871	0,0046
	VAR-VEC	0,1130	0,1012	0,6359	0,0071
Original	VAR	570.437,92	481.960,55	5,9077	0,0722
	VAR-VEC	871.697,50	774.033,87	9,5082	0,1103

Visando complementar a análise, foram construídos intervalos para os valores previstos, com nível de confiança de 95%. As Tab. 4.15 e 4.16 apresentam os intervalos de confiança e os valores observados, pela métrica original e logarítmica, para os modelos VAR e VAR-VEC respectivamente, evidenciando que todos os intervalos de confiança contém os valores observados, com exceção dos valores no ano de 2014. Esta diferença pode ter ocorrido devido ao baixo poder de ajuste do termo de correção de erros do modelo VEC residencial, somente 0,81% por ano.

Tabela 4.17 – Intervalos de confiança para o modelo VAR de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC –95%
logarítmica	2011	15,7551	15,7510	15,6949; 15,8071
	2012	15,8084	15,7751	15,6968; 15,8535
	2013	15,8633	15,7979	15,7026; 15,8932
	2014	15,9576	15,8194	15,7099; 15,9288
	2015	15,9018	15,8396	15,7178; 15,9614
	2016	15,9287	15,8588	15,7259; 15,9916
	2017	15,9351	15,8769	15,7342; 16,0196
original	2011	6.956.044,0	6.927.600,1	6.549.504,8; 7.327.170,3
	2012	7.336.379,0	7.096.696,8	6.561.960,7; 7.675.162,3
	2013	7.750.414,0	7.260.011,7	6.600.130,6; 7.985.995,5
	2014	8.516.881,0	7.417.510,2	6.648.487,9; 8.275.148,1
	2015	8.054.982,7	7.569.314,0	6.701.218,9; 8.549.642,3
	2016	8.275.110,0	7.715.624,6	6.755.719,3; 8.811.779,8
	2017	8.327.298,0	7.856.624,6	6.812.025,1; 9.061.996,4

Tabela 4.18 – Intervalos de confiança para o modelo VAR-VEC de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC -95%
logarítmica	2011	15,7551	15,7419	15,6863; 15,7974
	2012	15,8084	15,7565	15,6774; 15,8355
	2013	15,8633	15,7694	15,6706; 15,8682
	2014	15,9576	15,7807	15,6629; 15,8985
	2015	15,9018	15,7903	15,6530; 15,9275
	2016	15,9287	15,7983	15,6405; 15,9560
	2017	15,9351	15,8046	15,6250; 15,9842
original	2011	6.956.044,0	6.864.433,0	6.493.699,8; 7.256.331,8
	2012	7.336.379,0	6.965.549,1	6.436.149,4; 7.538.501,6
	2013	7.750.414,0	7.056.205,5	6.392.378,6; 7.788.968,5
	2014	8.516.881,0	7.136.093,1	6.343.067,2; 8.028.265,0
	2015	8.054.982,7	7.204.972,7	6.280.863,3; 8.265.047,2
	2016	8.275.110,0	7.262.640,3	6.202.661,3; 8.503.768,7
	2017	8.327.298,0	7.308.977,9	6.107.487,3; 8.746.822,1

Os gráficos que representam os modelos de previsão VAR e VAR - VEC podem ser visualizados nas Figuras 4.7 e 4.8 respectivamente.

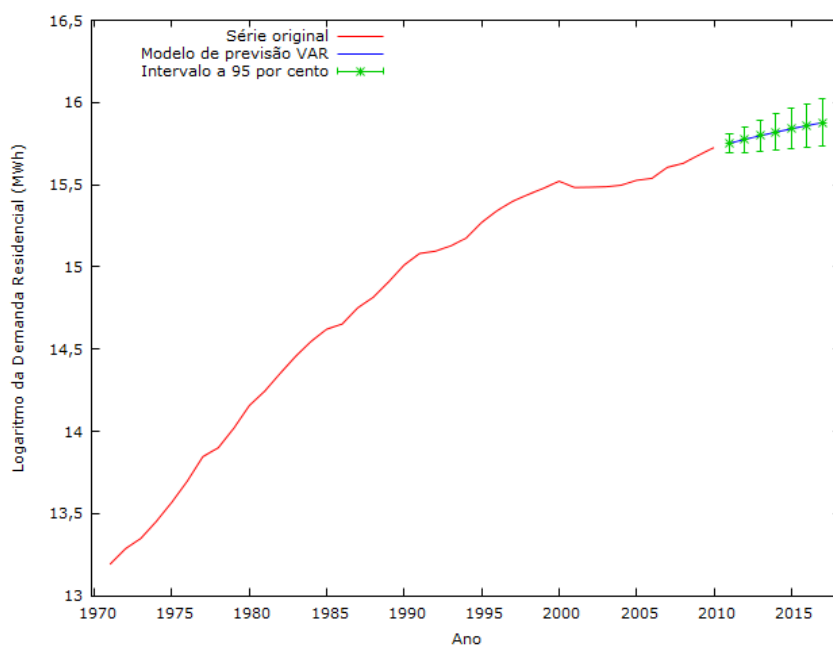


Figura 4.7 – Gráfico do modelo de previsão VAR para demanda residencial de energia elétrica/RS

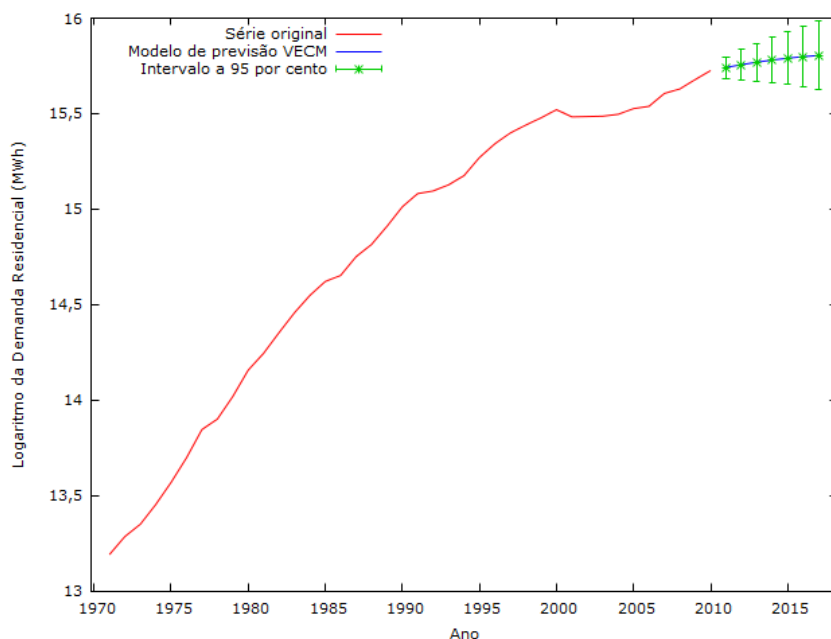


Figura 4.8 – Gráfico do modelo de previsão VAR-VEC para demanda residencial de energia elétrica/RS

4.1.3. Comparações para o setor Residencial

Técnica utilizada: ARIMA x VAR x VAR-VECM

Na análise dos resultados, tanto para a métrica original como para a métrica logarítmica, foi constatado que os valores previstos pelos modelos VAR e VAR-VEC ficaram sempre abaixo dos valores observados e forneceram intervalos de confiança de menor amplitude. Como consequência, um dos valores observados ficou fora do intervalo de confiança. Já os valores previstos pelo modelo ARIMA (1,1,1) ficaram acima dos valores observados, com exceção do ano de 2014, que ficou abaixo. Todos eles, entretanto, ficaram dentro dos intervalos de confiança.

Considerando os indicadores de qualidade (RMSE, MAE, MAPE e U de Theil) para os dados com a métrica original, o modelo VAR apresentou os melhores resultados, mas, se for considerada a métrica logarítmica, o modelo ARIMA (1,1,1) apresentou um MAE (0,0613) e MAPE (0,3851%) pouco abaixo do modelo VAR (MAE = 0,0616, MAPE = 0,3871%). Isto explica o porquê na análise com a métrica original, a

variabilidade retirada pela transformação logarítmica, causou uma inversão, deixando o modelo VAR melhor em todas as análises, inclusive um MAE (481.960,55) e um MAPE (5,91%) melhores do que o modelo ARIMA (1,1,1) (MAE = 527.826,08, MAPE = 6,46%). Uma pequena restrição que se pode fazer, diz respeito aos intervalos de confiança, nos quais o modelo ARIMA (1,1,1) talvez possa apresentar melhores resultados por ter todos os valores previstos dentro dos intervalos de confiança, embora eles tenham apresentado amplitude maior.

Comparação com o modelo Residencial de Silveira (2017)

A Tabela 4.17 apresenta os coeficientes para os modelos VAR encontrados em nível nacional por Silveira (2017) e neste estudo, visando facilitar a comparação.

Tabela 4.19 – Coeficientes do modelo VAR Residencial/RS e BR

	Dr	Rr	Tr	Er	Constante
Modelo RS	1	0,1714	-4,1806	-2,9280	-
Modelo BR	1	0,8098	-0,0802	-0,3550	-10,5280

O modelo encontrado para o setor residencial a nível nacional apresentou duas defasagens, já neste estudo apresentou uma defasagem. É possível notar uma elasticidade menor para a renda e maiores para a tarifa e preço dos eletrodomésticos, além da não utilização da constante.

A Tabela 4.18 mostra os coeficientes dos modelos VAR-VEC encontrados a nível nacional por Silveira (2017) e os encontrados neste estudo.

Tabela 4.20 – Coeficientes do modelo VAR-VEC Residencial/RS e BR

	ΔDr	$\Delta Dri-1$	$\Delta Rri-1$	Dr	u_i-1	Constante
Modelo RS	1	-	-	-	0,0081	0,4025
Modelo BR	1	0,318	-0,879	-0,077	-0,567	-

O modelo VAR-VEC residencial nacional apresentou com um vetor de cointegração e uma dummy que representa o período de quebra na série causada pela crise energética no ano de 2001. A renda e o preço dos eletrodomésticos não apresentaram significância. Já neste trabalho foi encontrado um vetor de cointegração. A validação do

modelo nacional de Silveira (2017) foi feita apenas para o próximo período, apresentando um MAPE de 3,33%. Neste estudo, considerando apenas o ano subsequente ao período utilizado para a construção do modelo, o modelo VAR-VEC apresentou um MAPE de 1,32%, mostrando que o modelo regional teve melhor desempenho na análise de curto prazo.

4.2. Comercial

4.2.1. Metodologia Box-Jenkins

O gráfico dos dados em nível para a demanda comercial de energia elétrica no Rio Grande do Sul pode ser visto na Fig. 4.9, em que é possível identificar uma tendência exponencial com uma queda no ano 1986 e retomada do crescimento no ano de 1987. Já entre os anos 2000 e 2002, houve uma forte redução do crescimento da demanda, voltando a ser razoável a partir do ano de 2003, razão pela qual também foi utilizada a transformação logarítmica. No ano de 2006 também houve uma pequena diminuição do crescimento.

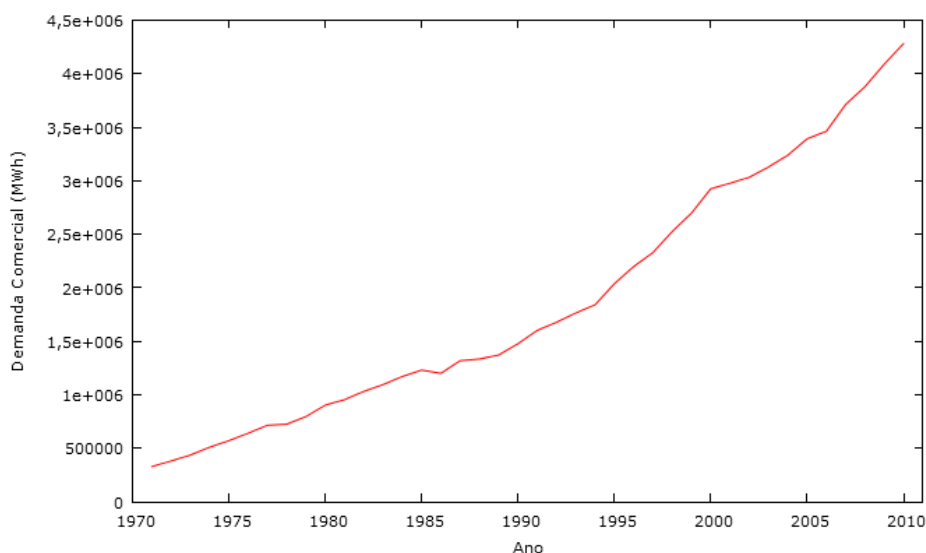


Figura 4.9 – Demanda comercial de energia elétrica no Rio Grande do Sul no período 1971 - 2010

Na Figura 4.10.a é possível visualizar o gráfico do logaritmo da demanda comercial, enquanto na Fig. 4.10.b, os gráficos FAC (superior) e FACP (inferior) destas observações, tomando limites com confiança de 95%. A lenta queda exponencial na FAC com muitas correlações significativas caracteriza a não estacionaridade.

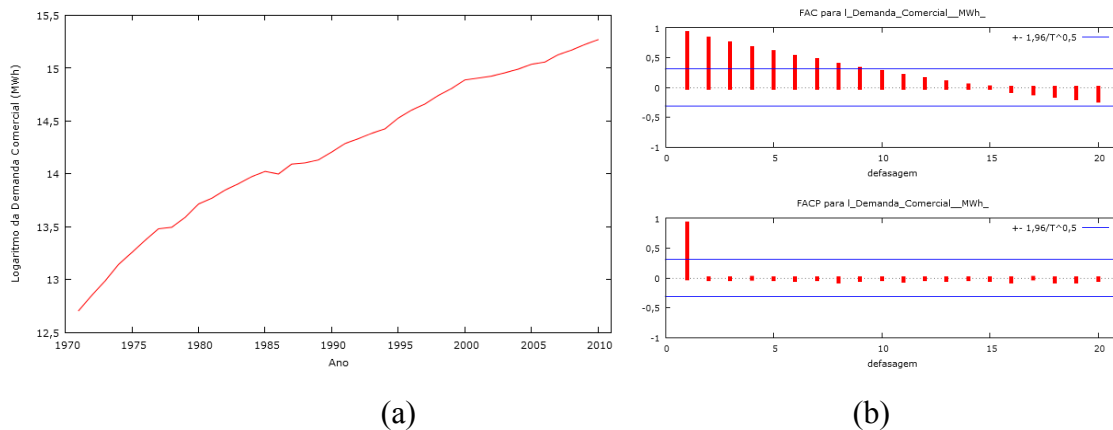


Figura 4.10 – Gráfico do logaritmo da demanda comercial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)

No teste ADF, aplicado para um número máximo de dez defasagens, não foram encontradas evidências para rejeitar a hipótese nula de que a série não é estacionária (*valor $p > 0,01$*) para modelo com tendência e constante. Já no teste KPSS, foram encontradas evidências para rejeitar a hipótese nula de que a série é estacionária (*valor $p < 0,01$*), o que fortalece a ideia de que a série deve ser considerada como não estacionária.

Diante do exposto, parte-se para análise do comportamento da primeira diferença. O gráfico em linhas (a) e os gráficos FAC e FACP (b) da primeira diferença do logaritmo das observações são mostrados na Fig. 4.11, tomando limites com confiança de 95%. No gráfico (a) desta figura, é possível identificar uma leve tendência decrescente entre os anos de 1970 e meados dos anos 1980, com oscilações a partir deste último ano, apresentando um comportamento aparentemente estacionário. No gráfico (b) superior da Fig. 4.11, observa-se que a FAC apresenta uma queda abrupta a partir da terceira defasagem, mantendo um comportamento senoidal. Já o gráfico inferior da Fig. 4.11.b apresenta um comportamento oscilatório com apenas a primeira correlação significativa. Segundo Morettin (2006), a série se caracteriza como um modelo AR(1). Por se modelar a primeira diferença, sugere-se um ARIMA (1, 1, 0).

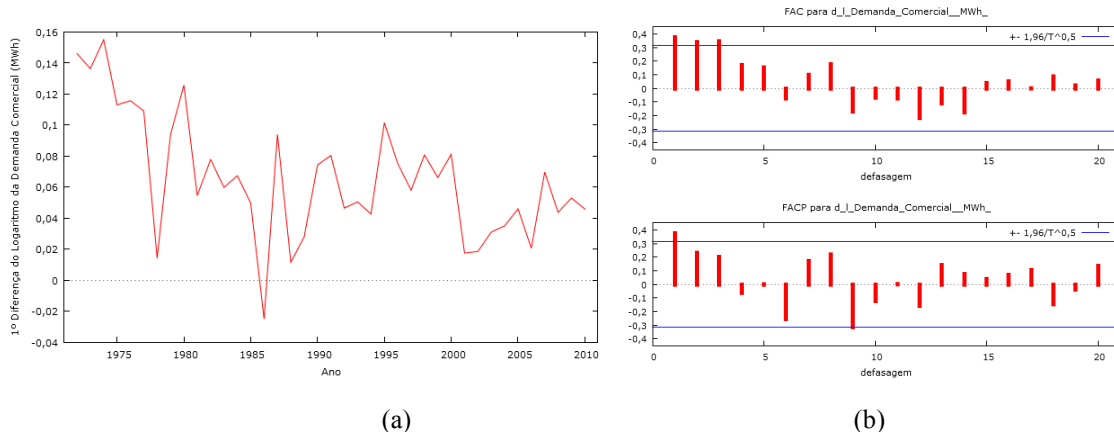


Figura 4.11 – Gráfico da primeira diferença do logaritmo da demanda comercial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)

No teste ADF, aplicado para um número máximo de dez defasagens, foram encontradas evidências para rejeitar a hipótese nula de que a série não é estacionária (*valor $p < 0,01$*) para modelo com tendência e constante, enquanto no teste KPSS não foram encontradas evidências para rejeitar a hipótese nula de que a série é estacionária (*valor $p > 0,10$*). Logo, ela pode ser considerada estacionária.

Assim, o modelo ARIMA (1,1,0) foi avaliado, admitindo-se variações de ± 1 na ordem do componente auto regressivo e +1 no componente de médias móveis. Neste caso, os modelos considerados como candidatos foram: ARIMA (0,1,0), ARIMA (1,1,0), ARIMA (2,1,0), ARIMA (0,1,1), ARIMA (1,1,1) e ARIMA (2,1,1).

Os modelos candidatos foram estimados com e sem constante e a significância de seus coeficientes está apresentada na Tab. 4.19, em que é possível constatar que, não havendo interferência da constante, os modelos ARIMA (0,1,0), ARIMA (1,1,0), ARIMA (2,1,0), ARIMA (0,1,1) e ARIMA (1,1,1) apresentaram todos os seus coeficientes significativos para um nível de significância de 0,01. Já com a inserção de uma constante, os modelos ARIMA(0,1,0), ARIMA (1,1,0) e ARIMA (1,1,1) apresentaram todos os seus coeficientes significativos para um nível de significância de 0,01. Para um nível de significância de 0,05, houve o acréscimo do modelo ARIMA (0,1,1) com constante.

Tabela 4.21 – Significância dos coeficientes dos modelos selecionados

ARIMA (p,1,q)	Coeficiente	valor p	
		Sem constante	Com constante
(0,1,0)	C	-	< 0,01*
(1,1,0)	C	-	< 0,01*
	ϕ_1	< 0,01*	< 0,01*
(2,1,0)	C	-	< 0,01*
	ϕ_1	< 0,01*	0,0388**
	ϕ_2	< 0,01*	0,0515
(0,1,1)	-	-	< 0,01*
	θ_1	< 0,01*	0,0371**
(1,1,1)	C	-	< 0,01*
	ϕ_1	< 0,01*	< 0,01*
	θ_1	< 0,01*	< 0,01*
(2,1,1)	C	-	< 0,01*
	ϕ_1	< 0,01*	< 0,01*
	ϕ_2	0,3544	0,3454
	θ_1	< 0,01*	0,0377**

Obs: * indica significativo para nível de significância 0,01. ** indica significativo para nível de significância 0,05.

Os resultados obtidos para avaliação da parcimonialidade dos modelos candidatos pelos critérios de informação de Akaike, Schwarz e Hannan-Quinn estão apresentados na Tab. 4.20. De acordo com estes critérios de informação, o melhor modelo foi o ARIMA (1,1,1) (CIS = -139,562, CIA = -144,552, HQ = -142,762), seguido pelos modelos ARIMA (2,1,1), ARIMA (2,1,0), ARIMA (1,1,0), ARIMA (0,1,1), e ARIMA (0,1,0), todos sem constante. Com o acréscimo da constante o melhor modelo foi o ARIMA (1,1,1), seguido pelos modelos ARIMA (2,1,0), ARIMA (2,1,1), ARIMA (1,1,0), ARIMA (0,1,1), e ARIMA (0,1,0).

Tabela 4.22 – Resultados dos critérios de informação nos modelos candidatos para modelos ARIMA (p,1,q) (continua)

Modelo	Critério de informação			
	Constante	Schwarz	Akaike	Hannan-Quinn
(1,1,1)	Sem	-139,562	-144,552	- 142,762
	Com	-138,381	-145,035	-142,648
(2,1,1)	Sem	-136,623	-143,277	-140,890
	Com	-135,327	-143,645	-140,660
(2,1,0)	Sem	-135,181	-140,171	-138,381
	Com	-136,591	-143,245	-140,857
(1,1,0)	Sem	-129,457	-132,784	-131,591
	Com	-137,124	-142,115	-140,324

Tabela 4.23 – Resultados dos critérios de informação nos modelos candidatos para modelos ARIMA (p,1,q) (conclusão)

Modelo	Critério de informação			
	Constante	Schwarz	Akaike	Hannan-Quinn
(0,1,1)	Sem	-105,216	-108,543	-107,349
	Com	-134,905	-139,896	-138,106
(0,1,0)	Sem	-86,434	-88,098	-87,501
	Com	-134,206	-137,533	-136,339

Em sequência, foi realizada a averiguação da presença de autocorrelação, normalidade e homoscedasticidade entre variâncias nos resíduos. Os resultados, apresentados na Tab. 4.21, mostram que todos os modelos cumpriram as suposições do modelo teórico a um nível de significância de 0,01. Entretanto, se for analisado a um nível de significância de 0,05, encontrou-se evidências de que os resíduos dos modelos ARIMA (1,1,1), ARIMA (2,1,0) e ARIMA (1,1,0) apresentam autocorrelação. O restante dos modelos atenderam as suposições do modelo teórico. Para a sequência do estudo, a análise foi feita sobre o nível de significância de 0,01.

Tabela 4.24 – Resultados dos testes de Ljung-Box, Jarque-Bera e ARCH para os resíduos dos modelos candidatos ARIMA (p,1,q)

Modelo	Ljung Box			Jarque-Bera		ARCH	
	Constante	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor
(1,1,1)	Sem	29,9225	0,0382	2,3620	0,3070	13,1336	0,3594
	Com	24,9733	0,1257	2,5096	0,2851	11,9615	0,4488
(2,1,1)	Sem	26,2453	0,0701	2,7596	0,2516	12,4475	0,4104
	Com	24,4824	0,1069	2,7394	0,2542	11,7152	0,4688
(2,1,0)	Sem	29,2881	0,0450	0,2097	0,9005	12,6522	0,3948
	Com	26,5422	0,0880	0,1996	0,9050	11,5421	0,4831
(1,1,0)	Sem	32,2051	0,0296	0,0986	0,9519	12,1190	0,4362
	Com	24,2073	0,1883	0,1082	0,9473	11,2115	0,5109
(0,1,1)	Sem	28,0880	0,0818	0,2194	0,8961	12,1223	0,4359
	Com	22,2676	0,2711	0,1320	0,9361	11,4419	0,4915
(0,1,0)	Sem	31,1778	0,0529	0,7508	0,6870	9,7104	0,6413
	Com	30,0048	0,0698	0,5794	0,7485	9,6487	0,6467

Assim, os modelos de melhor desempenho pelos critérios de informação, significância dos coeficientes e cumprimento aos pressupostos do modelo teórico foram

ARIMA (1,1,1) e ARIMA (2,1,0). Cabe destacar que, como o modelo ARIMA (2,1,1) não apresentou significância em um de seus coeficientes, foi substituído pelo modelo ARIMA (2,1,0).

A validação dos modelos encontrados foi realizada por meio das medidas: raiz do erro quadrático médio (RMSE), o erro absoluto médio (MAE), o erro absoluto percentual médio (MAPE) e o U de Theil, calculados com a métrica original e com a transformação logarítmica estão apresentadas na Tab. 4.22.

Tabela 4.25 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão

Métrica	ARIMA (p,1,q)	RMSE	MAE	MAPE (%)	U de Theil
Logarítmica	(2,1,0)	0,0716	0,0665	0,4309	0,0046
	(1,1,1)	0,0788	0,0696	0,4508	0,0051
Original	(2,1,0)	362.076,523	333.726,295	6,5900	0,0716
	(1,1,1)	405.825,703	353.297,292	7,0013	0,0803

O modelo que apresentou os menores erros de previsão foi o modelo ARIMA (2,1,0), seguido pelo modelo ARIMA (1,1,1).

A Tab. 4.23 apresenta os valores previstos para o período 2011-2017 com respectivos intervalos de confiança e seus valores de referência, pela métrica original e logarítmica. Observa-se que até 2015, os valores observados ficaram abaixo dos previstos, o mesmo não acontecendo em 2016 e 2017. Além disso, todos pertenceram aos respectivos intervalos de confiança.

Tabela 4.26 – Intervalos de confiança para o modelo ARIMA (2,1,0) de previsão pela métrica logarítmica e original (continua)

Métrica	Ano	Valor observado	Previsão	IC-95%
logarítmica	2011	15,3422	15,2914	15,2205; 15,3623
	2012	15,4193	15,3342	15,2080; 15,4603
	2013	15,4329	15,3759	15,1773; 15,5744
	2014	15,5124	15,4156	15,1392; 15,6920
	2015	15,4701	15,4539	15,0921; 15,8158

Tabela 4.27 – Intervalos de confiança para o modelo ARIMA (2,1,0) de previsão pela

métrica logarítmica e original (conclusão)

Métrica	Ano	Valor observado	Previsão	IC-95%
logarítmica	2016	15,4258	15,4907	15,0384; 15,9430
	2017	15,4314	15,5260	14,9785; 16,0735
original	2011	4.603.000,0	4.374.763,4	4.075.320,7; 4.696.208,3
	2012	4.972.000,0	4.566.013,2	4.024.913,6; 5.179.857,0
	2013	5.040.224,0	4.760.451,1	3.903.011,5; 5.806.253,0
	2014	5.456.817,0	4.953.539,8	3.757.367,0; 6.530.525,7
	2015	5.230.826,0	5.147.125,7	3.584.326,0; 7.391.320,8
	2016	5.004.438,0	5.339.812,0	3.397.039,9; 8.393.658,2
	2017	5.032.131,0	5.531.867,4	3.199.586,1; 9.564.223,8

As Figuras 4.12 e 4.13 ilustram o modelo de previsão selecionado e sua comparação com a série em nível, por meio da métrica logarítmica e original, respectivamente.

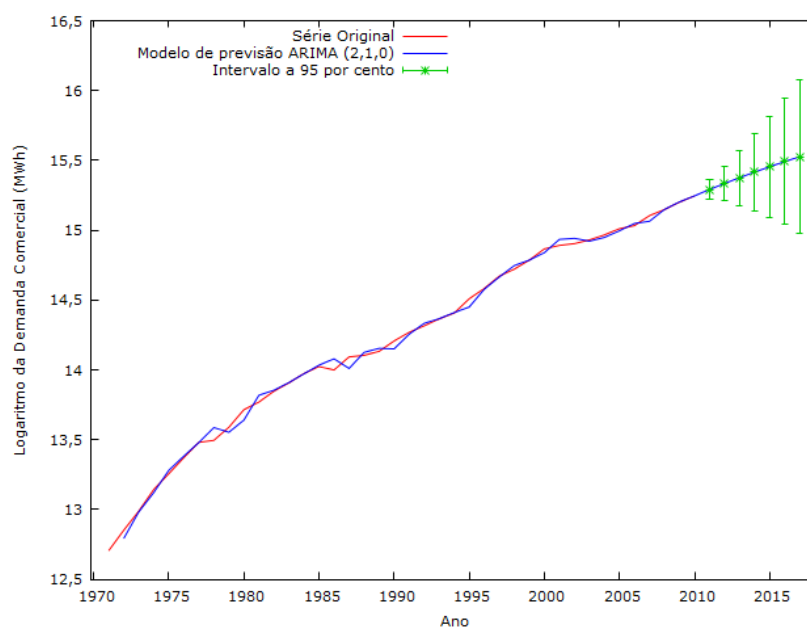


Figura 4.12 – Gráfico do logaritmo da demanda comercial de energia elétrica/RS com modelo de previsão ARIMA (2,1,0)

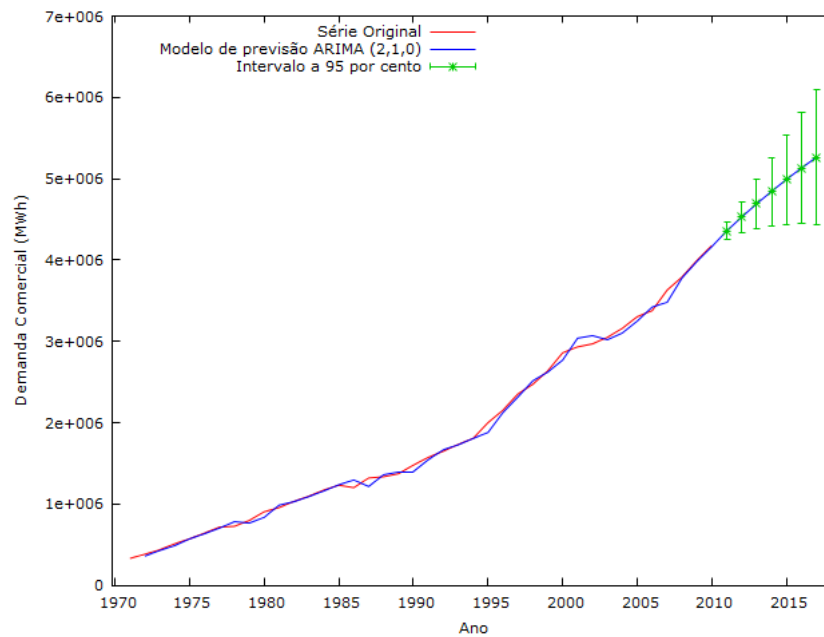
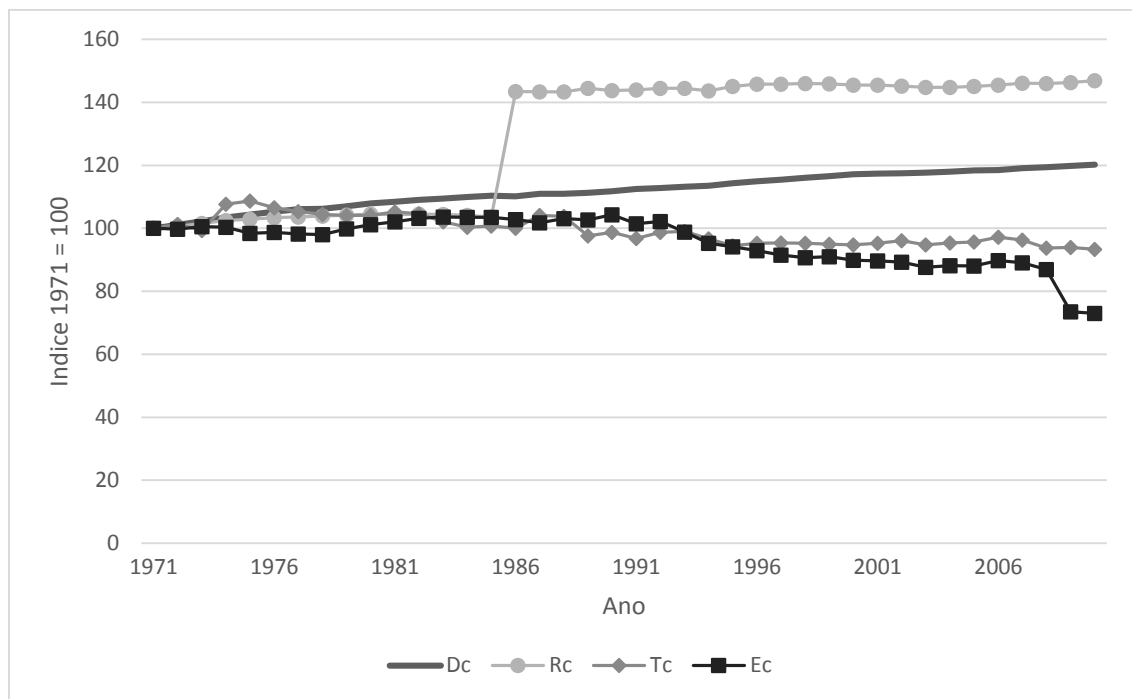


Figura 4.13 – Gráfico da demanda comercial de energia elétrica/RS com modelo de previsão ARIMA (2,1,0)

4.2.2. Metodologia VAR-VEC

Gráfico em linhas

Seguindo o mesmo processo aplicado no setor residencial, a Fig. 4.14 apresenta o gráfico em linhas das variáveis independentes e dependente para o setor comercial.



Fonte: Dados de Eletrobras (2018).

Obs: Dc = demanda; Rc = renda; Tc = tarifa; Ec = preço dos eletrointensivos.

Figura 4.14 – Comportamento das séries temporais deflacionadas ligadas à demanda comercial em relação ao primeiro valor observado em escala para 100

Na Figura 4.14 é possível observar uma variação acentuada na renda entre os anos de 1985 e 1986, no início da mudança de planos econômicos, desta maneira esta variável manteve os seus valores neste novo patamar. Além disso, no período analisado, a demanda de energia elétrica aumentou, assim como a renda, e houve uma pequena queda nas variáveis tarifa de energia elétrica e preço dos eletrointensivos, sendo que este último apresentou uma queda mais acentuada nos anos finais.

Teste de normalidade

Os testes de raiz unitária de Dickey-Fuller Aumentado (ADF) e Kwiatkowski–Phillips–Schmidt–Shin (KPSS) foram realizados depois da análise gráfica em cada uma das variáveis em nível e em primeira diferença. Os resultados dos testes ADF são mostrados na Tab. 4.24, enquanto os resultados dos testes KPSS são mostrados na Tab. 4.25.

Pelo teste ADF, todas as variáveis podem ser consideradas não estacionárias em nível, a um nível de significância de 0,01, e podem ser consideradas estacionárias em primeira diferença a um nível de significância de 0,01. A partir dos resultados do teste KPSS, pode-se observar que, considerando um nível de significância de 0,01, os dados

estudados apresentaram indícios de não estacionários apenas em nível.

Tabela 4.28 – Teste ADF no setor comercial

Variável	Termos na equação	Nº de defasagens	Valor-p
Dc – Demanda	Constante e Tendência	0	0,4785
ΔDc	Constante e Tendência	0	< 0,01
Rc – Renda	Sem Constante	0	0,9074
ΔRc	Sem Constante	0	< 0,01
Tc – Tarifa	Constante e Tendência	0	0,0296
ΔTc	Constante	3	< 0,01
Ec – Eletrointensivos	Constante	0	0,9972
ΔEc	Constante	0	< 0,01

Tabela 4.29 – Teste KPSS no setor comercial

Variável	Termos da equação	Nº de defasagens	Estatística de teste	Valores críticos (5%)	Valores críticos (1%)
Dc	Constante e Tendência	3	0,20	0,149	0,211
ΔDc	Constante e Tendência	3	0,15	0,149	0,211
Rc	Constante e Tendência	3	0,18	0,149	0,211
ΔRc	Constante	3	0,12	0,473	0,719
Tc	Constante	3	0,93	0,473	0,719
ΔTc	Constante	3	0,14	0,473	0,719
Ec	Constante e Tendência	3	0,18	0,149	0,211
ΔEc	Constante	3	0,27	0,473	0,719

Individualmente as séries Dc, Rc, Tc e Ec podem ser consideradas estacionárias em primeira diferença, tanto pelo teste ADF como pelo teste KPSS, para um nível de significância de 0,01.

Teste de Cointegração

Para a construção do modelo VAR é necessário determinar o número de defasagens, a presença de variáveis dummy e de termos determinísticos (constante e tendência), o que é feito pelos critérios de informação de Schwartz, Hannan-Quinn e Akaike. As variáveis dummy analisadas foram as apresentadas no capítulo 3, identificadas por *DYc1*, *DYc2*, *DYc3*, *DYc4*, *DYc5*, *DYc6*. Os resultados desta análise estão apresentados na Tab. 4.26.

Tabela 4.30 – Seleção de defasagens do modelo VAR (continua)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Dc; Rc; Tc; Ec; DYc3; DYc6	Constante e	1	-9,7747*	-8,3957*	-9,2841*
	Tendência**	2	-9,4935	-7,4249	-8,7575
	Constante	1	-9,6086*	-8,4019*	-9,1793*
		2	-9,3615	-7,4653	-8,6869
	Sem Constante	1	-7,9126	-6,8783*	-7,5446*
		2	-8,0749*	-6,3511	-7,4616
Dc; Rc; Tc; Ec; DYc3	Constante e	1	-8,7699*	-7,5633*	-8,3406*
	Tendência**	2	-8,4367	-6,5406	-7,7621
	Constante	1	-8,5075*	-7,4732*	-8,1395*
		2	-8,1726	-6,4488	-7,5593
	Sem Constante	1	-6,8117	-5,9498*	-6,5051*
		2	-6,8788*	-5,3274	-6,3268
Dc; Rc; Tc; Ec; DYc6	Constante e	1	-4,6882*	-3,4815*	-4,2589*
	Tendência	2	-4,4126	-2,5164	-3,7380
	Constante**	1	-4,8262*	-3,7920*	-4,4582*
		2	-4,5568	-2,8330	-3,9435
	Sem Constante	1	-4,3243*	-3,4624*	-4,0176*
		2	-4,1547	-2,6033	-3,6027

Tabela 4.31 – Seleção de defasagens do modelo VAR (conclusão)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Dc; Rc; Tc; Ec; DYc4	Constante e	1	-4,2126*	-3,0059*	-3,7833*
	Tendência	2	-3,9854	-2,0893	-3,3108
	Constante**	1	-4,1876*	-3,1533*	-3,8196*
		2	-3,8526	-2,1289	-3,2393
	Sem Constante	1	-3,7960*	-2,9341*	-3,4894*
		2	-3,4827	-1,9313	-2,9308
	Constante e	1	-3,7002*	-2,6660*	-3,3323*
	Tendência	2	-3,3665	-1,6427	-2,7532
Dc; Rc; Tc; Ec	Constante**	1	-3,7247*	-2,8629*	-3,4181*
		2	-3,3686	-1,8172	-2,8167
	Sem Constante	1	-3,2211*	-2,5316*	-2,9758*
		2	-2,9597	-1,5807	-2,4691
	Constante e	1	-3,5720*	-2,3654*	-3,1427*
	Tendência	2	-3,2573	-1,3611	-2,5826
Dc; Rc; Tc; Ec; DYc2	Constante**	1	-3,6352*	-2,6009*	-3,2672*
		2	-3,2985	-1,5747	-2,6852
	Sem Constante	1	-3,1030*	-2,2412*	-2,7964*
		2	-2,8427	-1,2913	-2,2907
	Constante e	1	-3,6382*	-2,4316*	-3,2089*
	Tendência	2	-3,3432	-1,4470	-2,6686
Dc; Rc; Tc; Ec; DYc5	Constante**	1	-3,6337*	-2,5994*	-3,2657*
		2	-3,2766	-1,5528	-2,6633
	Sem Constante	1	-3,2337*	-2,3718*	-2,9271*
		2	-2,9651	-1,4137	-2,4131
	Constante e	1	-3,5855*	-2,3789*	-3,1562*
	Tendência	2	-3,3191	-1,4229	-2,6444
Dc; Rc; Tc; Ec; DYc1	Constante**	1	-3,5856*	-2,5513*	-3,2176*
		2	-3,2332	-1,5094	-2,6199
	Sem Constante	1	-3,1507*	-2,2888*	-2,8441*
		2	-2,9151	-1,3637	-2,3631
	Constante e	1	-3,5855*	-2,3789*	-3,1562*

Obs: 1 - * indica o menor valor obtido pelo critério; 2 - ** indica o termo da equação que apresentou os menores valores dos critério de informação.

De acordo com os resultados dos critérios de informação, apresentados na Tabela 4.26, o modelo mais parcimonioso utiliza duas variáveis dummy, DYc3 e DYc6, apresenta constante e tendência determinística, além de utilizar uma defasagem.

Em sequência, para determinar o posto da matriz de cointegração pelos testes de Johansen, foram consideradas as situações: sem constante, com constante apenas no vetor de cointegração (restrita) ou fora do vetor de cointegração (irrestrita) e com tendência restrita ou irrestrita. Os resultados obtidos para o teste do traço e o teste do máximo autovalor para o modelo escolhido que utilizou tendência sem restrições podem ser visualizados na Tab. 4.27 e Tab. 4.28.

Tabela 4.32 – Teste do traço de Johansen para o modelo com Dc, Rc, Tc, Ec, DYc3 e DYc6 com tendência sem restrição

Teste	Hipótese nula (H_0)	Hipótese alternativa (H_1)	Valor-p
1º	$r = 0$	$r > 0$	0,0000
2º	$r \leq 1$	$r > 1$	0,1350
3º	$r \leq 2$	$r > 2$	0,7750
4º	$r \leq 3$	$r > 3$	0,3085

De acordo com a Tabela 4.27, o primeiro teste do traço indica a rejeição da hipótese nula de que não existe relação de longo prazo entre as variáveis, encontrando evidências de que exista pelo menos um vetor de cointegração entre elas. Por outro lado, no segundo teste do traço, a hipótese de que $r \leq 1$ não pode ser rejeitada para um nível de significância de 0,01, indicando a existência de no máximo um vetor de cointegração entre as variáveis.

Tabela 4.33 – Teste do máximo autovalor de Johansen para o modelo Dc, Rc, Tc, Ec, DYc3 e DYc6 com tendência sem restrição

Teste	Hipótese nula (H0)	Hipótese alternativa (H1)	Valor-p
1º	$r = 0$	$r = 1$	0,0000
2º	$r = 1$	$r = 2$	0,0465
3º	$r = 2$	$r = 3$	0,8144
4º	$r = 3$	$r = 4$	0,2908

O primeiro teste do máximo autovalor, cujos os resultados estão apresentados na Tab. 4.28, indica rejeição da hipótese de que não existe vetor de cointegração entre as variáveis, corroborando com o resultado do teste do traço. Ao testar a hipótese nula de que $r = 1$, o teste indica que a mesma não pode ser rejeitada a um nível de significância de 0,01. Assim sendo, como no teste do traço, segundo o teste do máximo autovalor, existem evidências que há apenas um vetor de cointegração entre as séries das variáveis relacionadas a demanda comercial de energia elétrica a um nível de significância de 0,01.

Deste modo os testes de cointegração indicam que as séries são cointegráveis, sugerindo a existência de relação de longo prazo entre elas. Uma vez que o objetivo deste trabalho é determinar a influência das variáveis estudadas na demanda de energia elétrica, optou-se por trabalhar com o vetor de cointegração normalizado em relação a variável Dc. O vetor normalizado, resultante do procedimento de Johansen é apresentado na Tab. 4.29.

Tabela 4.34 – Vetor de cointegração normalizado para o setor comercial

Dc	Rc	Tc	Ec
1	-0,88702	0,10373	-0,18031
Desvio Padrão	(0,01048)	(0,07963)	(0,04766)

A partir da Tabela 4.29 é possível escrever a equação que expressa a relação de longo prazo entre as variáveis, dada pela Eq. 4.3.

$$Dc = 0,88702Rc - 0,10373Tc + 0,18031Ec \quad (4.3)$$

Conforme a Equação 4.3, os coeficientes dos parâmetros renda e tarifa estão de acordo com o esperado pela teoria econômica, ou seja, a variável renda apresenta sinal positivo, e a variável tarifa apresenta sinal negativo. Além do mais, a demanda de energia elétrica comercial sofre mais alterações com a variação da renda, seguido pela alteração do preço dos materiais elétricos do que com a alteração da tarifa. Talvez, com o aumento da renda, o comerciante tenda a investir em seu negócio, com compra de novos equipamentos, aumentando a demanda de energia elétrica. Já com o aumento da tarifa, pode ser que este consumidor tenda a reduzir os seus gastos, desligando equipamentos e, conseqüentemente, diminuindo a demanda. Destacou-se o sinal diferente do esperado na variável E_c , representando uma relação direta entre demanda e preço de eletrointensivos. Isto talvez possa ser explicado pelo fato de que quando o mercado está aquecido, há aumento da demanda de energia elétrica pelo aumento da atividade comercial, existindo uma maior procura por eletrointensivos, o que pode aumentar seu preço.

A análise da qualidade do modelo, apresentada na Tabela 4.30, foi realizada pelos testes de Jarque-Bera, Ljung-Box e ARCH visando verificar a normalidade dos resíduos, a autocorrelação e a heterocedasticidade, respectivamente. Os resultados desta análise detectam evidências de que os resíduos do modelo são normalmente distribuídos, não apresentam autocorrelação e não apresentam heterocedasticidade.

Tabela 4.35 – Testes de qualidade do modelo VAR para o setor comercial

Teste	Estatística de Teste	Valor-p
Jarque-Bera	1,8018	0,4062
Ljung-Box Q'	10,6170	0,9360
ARCH	116,9680	0,1207

Até o momento foi encontrada a relação de longo prazo entre as variáveis estudadas. Na sequência será modelada a relação de curto prazo.

Modelo de correção de erros

No modelo vetorial de correção de erros (VEC), o vetor de cointegração de longo prazo é utilizado como variável explicativa (u_t) do termo de correção de erros. O resultado do VAR-VEC é expresso pela Eq. 4.2.

$$\Delta Dc_t = 0,1072DYc3 + 0,0215DYc6 - 0,0169(u_{t-1}) \quad (4.4)$$

O modelo apresentou coeficiente de determinação $R^2 = 0,4897$ e $R^2_{ajustado} = 0,4297$, sugerindo que alterações nas variáveis independentes explicam 48,97% das variações da demanda de energia elétrica comercial. A constante e a tendência não apresentaram os coeficientes significativos, sendo retiradas da Eq. 4.4.

O coeficiente do termo erro na Eq. 4.4 indica que a velocidade de ajustamento em relação ao equilíbrio de longo prazo é $0,0169$, ou seja, 1,69% da diferença entre a relações de longo e curto prazo é corrigida a cada ano.

Os resultados da análise das suposições teóricas do modelo, mostrados na Tab. 4.31, indicam evidências de resíduos normalmente distribuídos, sem autocorrelação e nem heterocedasticidade, o que é desejável.

Tabela 4.36 – Testes das suposições teóricas modelo VAR-VEC para o setor comercial

Teste	Estatística de Teste	Valor-p
Jarque-Bera	2,5182	0,2839
Ljung-Box Q'	11,2550	0,9150
ARCH	116,2820	0,1270

Previsões

A validação do modelo foi realizada por meio da previsão da demanda de energia elétrica para os anos de 2011 à 2017, utilizando o modelo VAR e o VAR-VEC. A validação destes modelos foi realizada por meio das medidas RMSE, MAE, MAPE e o U de Theil, calculadas com a métrica original e com a transformação logarítmica, e estão apresentadas na Tab. 4.32.

Tabela 4.37 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão

Métrica	Modelo	RMSE	MAE	MAPE (%)	U de Theil
Logarítmica	VAR-VEC	0,0856	0,0746	0,4830	0,0055
	VAR	0,0961	0,0714	0,4625	0,0062
Original	VAR-VEC	422.328,150	365.169,844	7,1106	0,0836
	VAR	524.245,256	382.753,403	7,5725	0,1037

Para complementar a análise, foram construídos intervalos de confiança para os valores previstos, com nível de confiança de 95%. As Tab. 4.33 e 4.34 apresentam os intervalos de confiança e os valores observados, pela métrica original e logarítmica, para os modelos VAR e VAR-VEC, respectivamente. No modelo VAR-VEC, os valores observados nos anos de 2012 e 2014 não pertencem aos respectivos intervalos de confiança, estando acima do limite superior do intervalo. Já para o modelo VAR, os valores observados para os anos de 2016 e 2017 não pertencem aos respectivos intervalos de confiança, estando abaixo do valor mínimo.

Tabela 4.38 – Intervalos de confiança para o modelo VAR de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC-95%
logarítmica	2011	15,3422	15,3237	15,2788; 15,3686
	2012	15,4193	15,3749	15,3139; 15,4360
	2013	15,4330	15,4253	15,3531; 15,4974
	2014	15,5124	15,4749	15,3947; 15,5552
	2015	15,4701	15,5242	15,4379; 15,6104
	2016	15,4258	15,5731	15,4824; 15,6637
	2017	15,4314	15,6217	15,5278; 15,7156
original	2011	4.603.000,0	4.518.709,5	4.320.302,6; 4.726.223,4
	2012	4.972.000,0	4.756.149,6	4474369,5; 5.055.675,3
	2013	5.040.224,0	5.001.542,9	4653224,6; 5.375.929,2
	2014	5.456.817,0	5.256.321,5	4850857,3; 5.695.682,6
	2015	5.230.826,7	5.521.604,2	5065386,5; 6.018.911,6
	2016	5.004.438,0	5.798.338,7	5295870,9; 6.348.486,5
	2017	5.032.131,0	6.087.408,4	5541772,8; 6.686.773,3

Tabela 4.39 – Intervalos de confiança para o modelo VAR-VEC de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC-95%
Logarítmica	2011	15,3422	15,2976	15,2424; 15,3528
	2012	15,4193	15,3215	15,2433; 15,3997
	2013	15,4330	15,3433	15,2475; 15,4391
	2014	15,5124	15,3630	15,2524; 15,4737
	2015	15,4701	15,3806	15,2569; 15,5043
	2016	15,4258	15,3961	15,2606; 15,5317
	2017	15,4314	15,4095	15,2631; 15,5559
Original	2011	4.603.000,0	4.402.103,3	4.165.725,5; 4.651.894,0
	2012	4.972.000,0	4.508.666,5	4.169.651,4; 4.875.240,5
	2013	5.040.224,0	4.608.066,9	4.187.167,3; 5.071.275,9
	2014	5.456.817,0	4.699.769,4	4.207.566,5; 5.249.550,4
	2015	5.230.826,7	4.783.231,9	4.226.653,1; 5.413.097,3
	2016	5.004.438,0	4.857.944,7	4.242.231,6; 5.563.021,8
	2017	5.032.131,0	4.923.464,5	4.252.969,5; 5.699.670,9

Talvez a falha nas previsões obtidas com o modelo VAR-VEC para o setor comercial possa ter ocorrido devido ao baixo poder de ajuste de seu termo de correção de erros que foi de somente 1,69% por ano. Já a falha na previsão da demanda de energia pelo o setor comercial pelo modelo VAR pode ter ocorrido devido a oscilação da demanda energética nos últimos anos do período analisado, havendo queda nos anos de 2015 e 2016, com uma aparente estabilização no ano de 2017.

Os gráficos que representam os modelos de previsão VAR e VAR-VEC podem ser visualizados nas Fig. 4.15 e 4.16 respectivamente.

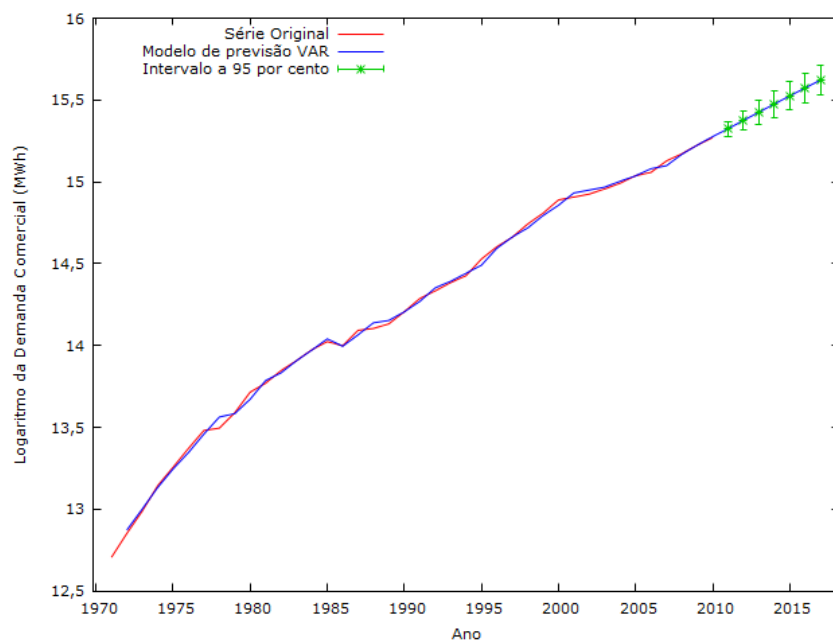


Figura 4.15 – Gráfico do modelo de previsão VAR para demanda comercial de energia elétrica/RS

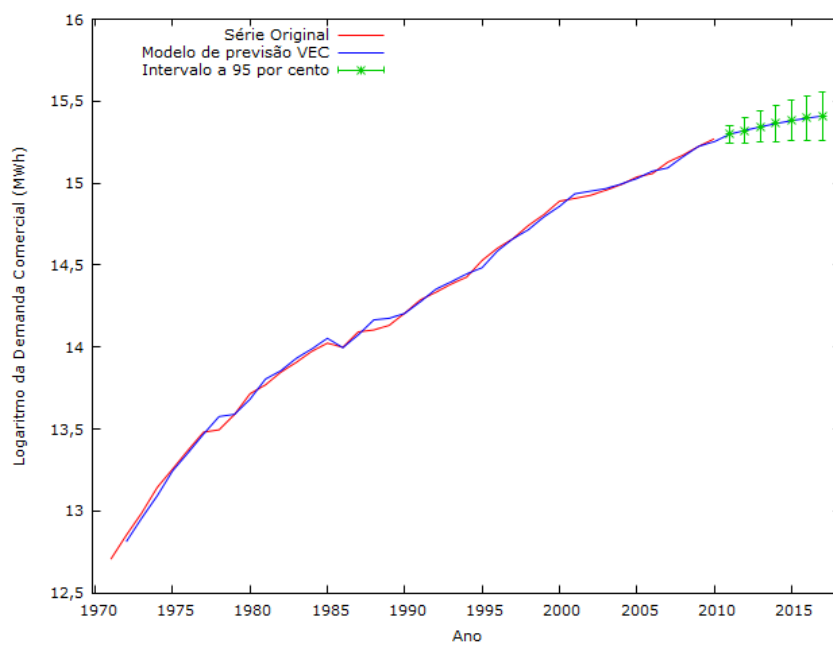


Figura 4.16 – Gráfico do modelo de previsão VEC para demanda comercial de energia elétrica/RS

4.2.3. Comparações para o setor Comercial

Técnica utilizada: ARIMA x VAR x VAR-VEC

Na comparação entre os indicadores de qualidade dos modelos, o ARIMA (2,1,0) apresentou os melhores resultados para a previsão, tanto na métrica logarítmica quanto na métrica original. Outro fato observado é sobre a estimativa pontual das previsões, em que modelo ARIMA (2,1,0) forneceu previsões acima do observado nos dois últimos anos, enquanto no modelo VAR isto ocorreu nos últimos três anos de previsão. Já o modelo VAR-VEC apresentou valores previstos abaixo dos observados em todo o período.

Nos intervalos de confiança, o modelo ARIMA (2,1,0) apresentou maiores amplitudes e conseguiu abranger todos os valores observados. O modelo VAR forneceu os intervalos de menor amplitude, mas não abrangeu os valores observados dos anos de 2016 e 2017, em que o valor observado ficou abaixo do valor mínimo. O modelo VAR-VEC apresentou os intervalos de confiança com amplitude intermediária, não abrangendo os valores observados dos anos de 2012 e 2014.

Finalizando, o melhor resultado para o MAPE foi o encontrado para o ARIMA (2,1,0), sendo de 0,43% para a escala logarítmica e 6,59% para a métrica original. Se considerar a escala logarítmica, o segundo melhor modelo foi o VAR (0,46%) seguido pelo modelo VAR-VEC (0,48%), enquanto que se considerar a escala original, o VAR-VEC encontra resultado melhor (7,11%) em relação ao VAR (7,57%). Se analisarmos no curto prazo, apenas um ano a frente, o modelo ARIMA apresenta o pior resultado com MAPE de 0,0509% na escala logarítmica, o modelo VAR foi o melhor (0,0185%), seguido pelo modelo VAR-VEC (0,0446%), o mesmo acontecendo para a escala original.

Comparação com o modelo de Silveira(2017)

A Tabela 4.35 apresenta os coeficientes para os modelos VAR encontrados por Silveira (2017) (Modelo BR) e neste estudo (Modelo RS), sendo possível notar que o Modelo RS apresentou elasticidade maior para a renda e menores para a tarifa e preço dos eletrointensivos, embora o sentido da relação destas variáveis com a demanda tenha sido o mesmo.

Tabela 4.40 – Coeficientes do modelo VAR Comercial/RS e BR

	Dc	Rc	Tc	Ec	Constante
Modelo RS	1	0,88702	-0,10373	0,18031	-
Modelo BR	1	0,4366	-0,6316	-0,4594	8,616

Já a Tabela 4.36 mostra os coeficientes dos modelos VAR-VEC dos estudos considerados nesta análise, observando-se que o Modelo RS comercial considerou duas variáveis dummy, enquanto o Modelo BR comercial apenas uma de importância intermediária, se comparada com as variáveis dummy do Modelo RS.

Tabela 4.41 – Coeficientes do modelo VAR-VEC Comercial/RS e BR

	ΔDc	DYc3	DYc6	Dc2	u_{i-1}
Modelo RS	1	0,1072	0,0215	-	0,0169
Modelo BR	1	-	-	-0,077	-0,1157

Na validação, o modelo BR obtido pelo VAR-VEC apresentou um MAPE 4,78% considerando apenas um ano após o período estudado, enquanto que no Modelo RS, analisando apenas o ano subsequente ao período utilizado para a modelagem, esta medida assume o valor 0,29%, mostrando novamente que o Modelo RS foi melhor que o Modelo BR.

4.3. Industrial

4.3.1. Metodologia Box-Jenkins

O gráfico dos dados em nível para a demanda industrial de energia elétrica no Rio Grande do Sul pode ser visto na Fig. 4.17, em que é possível identificar uma tendência exponencial com uma oscilação entre os anos de 1987 e 1991 e outra entre os anos 2000 e 2002. Também se identifica no ano de 2009 uma queda acentuada nesta demanda devido à crise econômica mundial, sentida de maneira mais acentuada pelo setor industrial. O uso da transformação logarítmica para estabilizar a variância amenizou estas variações.

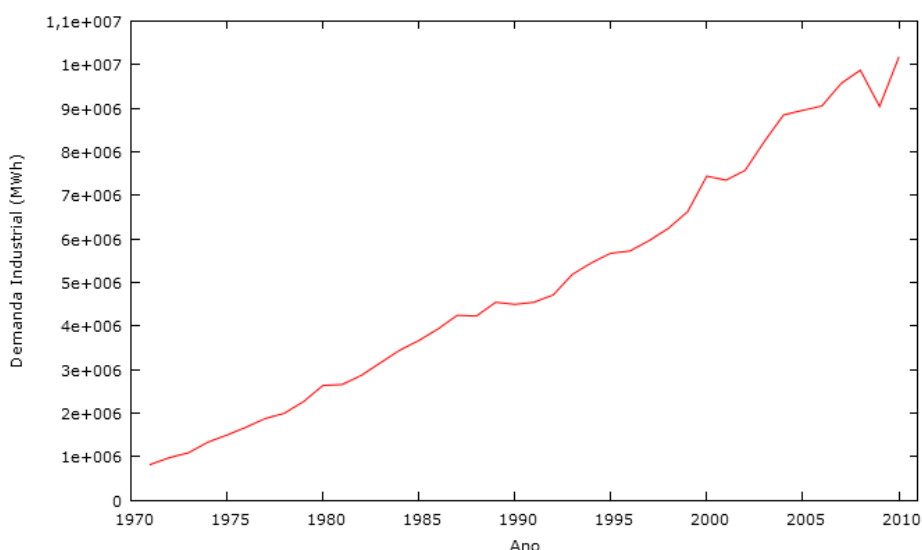


Figura 4.17 – Demanda industrial de energia elétrica no Rio Grande do Sul no período 1961 - 2010

Na Figura 4.18.a é possível visualizar o gráfico do logaritmo da demanda industrial, enquanto na Fig. 4.18.b, os gráficos FAC (superior) e FACP (inferior) destas observações, tomando limites com confiança de 95%. A lenta queda exponencial na FAC com muitas correlações significativas caracteriza a não estacionaridade.

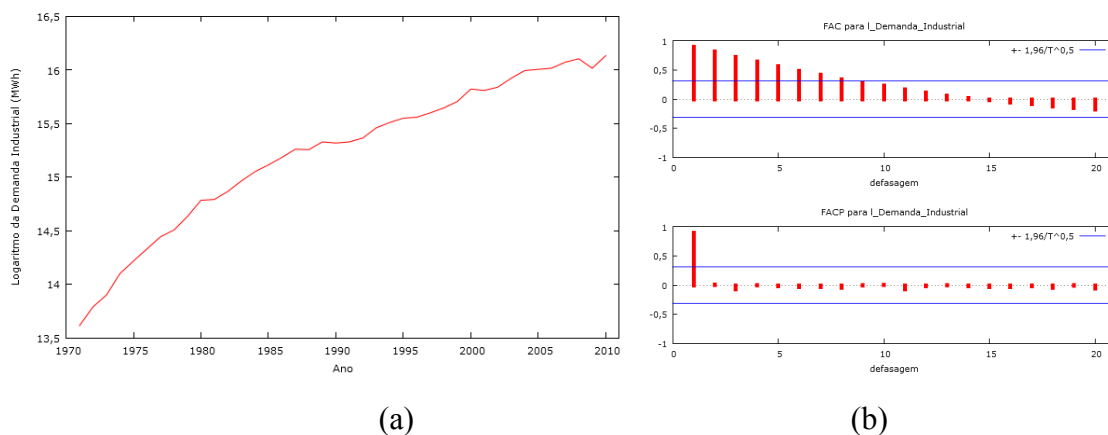


Figura 4.18 – Gráfico do logaritmo da demanda industrial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)

No teste ADF, aplicado para um número máximo de dez defasagens, não foram encontradas evidências para rejeitar a hipótese nula de que a série não é estacionária (*valor $p > 0,01$*) para modelo com tendência e constante. Já no teste KPSS, foram encontradas evidências para rejeitar a hipótese nula de que a série é estacionária (*valor $p < 0,01$*), o que fortalece a ideia de que a série deve ser considerada como não estacionária.

Parte-se para a análise do comportamento da primeira diferença. O gráfico em linhas (a) e os gráficos FAC e FACP (b) da primeira diferença do logaritmo das observações são mostrados na Fig. 4.19, tomando limites de confiança de 95%. No gráfico (a) desta figura, é possível identificar um comportamento aparentemente estacionário. No gráfico (b) superior da Fig. 4.19, observa-se que a FAC apresenta uma queda lenta sem a existência de correlações significativas, enquanto que no gráfico (b) inferior desta mesma figura, a FACP não possui correlações significativas. Estas propriedades sugerem um modelo ARMA(0,0). Por se modelar a primeira diferença, sugere-se um ARIMA (0, 1, 0).

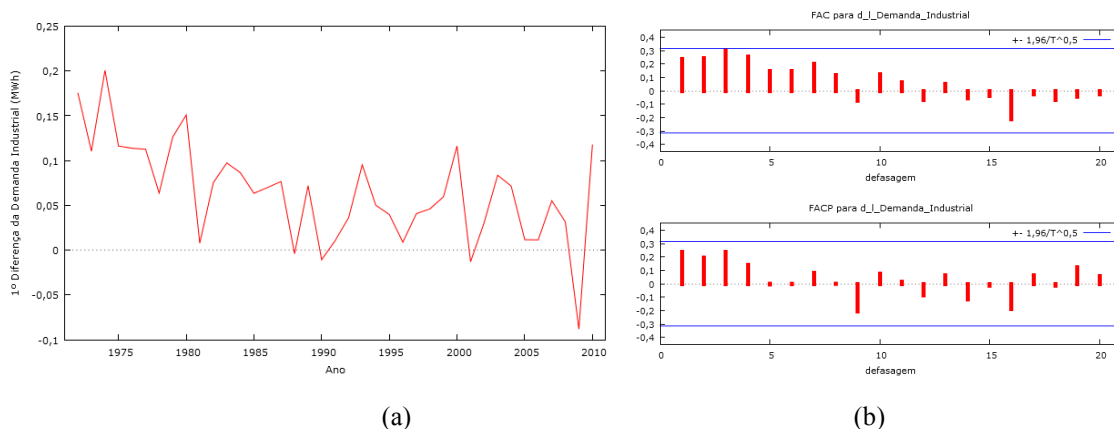


Figura 4.19 – Gráfico da primeira diferença do logaritmo da demanda industrial de energia elétrica no Rio Grande do Sul ao longo do tempo (a) e respectivos correlogramas da FAC e FACP (b)

No teste ADF, aplicado para um número máximo de dez defasagens, foram encontradas evidências para rejeitar a hipótese nula de que a série não é estacionária (*valor $p < 0,01$*) para modelo com tendência e constante, enquanto no teste KPSS não foram encontradas evidências para rejeitar a hipótese nula de que a série é estacionária (*valor $p > 0,10$*). Logo, ela pode ser considerada estacionária.

Assim, o modelo ARIMA (0,1,0) foi avaliado, admitindo-se variações de até +2 nas ordens dos componentes auto regressivos e de médias móveis. Neste caso, os modelos considerados como candidatos foram: ARIMA (0,1,0), ARIMA (0,1,1), ARIMA (0,1,2), ARIMA (1,1,0), ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (2,1,0), ARIMA (2,1,1) e ARIMA (2,1,2).

Os modelos candidatos foram estimados com e sem constante e a significância de seus coeficientes está apresentada na Tab. 4.37, em que é possível constatar que, sem uso de constante, os modelos ARIMA (0,1,0), ARIMA (1,1,0), ARIMA (0,1,1) e ARIMA (1,1,1) apresentaram todos os seus coeficientes significativos para um nível de significância de 0,01, não havendo interferência da constante. Já se houver a inserção de constante, apenas o modelo ARIMA (0,1,0) teve os seus coeficientes significativos a um nível de significância de 0,01. Para um nível de significância de 0,05, sem o acréscimo de uma constante, três modelos juntam-se aos já apresentados: ARIMA (2,1,0), ARIMA (0,1,2) e ARIMA (1,1,2). Com o acréscimo de constante, também houve o acréscimo do modelo ARIMA (1,1,1) ao conjunto dos modelos que apresentam coeficientes significativos.

Tabela 4.42 – Significância dos coeficientes dos modelos selecionados

ARIMA (p,1,q)	Coeficiente	valor p	
		Sem constante	Com constante
(0,1,0)	C	-	< 0,01*
(1,1,0)	C	-	< 0,01*
	ϕ_1	< 0,01*	0,1075
(2,1,0)	C	-	< 0,01*
	ϕ_1	0,0327**	0,2814
	ϕ_2	< 0,01*	0,1482
(0,1,1)	-	-	< 0,01*
	θ_1	< 0,01*	0,1881
(1,1,1)	C	-	0,0179**
	ϕ_1	< 0,01*	< 0,01*
	θ_1	< 0,01*	< 0,01*
(2,1,1)	C	-	0,0350**
	ϕ_1	< 0,01*	< 0,01*
	ϕ_2	0,1252	0,2139
	θ_1	< 0,01*	< 0,01*
(0,1,2)	C	-	< 0,01*
	θ_1	0,0162**	0,4493
	θ_2	0,0209**	0,3937
(1,1,2)	C	-	0,0628
	ϕ_1	< 0,01*	< 0,01*
	θ_1	< 0,01*	< 0,01*
	θ_2	0,0367**	0,0557
(2,1,2)	C	-	0,0655
	ϕ_1	< 0,01*	< 0,01*
	ϕ_2	0,0795	0,0636
	θ_1	< 0,01*	< 0,01*
	θ_2	< 0,01*	< 0,01*

Obs: * indica significativo para nível de significância 0,01. ** indica significativo para nível de significância 0,05.

Os resultados obtidos para avaliação da parcimonia dos modelos candidatos pelos critérios de informação estão apresentados na Tab. 4.38. O melhor modelo sem o acréscimo de constante foi o modelo ARIMA (1,1,2) (CIS = -109,786, CIA = -116,440, HQ = -114,053), seguido pelos modelos ARIMA (1,1,1), ARIMA (2,1,2), ARIMA (2,1,1), ARIMA (2,1,0), ARIMA (1,1,0), ARIMA (0,1,2), ARIMA (0,1,1) e ARIMA (0,1,0). Com o acréscimo de uma constante, a sequência não foi alterada, sendo seus coeficientes sempre superiores ao correspondente sem constante.

Tabela 4.43 – Resultados dos critérios de informação nos modelos candidatos para modelos ARIMA (p,1,q)

Modelo	Critério de informação			
	Constante	Schwarz	Akaike	Hannan-Quinn
(1,1,2)	Sem	-109,786	-116,440	-114,053
	Com	-107,631	-115,949	-112,965
(1,1,1)	Sem	-110,104	-115,095	-113,304
	Com	-108,419	-115,073	-112,686
(2,1,2)	Sem	-107,729	-116,047	-113,063
	Com	-105,571	-115,552	-111,971
(2,1,1)	Sem	-108,098	-114,752	-112,364
	Com	-106,170	-114,488	-111,503
(2,1,0)	Sem	-102,179	-107,170	-105,379
	Com	-104,859	-111,513	-109,126
(1,1,0)	Sem	-97,6034	-100,931	-99,737
	Com	-106,510	-111,501	-109,710
(0,1,2)	Sem	-87,132	-92,122	-90,332
	Com	-102,910	-109,565	-107,177
(0,1,1)	Sem	-86,547	-89,874	-88,680
	Com	-105,856	-110,847	-109,056
(0,1,0)	Sem	-77,833	-79,496	-78,899
	Com	-107,678	-111,005	-109,811

Em sequência, foi realizada a averiguação da presença de autocorrelação, normalidade e homoscedasticidade entre variâncias nos resíduos. Os resultados, apresentados na Tab. 4.38, mostram que todos os modelos cumprem as suposições do modelo teórico.

Tabela 4.38 – Resultados dos testes de Ljung-Box, Jarque-Bera e ARCH para os resíduos dos modelos candidatos ARIMA (p,1,q) (continua)

Modelo	Ljung Box			Jarque-Bera		ARCH	
	Constante	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor
(1,1,2)	Sem	10,7679	0,8684	0,1008	0,9508	12,6476	0,3952
	Com	11,9940	0,8005	0,0558	0,9725	12,5575	0,4020
(1,1,1)	Sem	14,2890	0,7101	0,1635	0,9215	9,1922	0,6864
	Com	13,3385	0,7711	0,3547	0,8375	8,0778	0,7790
(2,1,2)	Sem	9,4697	0,8928	0,2017	0,9041	14,0808	0,2956
	Com	10,1314	0,8597	0,3112	0,8559	12,9575	0,3721

Tabela 4.38 – Resultados dos testes de Ljung-Box, Jarque-Bera e ARCH para os resíduos dos modelos candidatos ARIMA (p,1,q) (conclusão)

Modelo	Ljung Box			Jarque-Bera		ARCH	
	Constante	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor
(2,1,1)	Sem	12,2501	0,7848	0,0431	0,9787	10,1444	0,6033
	Com	12,5293	0,7671	0,1577	0,9242	8,5907	0,7374
(2,1,0)	Sem	20,2945	0,3165	0,1310	0,9366	8,7663	0,7227
	Com	12,6770	0,8104	0,5256	0,7689	9,4198	0,6667
(1,1,0)	Sem	21,9559	0,2865	0,4908	0,7824	14,2201	0,2869
	Com	13,4908	0,8125	0,4490	0,7989	11,0446	0,5251
(0,1,2)	Sem	17,4407	0,4930	0,1626	0,9219	12,6555	0,3946
	Com	13,6951	0,7487	0,5039	0,7773	9,5077	0,6991
(0,1,1)	Sem	23,3916	0,2205	0,2614	0,8775	14,6852	0,2591
	Com	15,2837	0,7044	0,4680	0,7913	10,7384	0,5515
(0,1,0)	Sem	22,5682	0,3105	0,4113	0,8141	11,0517	0,5245
	Com	22,5682	0,3105	0,4113	0,8141	7,1373	0,8484

Assim, o modelo de melhor desempenho pelos critérios de informação, significância dos coeficientes e cumprimento dos pressupostos do modelo teórico foram os modelos ARIMA (1,1,1) e ARIMA (1,1,2) sem constante.

A validação do modelo encontrado, foi realizada por meio das medidas: raiz do erro quadrático médio (RMSE), o erro absoluto médio (MAE), o erro absoluto percentual médio (MAPE) e o U de Theil, calculados com a métrica original e com a transformação logarítmica estão apresentadas na Tab. 4.39. O modelo de previsão selecionado foi o ARIMA (1,1,2), pois apresentou os menores erros.

Tabela 4.44 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão

Métrica	ARIMA (p,2,q)	RMSE	MAE	MAPE (%)	U de Theil
Logarítmica	(1,1,2)	0,0838	0,0644	0,3999	0,0052
	(1,1,1)	0,1660	0,1345	0,8352	0,0103
Original	(1,1,2)	867.030,109	667.835,694	6,7630	0,0851
	(1,1,1)	1.829.996,537	1.478.264,387	14,9443	0,1797

A Tab. 4.40 apresenta os valores previstos para o período 1911-1917 com respectivos intervalos de confiança e seus valores de referência, pela métrica original e logarítmica. Os valores observados em todo o período considerado pertencem aos respectivos intervalos de confiança.

Tabela 4.45 – Intervalos de confiança para o modelo ARIMA (1,1,2) de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC- 95%
logarítmica	2011	16,1759	16,1249	16,0203; 16,2389
	2012	16,1426	16,1429	16,0051; 16,3711
	2013	16,1825	16,1608	15,9227; 16,4884
	2014	16,1645	16,1786	15,8584; 16,6248
	2015	16,1067	16,1962	15,7709; 16,7548
	2016	16,0658	16,2136	15,6862; 16,8911
	2017	16,1045	16,2309	15,5925; 17,0243
original	2011	10.595.000,0	10.067.429,8	9.068.079,0; 11.284.075,0
	2012	10.248.000,0	10.251.247,9	8.931.884,9; 12.878.940,3
	2013	10.665.149,0	10.436.679,1	8.225.182,1; 14.481.796,4
	2014	10.474.837,0	10.623.700,9	7.712.886,0; 16.597.423,0
	2015	9.886.935,4	10.812.300,7	7.066.642,3; 18.903.090,8
	2016	9.490.690,0	11.002.466,1	6.492.920,6; 21.662.170,5
	2017	9.864.617,2	11.194.173,7	5.912.048,8; 24.749.380,1

As Figuras 4.20 e 4.21 ilustram o modelo de previsão selecionado e sua comparação com a série em nível, por meio da métrica logarítmica e original, respectivamente.

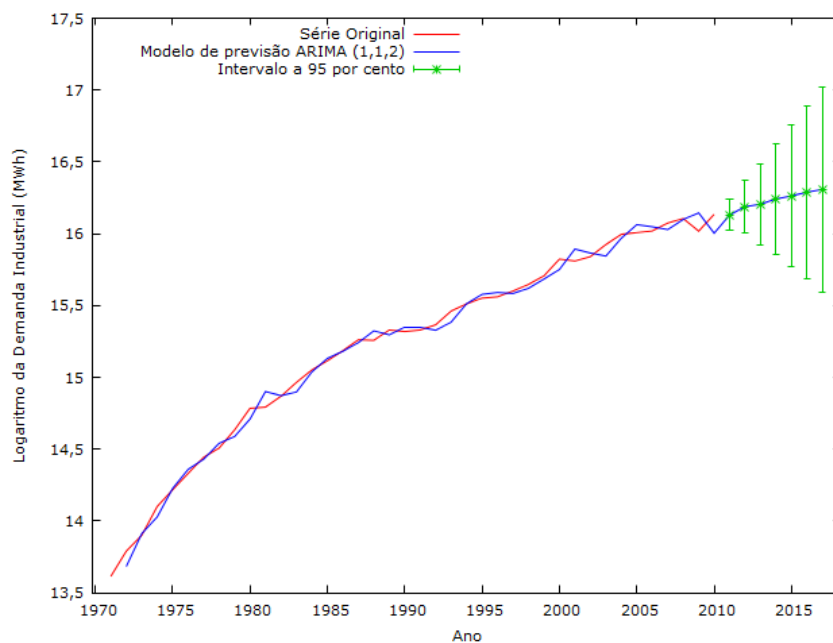


Figura 4.20 – Gráfico do logaritmo da demanda industrial de energia elétrica/RS com modelo de previsão ARIMA (1,1,2)

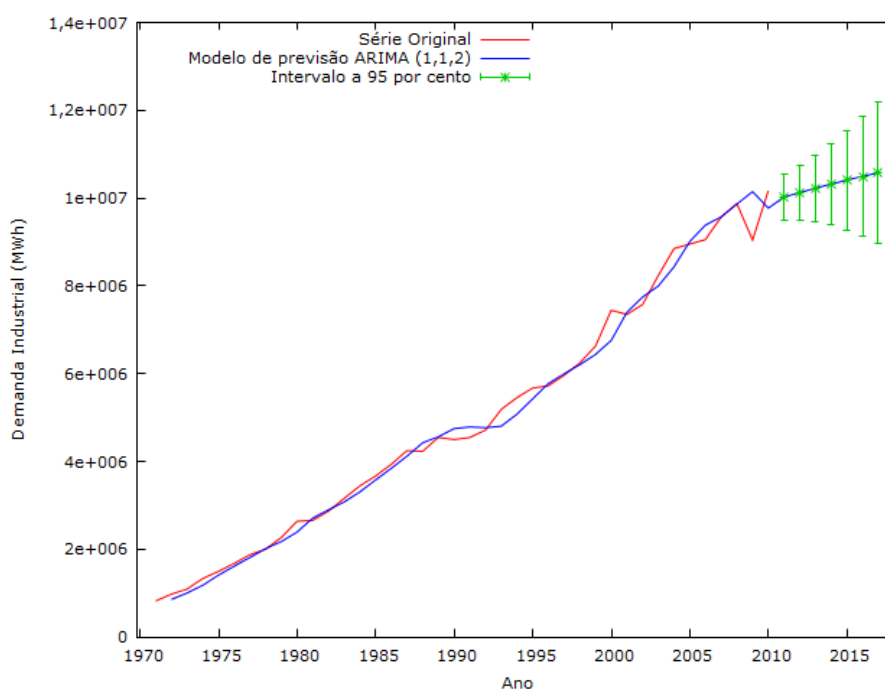
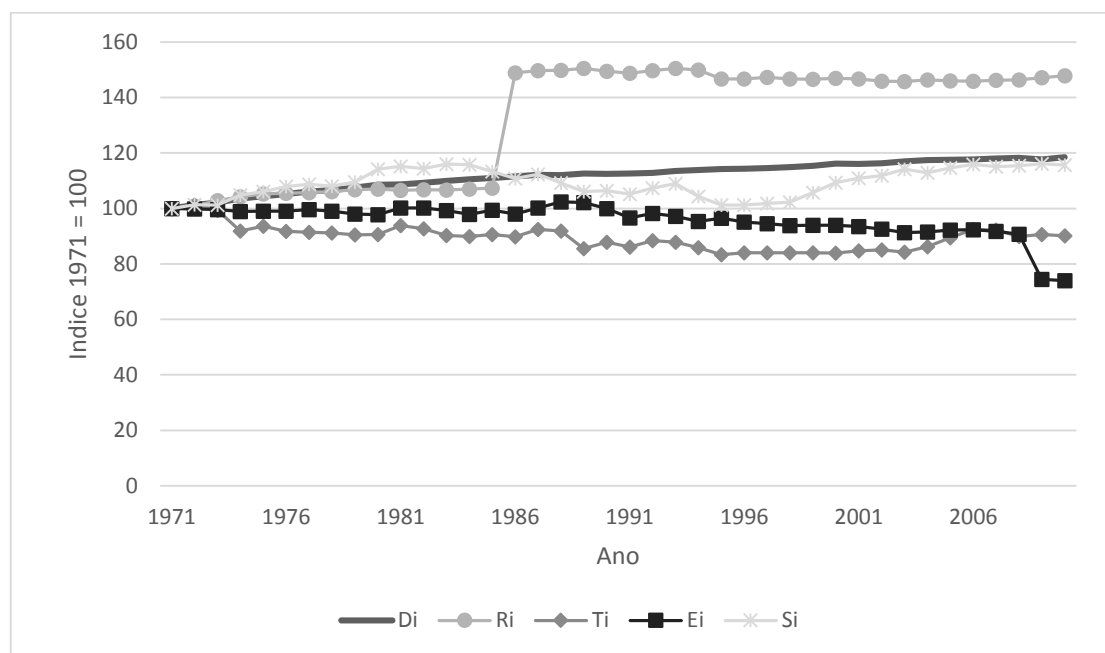


Figura 4.21 – Gráfico da demanda industrial de energia elétrica/RS com modelo de previsão ARIMA (1,1,2)

4.3.2. Metodologia VAR-VEC

Gráfico em linhas

Seguindo o mesmo processo aplicado nos setores residencial e comercial, a Fig. 4.22 apresenta o gráfico em linhas das variáveis independentes e dependente consideradas neste estudo para o setor industrial.



Fonte: Dados de Eletrobrás (2018).

Obs: Di = demanda; Ri = renda; Ti = tarifa; Ei = preço dos eletrointensivos; Si = preço substituto de energia elétrica.

Figura 4.22 – Comportamento das séries temporais deflacionadas ligadas à demanda industrial em relação ao primeiro valor observado em escala para 100

Na Figura 4.22 é possível observar um aumento bastante grande no PIB quando ocorreram as mudanças na moeda brasileira, no mesmo período que este fenômeno ocorreu no setor comercial. Também é possível observar aumento gradual na demanda de energia elétrica e no preço substituto de energia elétrica. As variáveis tarifa de energia elétrica e preço dos eletrointensivos apresentaram queda, sendo a deste último mais acentuada nos anos finais.

Teste de normalidade

Os resultados dos testes ADF são mostrados na Tab. 4.41, enquanto os dos testes KPSS na Tab. 4.42. Pelo teste ADF, todas as variáveis podem ser consideradas não

estacionárias em nível, utilizando um grau de significância de 0,01, e podem ser consideradas estacionárias em primeira diferença para este mesmo nível de significância. Com base nos resultados do teste KPSS, pode-se concluir que, ao grau de significância de 0,01, os dados estudados apresentaram indícios de não estacionária apenas em nível, com exceção da variável que substitui a energia elétrica (Si). Assim, as séries estudadas se tornaram estacionárias em primeira diferença, exceto a variável Si, que se tornou estacionária já em nível.

Tabela 4.46 – Teste ADF no setor industrial

Variável	Termos da equação	Nº de defasagens	Valor-p
Di – Demanda	Constante e Tendência	0	0,0613
Δ Di	Constante e Tendência	0	< 0,01
Ri – Renda	Sem constante	0	0,8993
Δ Ri	Sem constante	0	< 0,01
Ti – Tarifa	Constante	0	0,0980
Δ Ti	Sem constante	0	< 0,01
Ei – Máquinas e equipamentos	Sem constante	0	0,1423
Δ Ei	Sem constante	0	< 0,01
Si – Substituto de energia	Constante	3	0,2292
Δ Si	Constante	0	< 0,01

Tabela 4.47 – Teste KPSS no setor industrial (continua)

Variável	Termos da equação	Nº de defasagens	Estatística de teste	Valores críticos (5%)	Valores críticos (1%)
Di	Constante e Tendência	3	0,25	0,149	0,212
Δ Di	Constante e Tendência	3	0,19	0,149	0,212
Ri	Constante e Tendência	3	0,19	0,149	0,212
Δ Ri	Constante	3	0,13	0,473	0,719
Ti	Constante e Tendência	3	0,20	0,149	0,212

Tabela 4.48 – Teste KPSS no setor industrial (conclusão)

Variável	Termos da equação	Nº de defasagens	Estatística de teste	Valores críticos (5%)	Valores críticos (1%)
ΔTi	Constante	3	0,30	0,473	0,719
Ei	Constante e Tendência	3	0,21	0,149	0,212
ΔEi	Constante	3	0,35	0,473	0,719
Si	Constante e Tendência	3	0,13	0,149	0,212
ΔSi	Constante	3	0,13	0,473	0,719

Deste modo, o teste ADF encontrou evidências de que todas as variáveis estudadas são estacionárias em primeira diferença a um nível de significância de 0,01. O teste KPSS discordou do ADF, apenas para a variável Si , não encontrando evidências de não estacionaridade em nível.

Teste de Cointegração

Para a construção do modelo VAR é necessário determinar o número de defasagens, a presença de variáveis dummy e de termos determinísticos (constante e tendência), o que é feito pelos critérios de informação de Schwartz, Hannan-Quinn e Akaike. Já as variáveis dummy analisadas foram as apresentadas no capítulo 3, identificadas por $DYi1$, $DYi2$, $DYi3$, $DYi4$, $DYi5$, $DYi6$. Os resultados desta análise estão apresentados na Tab. 4.43.

Tabela 4.49 – Seleção de defasagens do modelo VAR (continua)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Di; Ri; Ti; Ei; Si; DYi3	Constante e Tendência**	1	-8,7236*	-6,9999*	-8,1103*
		2	-8,6505	-5,8494	-7,6539
	Constante	1	-7,7744*	-6,2661*	-7,2378*
		2	-7,6965	-5,1109	-6,7766
	Sem Constante	1	-6,6160	-5,3232*	-6,1560
		2	-7,3578*	-4,9876	-6,5145*

Tabela 4.50 – Seleção de defasagens do modelo VAR (continua)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Di; Ri; Ti; Ei; Si; DYi6	Constante e	1	-6,1631*	-4,4393*	-5,5498*
	Tendência**	2	-5,9566	-3,1554	-4,9599
	Constante	1	-6,0852*	-4,5769*	-5,5485*
		2	-5,7123	-3,1267	-4,7924
	Sem Constante	1	-5,7479*	-4,4551*	-5,2879*
		2	-5,6245	-3,2543	-4,7812
	Constante e	1	-5,2869	-3,5632*	-4,6736*
	Tendência**	2	-5,4800*	-2,6789	-4,4834
Di; Ri; Ti; Ei; Si; DYi4	Constante	1	-5,1690	-3,6607*	-4,6324*
		2	-5,4441*	-2,8584	-4,5241
	Sem Constante	1	-4,7253	-3,4324*	-4,2653
		2	-5,4202*	-3,0500	-4,5769*
	Constante e	1	-4,5795	-2,8558*	-3,9662*
	Tendência**	2	-4,6297*	-1,8286	-3,6331
Di; Ri; Ti; Ei; Si; DYi5	Constante	1	-4,3944	-2,8861*	-3,8577*
		2	-4,5178*	-1,9322	-3,5979
	Sem Constante	1	-3,9784	-2,6856*	-3,5184
		2	-4,5006*	-2,1304	-3,6573*
	Constante e	1	-4,5681*	-2,8443*	-3,9548*
	Tendência**	2	-4,2400	-1,4389	-3,2434
Di; Ri; Ti; Ei; Si; DYi2	Constante	1	-4,3597*	-2,8514*	-3,8231*
		2	-4,0599	-1,4742	-3,1399
	Sem Constante	1	-3,8768	-2,5840*	-3,4168*
		2	-3,9792*	-1,6090	-3,1359
	Constante e	1	-4,4845*	-2,9762*	-3,9479*
	Tendência**	2	-4,0282	-1,4425	-3,1082
Di; Ri; Ti; Ei; Si	Constante	1	-4,4060*	-3,1131*	-3,9460*
		2	-4,0199	-1,6497	-3,1766
	Sem Constante	1	-4,0213*	-2,9439*	-3,6379*
		2	-4,0155	-1,8608	-3,2489

Tabela 4.51 – Seleção de defasagens do modelo VAR (conclusão)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Di; Ri; Ti; Ei; Si; DYi1	Constante e	1	-4,4214	-2,6976*	-3,8081*
	Tendência	2	-4,4288*	-1,6276	-3,4322
	Constante**	1	-4,3464	-2,8381*	-3,8098*
		2	-4,4432*	-1,8575	-3,5232
	Sem Constante	1	-3,8686	-2,5757*	-3,4086
		2	-4,4466*	-2,0764	-3,6033*

Obs: 1 - * indica o menor valor obtido pelo critério; 2 - ** indica o termo da equação que apresentou os menores valores dos critério de informação.

De acordo com os resultados dos critérios de informação, apresentados na Tabela 4.43, o modelo mais parcimonioso utiliza uma variável dummy, DYi3, apresenta constante e tendência determinística, além de utilizar uma defasagem.

Para determinar o posto da matriz de cointegração pelos testes de Johansen, foram consideradas as situações: sem constante, com constante apenas no vetor de cointegração (restrita) ou fora do vetor de cointegração (irrestrita) e com tendência restrita ou irrestrita. Os resultados obtidos para o teste do traço e o teste do máximo autovalor para o modelo escolhido que utilizou tendência restrita podem ser visualizados na Tab. 4.44 e Tab. 4.45.

Tabela 4.52 – Teste do traço de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3 e tendência restrita

Teste	Hipótese nula (H_0)	Hipótese alternativa (H_1)	Valor-p
1º	$r = 0$	$r > 0$	0,0000
2º	$r \leq 1$	$r > 1$	0,3786
3º	$r \leq 2$	$r > 2$	0,6471
4º	$r \leq 3$	$r > 3$	0,5189
5º	$r \leq 4$	$r > 4$	0,5592

A partir da Tabela 4.44, foi possível notar que o primeiro teste do traço indica a rejeição da hipótese nula, encontrando evidências de que exista pelo menos um vetor de

cointegração entre elas. No segundo teste do traço, a hipótese de que $r \leq 1$ não pode ser rejeitada para um nível de significância de 0,01, indicando a existência de no máximo um vetor de cointegração entre as variáveis.

Tabela 4.53 – Teste do máximo autovalor de Johansen para o modelo com D_i , R_i , T_i , E_i , DY_i3 e tendência restrita

Teste	Hipótese nula (H_0)	Hipótese alternativa (H_1)	Valor-p
1º	$r = 0$	$r = 1$	0,0000
2º	$r = 1$	$r = 2$	0,2506
3º	$r = 2$	$r = 3$	0,8117
4º	$r = 3$	$r = 4$	0,5293
5º	$r = 4$	$r = 5$	0,5546

Já na Tabela 4.45, o primeiro teste do máximo autovalor indica rejeição da hipótese de que não existe vetor de cointegração entre as variáveis, legitimando o resultado do teste do traço. Ao testar a hipótese nula de que $r = 1$, o teste indica que a mesma não pode ser rejeitada. Desta forma, como no teste do traço, segundo o teste do máximo autovalor, existem evidências que há apenas um vetor de cointegração entre as séries das variáveis relacionadas a demanda industrial de energia elétrica a um nível de significância de 0,01.

Os testes de cointegração indicaram que as séries cointegram, sugerindo a existência de relação de longo prazo entre elas. Uma vez que o objetivo deste trabalho é determinar a influência das variáveis estudadas na demanda de energia elétrica, optou-se por trabalhar com o vetor de cointegração normalizado em relação a variável D_i . O vetor normalizado, resultante do procedimento de Johansen pode ser visto na Tab. 4.46.

Tabela 4.54 – Vetor de cointegração normalizado para o setor Industrial

D_i	R_i	T_i	E_i	S_i	Tendência
1	-0,48491	0,13113	-0,28828	0,06954	-0,072566
Desvio Padrão	(0,00937)	(0,05170)	(0,05752)	(0,04227)	0,0030653

A partir da Tabela 4.46 foi possível escrever a equação que expressa a relação de longo prazo entre as variáveis, dados pela Eq. 4.5.

$$Di = 0,48491Ri - 0,13113Ti + 0,28828Ei - 0,06954Si + 0,072566t \quad (4.5)$$

Segundo a Equação 4.5, a variável renda apresenta sinal positivo, já as variáveis tarifa e preço do substituto da energia elétrica exibiram sinais negativos, concordando com a teoria econômica. A tendência também apresentou sinal positivo, assim como a variável que representa o preço das máquinas e equipamentos. A demanda de energia elétrica industrial sofre mais alterações com a variação da renda, seguida pela alteração do preço das máquinas e equipamentos, preço da tarifa e por final, com o preço do substituto da energia elétrica. Parece que, com a elevação da tarifa e do preço do substituto para energia elétrica, a indústria tende a gastar menos energia elétrica. No entanto, também parece que, com o aumento da renda da empresa, esta tende a aumentar o investimento na própria empresa, aumentando a demanda. Somente a variável Ei que apresentou um sinal oposto ao esperado, o que também foi encontrado por Schmidt e Lima (2004). No entanto, os autores resolveram o problema com a inclusão da variável Si , para o bem substituto à energia, em que foi utilizado como *proxy* o índice de preços de óleos e lubrificantes. Assim, com a inclusão de uma nova variável os resultados foram alterados significativamente. Já no trabalho de Mattos (2004), a inserção da variável Si causou multicolinearidade no modelo, desse modo o autor optou pela não inclusão desta variável. No trabalho de Viana (2010), a inclusão da variável Si fez com que a hipótese da não existência de cointegração entre as variáveis não pudesse ser rejeitada, resultando na alteração do *proxy* utilizado para a esta variável. Diante do exposto, foi proposto uma mudança na análise da variável Si , em que deixou de ser uma variável endógena e foi considerada uma variável exógena.

O resultado para a determinação do número de defasagens do novo modelo VAR a ser testado é apresentado na Tab 4.47.

Tabela 4.55 – Seleção de defasagens do modelo VAR (continua)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Di; Ri; Ti; Ei; Si; DYi3	Constante e	1	-7,2992	-5,9202*	-6,8085*
	Tendência**	2	-7,3891*	-5,3206	-6,6532
	Constante	1	-5,8811	-4,6745*	-5,4518*
		2	-6,0993*	-4,2031	-5,4246
	Sem	1	-4,7883	-3,7540	-4,4203
	Constante	2	-5,5838*	-3,8600*	-4,9705*
	Constante e	1	-4,7631	-3,3840*	-4,2724*
	Tendência**	2	-4,7916*	-2,7231	-4,0556
	Constante	1	-4,6143*	-3,4076*	-4,1850*
		2	-4,3876	-2,4914	-3,7129
Di; Ri; Ti; Ei; Si; DYi6	Sem	1	-4,0337*	-2,9994*	-3,6657*
	Constante	2	-3,9085	-2,1848	-3,2952
	Constante e	1	-3,8872	-2,5082*	-3,3965
	Tendência**	2	-4,2529*	-2,1844	-3,5169*
	Constante	1	-3,7179	-2,5113*	-3,2886
		2	-4,0899*	-2,1938	-3,4153*
	Sem	1	-3,0805	-2,0462	-2,7125
	Constante	2	-3,7750*	-2,0513*	-3,1617*
	Constante e	1	-3,2086*	-1,8296*	-2,7179*
	Tendência**	2	-3,0324	-0,9638	-2,2964
Di; Ri; Ti; Ei; Si; DYi2	Constante	1	-2,8290*	-1,6224*	-2,3997*
		2	-2,6420	-0,7458	-1,9673
	Sem	1	-2,1030*	-1,0687*	-1,7350*
	Constante	2	-2,0945	-0,3707	-1,4812
	Constante e	1	-3,0632*	-1,8565*	-2,6338*
	Tendência**	2	-2,8284	-0,9323	-2,1538
	Constante	1	-2,8792*	-1,8449*	-2,5112*
		2	-2,6318	-0,9080	-2,0185
	Sem	1	-2,2614*	-1,3995*	-1,9547*
	Constante	2	-2,2393	-0,6879	-1,6874

Tabela 4.56 – Seleção de defasagens do modelo VAR (conclusão)

Variáveis	Termos da Equação	Defasagens	Akaike	Schwartz	Hannan-Quinn
Di; Ri; Ti; Ei; Si; DYi5	Constante e	1	-3,0463	-1,6672*	-2,5556*
	Tendência**	2	-3,1452*	-1,0767	-2,4092
	Constante	1	-2,9321	-1,7254*	-2,5027*
		2	-2,9789*	-1,0827	-2,3043
	Sem	1	-2,2296	-1,1953*	-1,8616
	Constante	2	-2,6645*	-0,9407	-2,0512*
	Constante e	1	-2,9926	-1,6136*	-2,5020*
	Tendência	2	-3,1164*	-1,0479	-2,3805
Di; Ri; Ti; Ei; Si; DYi1	Constante**	1	-2,9669	-1,7603*	-2,5376*
		2	-3,0372*	-1,1411	-2,3626
	Sem	1	-2,2142	-1,1800*	-1,8463
	Constante	2	-2,8574*	-1,1337	-2,2441*

Obs: 1 - * indica o menor valor obtido pelo critério; 2 - ** indica o termo da equação que apresentou os menores valores dos critério de informação.

De acordo com os resultados dos critérios de informação, apresentados na Tabela 4.47, o modelo mais parcimonioso utiliza uma variável dummy, DYi3, apresenta constante e tendência determinística, além de utilizar uma defasagem.

Para determinar o posto da matriz de cointegração pelos testes de Johansen, foram consideradas as situações: sem constante, com constante apenas no vetor de cointegração (restrita) ou fora do vetor de cointegração (irrestrita) e com tendência restrita ou irrestrita. Os resultados obtidos para o teste do traço e o teste do máximo autovalor para o modelo escolhido que utilizou constante sem restrições podem ser visualizados na Tab. 4.48 e Tab. 4.49.

Tabela 4.57 – Teste do traço de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3 e sem constante

Teste	Hipótese nula (H_0)	Hipótese alternativa (H_1)	Valor-p
1º	$r = 0$	$r > 0$	0,0000
2º	$r \leq 1$	$r > 1$	0,4714
3º	$r \leq 2$	$r > 2$	0,7774
4º	$r \leq 3$	$r > 3$	0,7816

Por meio da Tabela 4.48, foi possível notar que o primeiro teste do traço indica a rejeição da hipótese nula, encontrando evidências de que exista pelo menos um vetor de cointegração entre elas. No entanto, no segundo teste do traço, a hipótese de que $r \leq 1$ não pode ser rejeitada para um nível de significância de 0,01, indicando a existência de no máximo um vetor de cointegração entre as variáveis.

Tabela 4.58 – Teste do máximo autovalor de Johansen para o modelo com Di, Ri, Ti, Ei, DYi3, sem constante

Teste	Hipótese nula (H_0)	Hipótese alternativa (H_1)	Valor-p
1º	$r = 0$	$r = 1$	0,0000
2º	$r = 1$	$r = 2$	0,3094
3º	$r = 2$	$r = 3$	0,7109
4º	$r = 3$	$r = 4$	0,7762

Já na Tabela 4.49, o primeiro teste do máximo autovalor indica rejeição da hipótese de que não existe vetor de cointegração entre as variáveis, legitimando o resultado do teste do traço. Ao testar a hipótese nula de que $r = 1$, o teste indica que a mesma não pode ser rejeitada. Desta forma, como no teste do traço, segundo o teste do máximo autovalor, existem evidências que há apenas um vetor de cointegração entre as séries das variáveis relacionadas a demanda industrial de energia elétrica a um nível de significância de 0,01.

Os testes de cointegração indicaram que as séries cointegram, sugerindo a

existência de relação de longo prazo entre elas. Assim, optou-se por trabalhar com o vetor de cointegração normalizado em relação a variável D_i , que pode ser visto na Tab. 4.50.

Tabela 4.59 – Vetor de cointegração normalizado para o setor Industrial

D_i	R_i	T_i	E_i
1	-1,1420	0,70779	1,4288
Desvio Padrão	(0,04627)	(0,24821)	(0,18934)

A partir da Tabela 4.50 foi possível escrever a equação que expressa a relação de longo prazo entre as variáveis, dados pela Eq. 4.6.

$$D_i = 1,1420R_i - 0,70779T_i - 1,4288E_i \quad (4.6)$$

Segundo a Equação 4.6, a variável renda apresenta sinal positivo, mostrando que com o aumento da renda ocorre o aumento da demanda, já as variáveis tarifa e preço das máquinas e equipamentos exibiram sinais negativos, mostrando que com o aumento destas variáveis, a indústria tende a diminuir a sua demanda de energia elétrica, concordando com a teoria econômica e trabalhos como Schmidt e Lima (2004), Mattos (2004), Viana (2010) e Silveira (2017), entre outros. Notou-se que a demanda sofre mais alterações com a variação do preço dos equipamentos e máquinas, seguida pela influência da variação da renda, do que com a variação da tarifa.

Os resultados da avaliação das suposições do modelo teórico por meio do diagnóstico dos resíduos, apresentados na Tab. 4.51, indicaram a normalidade dos resíduos do modelo, não apresentando autocorrelação e nem heterocedasticidade.

Tabela 4.60 – Resultados da avaliação das suposições do modelo VAR para o setor Industrial

Teste	Estatística de Teste	Valor-p
Jarque-Bera	2,7864	0,2483
Ljung-Box Q'	15,8480	0,6674
ARCH	111,3050	0,2067

O modelo apresentou uma relação de longo prazo dentro do esperado pela teoria econômica. Em sequência, partiu-se para a aplicação do modelo de correção de erros.

Modelo de correção de erros

Com a finalidade de realizar os ajustes dos modelos para que se possa determinar a variação de curto prazo e realizar previsões também de curto prazo, é aplicado um modelo de correção de erros vetorial (VEC). Nesse modelo, o vetor de cointegração de longo prazo é utilizado como variável explicativa (u_t) do termo de correção do erro. O resultado do VAR-VEC é expresso pela Eq. 4.7.

$$\Delta Di_t = 0,0063Si + 0,0553DYi3 + 0,0016(u_{t-1}) \quad (4.7)$$

O modelo apresentou coeficiente de determinação $R^2 = 0,7115$ e $R^2_{ajustado} = 0,6955$, sugerindo que alterações nas variáveis consideradas explicam 71,15% das variações da demanda de energia elétrica industrial.

O coeficiente do termo erro na Eq. 4.5 indica que a velocidade de ajustamento em relação ao equilíbrio de longo prazo é $0,0016$, ou seja, 0,16% da diferença entre as relações de longo e curto prazo é corrigida a cada ano. A constante não apresentou coeficiente significativo, por isto não fez parte da Eq. 4.7.

Os resultados da análise das suposições do modelo teórico estão mostrados na Tab. 4.52, observando-se que resíduos são normalmente distribuídos, não apresentam autocorrelação e não apresentam heterocedasticidade.

Tabela 4.61 – Testes de qualidade do modelo VAR-VEC para o setor industrial

Teste	Estatística de Teste	Valor-p
Jarque-Bera	1,7482	0,4172
Ljung-Box Q'	11,5770	0,9029
ARCH	317,7550	0,2302

Previsões

Para validação do modelo, foi realizada a previsão para a demanda de energia elétrica para os anos de 2011 a 2017, utilizando o modelo VAR e o modelo VAR-VEC. A validação destes modelos foi realizada por meio das medidas: raiz do erro quadrático médio (RMSE), erro absoluto médio (MAE), erro absoluto percentual médio (MAPE) e U de Theil, calculados com a métrica original e com a transformação logarítmica, e estão apresentadas na Tab. 4.53.

Tabela 4.62 – Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE), Erro Absoluto Percentual Médio (MAPE) e estatística U de Theil para os modelos de previsão

Métrica	Modelo	RMSE	MAE	MAPE (%)	U de Theil
Logarítmica	VAR-VEC	0,2018	0,1646	1,0219	0,0129
	VAR	0,3150	0,2646	1,6425	0,0231
Original	VAR-VEC	2.283.083,857	1.851.625,976	18,702	0,4635
	VAR	3.862.966,129	3.197.038,980	32,215	0,2321

Visando complementar a análise, foram construídos intervalos de confiança para os valores previstos, com nível de confiança de 95%. As Tabs. 4.54 e 4.55 apresentam os intervalos de confiança e os valores observados, pela métrica original e logarítmica. Para o modelo VAR apenas o ano de 2011 ficou dentro do intervalo de confiança em ambas as métricas, já no modelo VAR-VEC, nos anos de 2015, 2016 e 2017 os valores observados ficaram de fora do intervalo de confiança.

Tabela 4.63 – Intervalos de confiança para o modelo VAR de previsão pela métrica logarítmica e original (continua)

Métrica	Ano	Valor observado	Previsão	IC-95%
logarítmica	2011	16,1759	16,1968	16,1582; 16,3080
	2012	16,1426	16,2579	16,2189; 16,4152
	2013	16,1825	16,3277	16,2798; 16,5066
	2014	16,1645	16,4005	16,3386; 16,5871
	2015	16,1067	16,4714	16,3936; 16,6575
	2016	16,0658	16,5382	16,4446; 16,7193
	2017	16,1045	16,6024	16,4943; 16,7764

Tabela 4.64 – Intervalos de confiança para o modelo VAR de previsão pela métrica logarítmica e original (conclusão)

Métrica	Ano	Valor observado	Previsão	IC-95%
original	2011	10.595.000,0	10.819.212,0	10409.162,6; 12.091.097,5
	2012	10.248.000,0	11.501.051,2	11060.436,3; 13.460.028,5
	2013	10.665.149,0	12.331.838,9	11754.621,0 14.747.748,7
	2014	10.474.837,0	13.263.865,3	12466.704,4; 15.983.332,2
	2015	9.886.935,4	14.220.731,9	13171.487,3; 17.149.318,6
	2016	9.490.690,0	15.220.731,9	13861.031,9; 18.243.218,9
	2017	9.865.617,2	16.229.773,8	14.567.331,3; 19.315.279,0

Tabela 4.65 – Intervalos de confiança para o modelo VAR-VEC de previsão pela métrica logarítmica e original

Métrica	Ano	Valor observado	Previsão	IC-95%
Logarítmica	2011	16,1759	16,1740	16,0843; 16,2636
	2012	16,1426	16,2153	16,0886; 16,3420
	2013	16,1825	16,2570	16,1019; 16,4121
	2014	16,1645	16,2990	16,1199; 16,4781
	2015	16,1067	16,3407	16,1405; 16,5411
	2016	16,0658	16,3819	16,1625; 16,6013
	2017	16,1045	16,4230	16,1860; 16,6599
Original	2011	10.595.000,0	10.574.732,6	9.668.415,5; 11.566.007,9
	2012	10.248.000,0	11.020.602,1	9.709.283,0; 12.509.013,2
	2013	10.665.149,0	11.490.187,9	9.838.993,7; 13.418.487,8
	2014	10.474.837,0	11.983.017,5	10.017.749,2; 14.333.815,2
	2015	9.886.935,4	12.494.298,8	10.226.705,1; 15.264.691,9
	2016	9.490.690,0	13.018.474,2	10.454.049,8; 16.211.963,3
	2017	9.865.617,2	13.564.762,6	10.702.971,8; 17.191.764,2

Esta diferença pode ter ocorrido devido à queda da demanda de energia elétrica ocorrida no ano de 2012 e nos anos de 2014 à 2016, o que pode sinalizar uma diminuição na produção das empresas.

Os gráficos que representam os modelos de previsão VAR e VAR-VEC podem ser

visualizados nas Figuras 4.23 e 4.24 respectivamente.

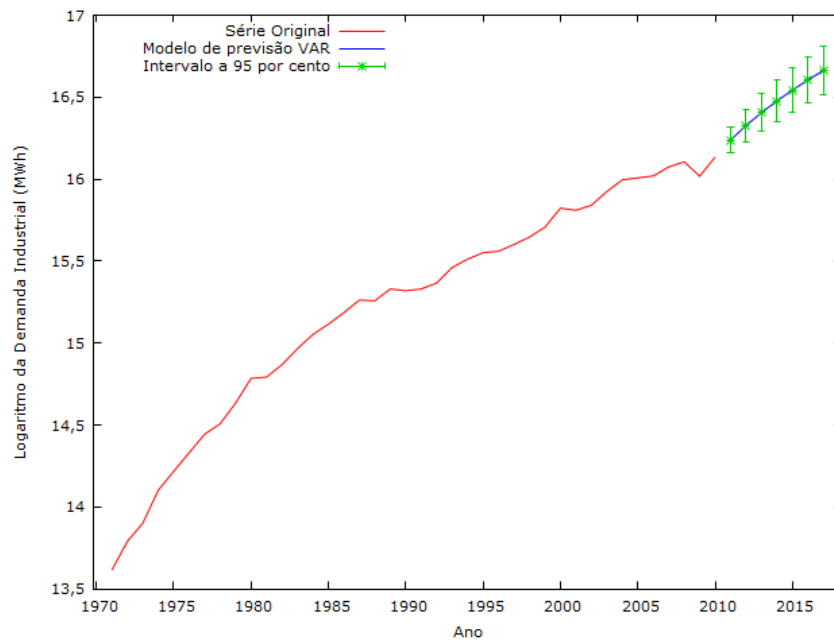


Figura 4.23 – Gráfico do modelo de previsão VAR para demanda industrial de energia elétrica/RS

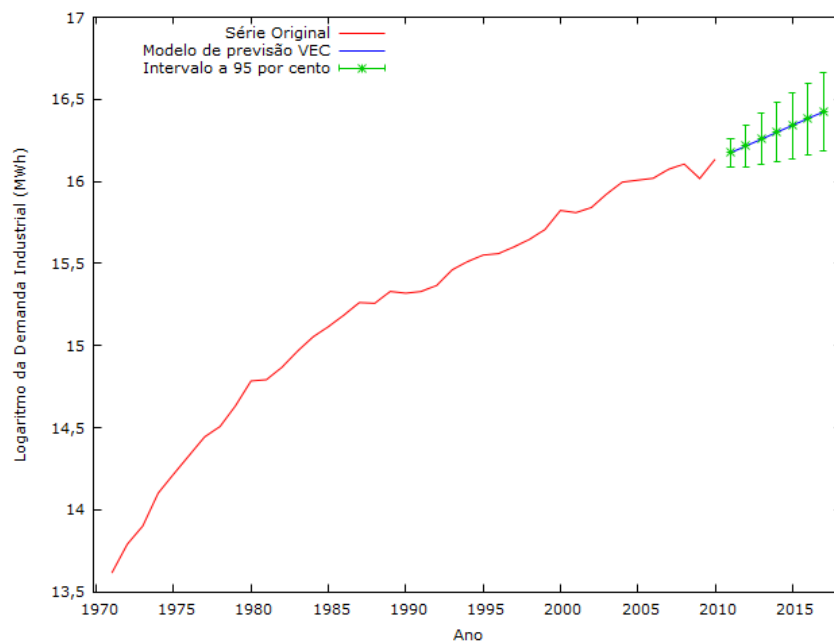


Figura 4.24 – Gráfico do modelo de previsão VAR-VEC para demanda industrial de energia elétrica/RS

4.3.3. Comparações para o setor Industrial

Técnica utilizada: ARIMA x VAR x VAR-VEC

Pela análise dos indicadores de qualidade, entre os modelos estudados, o modelo ARIMA (1,1,2) apresentou os melhores resultados para a previsão em ambas as métricas. O modelo VAR foi o que apresentou os maiores erros, sendo o modelo VAR-VEC melhor do que este, principalmente a curto prazo.

Nas previsões, o modelo ARIMA (1,1,2) obteve estimativas pontuais abaixo dos valores observados nos anos de 2011 e 2013 e no ano de 2012 apresentou estimativas pontuais muito próximas dos valores observados. Os intervalos de confiança foram os mais amplos, abrangendo todos os valores observados. Já os intervalos de confiança determinados a partir do modelo VAR conseguiram abranger os valores observados apenas no ano de 2011, sendo que nos demais anos, os valores observados ficaram abaixo do limite inferior do intervalo de confiança. Os intervalos de confiança dos modelos VAR-VEC conseguiram abranger os valores observados até 2014, pois no período de 2015 à 2017 ocorreram valores observados abaixo do limite inferior do intervalo de confiança.

Para finalizar, o melhor resultado para o MAPE foi o encontrado para o ARIMA (1,1,2), que foi de 0,40% para a escala logarítmica e 6,76% para a métrica original. Para a escala logarítmica, o modelo VAR apresentou o pior resultado (1,6425%) e o modelo VAR-VEC ficou com resultado intermediário (1,0219%), o mesmo acontecendo para a métrica original (32,215% e 18,702%, respectivamente).

Se analisarmos no curto prazo (um ano a frente), o modelo ARIMA (1,1,2) apresenta os piores resultados: na escala logarítmica (0,316%) e pela métrica original (4,979%). O modelo VAR foi intermediário tanto para a escala logarítmica (0,129%), quanto para a escala original (2,116%), enquanto o modelo VAR-VEC foi o melhor na métrica logarítmica (0,012%) e a métrica original (0,191%).

Comparação com o modelos de Silveira (2017)

A Tabela 4.56 apresenta os coeficientes encontrados para os modelos VAR encontrados por Silveira (2017) (Modelo BR) e neste estudo (Modelo RS).

Tabela 4.66 – Coeficientes do modelo VAR Industrial/RS e BR

	Di	Ri	Ti	Ei	Constante
Modelo RS	1	1,1420	-0,70779	-1,4288	-
Modelo BR	1	1,8640	-0,4690	-0,09774	2,810

O modelo VAR mostrado por Silveira (2017) para o setor industrial a nível nacional utilizou uma defasagem e eliminou a variável Si. O modelo desenvolvido neste trabalho também apresentou apenas uma defasagem e considerou a variável Si como exógena, sendo possível notar uma elasticidade menor para a renda e maiores para a tarifa e preço dos eletrodomésticos, e sem o acréscimo da constante.

A Tabela 4.57 mostra os coeficientes dos modelos VAR-VEC encontrados por Silveira (2017) e neste estudo.

Tabela 4.67 – Coeficientes do modelo VAR-VEC Industrial/RS e BR

	ΔDi	Di1-BR	Di2-BR	DYi3-RS	u_i-1	Si
Modelo RS	1	-	-	0,0553	0,0016	0,0063
Modelo BR	1	-0,1208	-0,1068	-	-0,1966	-

Em ambos os estudos foi encontrado apenas um vetor de cointegração, mas os modelos foram compostos por variáveis diferentes, sendo que no modelo de Silveira (2017) houve uma correção maior. O modelo VAR-VEC nacional apresentou um MAPE de 6,80%, considerando um ano após o período estudado, enquanto que o modelo encontrado neste estudo apresentou MAPE de apenas 0,191%. Estes achados concordam com o encontrado nos demais setores estudados neste trabalho, mostrando a importância da regionalização das séries estudadas para uma melhor previsão de demanda de energia elétrica.

5. CONSIDERAÇÕES FINAIS

Este trabalho buscou modelar a demanda de energia elétrica ocorrido no Rio Grande do Sul nas últimas décadas, para cada uma das três principais classes consumidoras: residencial, comercial e industrial, utilizando três técnicas: modelos Box-Jenkins, modelos VAR e modelos VAR-VEC. Posteriormente foi realizada uma comparação com estudo realizado utilizando a metodologia VAR-VEC a nível nacional, evidenciando a necessidade da regionalização deste tipo de estudo.

No setor residencial, as melhores estimativas de previsão foram encontradas pelo modelo VAR que apresentou os melhores indicadores. Já para os setores comercial e industrial, os melhores modelos para previsão foram os encontrados pela metodologia de Box-Jenkins. Neste sentido, os resultados encontrados não foram os esperados, pois em dois dos três setores estudados, os modelos Box-Jenkins apresentaram os melhores indicadores de qualidade. Provavelmente isto tenha ocorrido em decorrência da mudança no padrão de comportamento das séries no período de validação, com uma queda a partir do ano de 2014. Como os modelos de Box-Jenkins são mais robustos, conseguiram não se deixar afetar tanto por esta mudança. Observe que os intervalos de confiança obtidos por esta metodologia foram bem mais amplos. Aqui se verifica o princípio da incompatibilidade, estabelecido por Zadeh, que diz que alta complexidade é incompatível com alta precisão. Os modelos de Box-Jenkins, mais simples e menos precisos, apresentaram melhores resultados.

Dentre os melhores modelos encontrados em cada setor, o modelo VAR do setor residencial apresentou os melhores resultados, com um MAPE de 5,91%, após veio o setor comercial com um MAPE de 6,59%, e por fim o setor industrial com um MAPE de 6,76%, pela métrica original.

Na comparação dos modelos regionais, obtidos pelas metodologias VAR e VAR-VEC, com os modelos a nível nacional, desenvolvidos por Silveira (2017), constatou-se que, em todos os setores, os modelos para o estado do Rio Grande do Sul apresentaram MAPE melhor, confirmando a ideia da necessidade de regionalização dos modelos de previsão de demanda de energia elétrica em países como o Brasil, que possui grande dimensão territorial. Isto evidencia a importância deste estudo, sugerindo a sequência deste tipo de trabalho para outros estados do Brasil.

Para trabalhos futuros, também se sugere a análise de outras variáveis que possam

estar relacionadas à demanda de energia elétrica, além de revisão de alguns proxies, principalmente para o caso do setor industrial. Salienta-se que é possível que a demanda industrial de energia elétrica possa ser afetada por outras *proxies* além das aqui estudadas, como: taxas de câmbio, tipo de produção dos grandes consumidores industriais da região, autoprodução de energia e prática de cogeração de energia, entre outros. Também pode ser estudada a influência de fatos sociopolíticos, em que políticas econômicas como isenções fiscais podem decidir a instalação ou não de um grande consumidor. Em relação ao setor comercial, poderia ser avaliada a possibilidade de encontrar outro *proxy* capaz de representar melhor os equipamentos elétricos utilizados por esse setor. Outra sugestão se refere ao estudo do efeito da inflação nas previsões. Além disso, estudar as propriedades dos dados que compõem as séries estudadas, verificando a influência de algumas de suas propriedades no comportamento de cada método estudado também pode ser pertinente, além do estudo da utilização de outros tipos de função matemática na modelagem.

Por fim, cabe considerar que os objetivos deste trabalho foram alcançados, uma vez que foram encontradas as relações de curto e longo prazo entre a demanda de energia elétrica para o estado do Rio Grande do Sul para os três principais setores consumidores, por três metodologias diferentes.

REFERÊNCIAS

ABID, C. A.; História da Energia Elétrica no Rio Grande do Sul, 2007. Disponível em: http://www.carlosadib.com.br/elet_fatos%20-%20interior%20%20RS.html. Acesso em: dez. de 2017.

AKAIKE, Maximum likelihood identification of Gaussian autoregressive moving average models. **Biometrika**, v.60, n.2, p. 255-265, 1973. Disponível em: http://www.jstor.org/stable/2334537?seq=1#page_scan_tab_contents. Acesso em: jan de 2016.

ANDRADE, T.A.; LOBÃO, W. J. A. **Elasticidade renda e preço da demanda residencial de energia elétrica no Brasil**. Texto para discussão n 489, IPEA, 1997. Disponível em: http://repositorio.ipea.gov.br/bitstream/11058/2162/1/td_0489.pdf. Acesso em: jan de 2016.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, ANEEL. **Atlas de Energia Elétrica do Brasil**, 2008. Brasília. Disponível em: <http://www2.aneel.gov.br/arquivos/PDF/atlas3ed.pdf>. Acesso em: jan de 2018.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, ANEEL. **Por Dentro da Conta de Luz**, 2017. Brasília. Disponível em: http://www2.aneel.gov.br/arquivos/PDF/Cartilha_tarifas_Enersul.pdf. Acesso em: maio de 2018.

BUENO, R. D. L.S. **Econometria de séries temporais**. 2 ed. São Paulo: Editora Cengage Learning, 2011. 360p.

COMPANHIA ESTADUAL DE ENERGIA ELÉTRICA, CEEE. “**História do Grupo CEEE**”. Disponível em: <http://www.cee.com.br/pportal/cee/Component/Controller.aspx?CC=3237>. Acesso em: dez. de 2017.

DE LORENZO, H.C. O setor elétrico brasileiro: passado e futuro. **Perspectivas: Revista de Ciências Sociais**, v. 24/25, p. 147-170, 2002. Disponível em: <http://seer.fclar.unesp.br/perspectivas/article/view/406/291>. Acesso em: dez de 2017.

DEBNATH, K. B.; MOURSHED, M. Forecasting methods in energy planning models. **Renewable and Sustainable Energy Reviews**. v. 88, p. 297–325, 2018. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1364032118300200>. Acesso em: maio de 2018.

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. Acesso em: jan de 2018.

ELETROBRAS. “**Boletim siese: consumo de energia elétrica no Brasil**”. Disponível em: <http://www.ipeadata.gov.br>. Acesso em: jan de 2018.

ENGLE, R. F.; GRANGER, C. W. Co-integration and error correction: representation, estimation, and testing. *Econometrica*, v.55, n.2, p.251–276, 1987. Disponível em: http://www.jstor.org/stable/1913236?seq=1#page_scan_tab_contents. Acesso em: mar de 2018.

EMPRESA DE PESQUISA ENERGÉTICA, EPE. Nota Técnica DEA 19/15. **Projeção da demanda de energia elétrica 2016-2020**, 2015. Disponível em <http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-160/topico-168/Anu%C3%A1rio%20Estat%C3%ADstico%20de%20Energia%20El%C3%A9trica%202015.pdf>. Acesso em: abr de 2018.

EMPRESA DE PESQUISA ENERGÉTICA, EPE. Nota Técnica DEA 19/15. **Projeção da demanda de energia elétrica 2016-2020**, 2017a. Disponível em <http://www.epe.gov.br/default.aspx>. Acesso em: jan de 2018.

EMPRESA DE PESQUISA ENERGÉTICA, EPE. Nota Técnica DEA 03/15. **Projeção da demanda de energia elétrica 2015-2024**, 2017b. Disponível em [http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-245/topico-261/DEA%20001_2017%20-%20Proje%C3%A7%C3%B5es%20da%20Demanda%20de%20Energia%20El%C3%A9trica%202017-2026_VF\[1\].pdf](http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-245/topico-261/DEA%20001_2017%20-%20Proje%C3%A7%C3%B5es%20da%20Demanda%20de%20Energia%20El%C3%A9trica%202017-2026_VF[1].pdf). Acesso em: jan de 2018.

EMPRESA DE PESQUISA ENERGÉTICA, EPE. **Anuário estatístico de energia elétrica 2018 ano base 2017**, 2018. Disponível em: <http://epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-160/topico-168/Anuario2018vf.pdf>. Acesso em: mar de 2019.

ETTESTOL, I. Estimating residential demand for electricity with smooth transition regression. **Unpublished Manuscript**, NTNU, Trondheim, Norway, 2002. Disponível em: http://www.scielo.br/scielo.php?script=sci_nlinks&ref=000194&pid=S1413-8050200900010000400012&lng=en. Acesso em: dez de 2017.

FERREIRA, J. S. P.; SILVA, N. R. Determinação dos parâmetros α e b da função de Cobb-Douglas, através do método dos mínimos quadrados. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 3, n. 1, p. 1-2, 2015.

FUNDAÇÃO GETÚLIO VARGAS, FGV. Alterações na lista de produtos e nas ponderações do Índice de Preços por Atacado (IPA). **Nota Técnica - Fundação Getúlio Vargas**, 2007. Disponível em: <http://www.ipeadata.gov.br/doc/Nota%20reforma%20IPA.pdf>. Acesso em: jan de 2018.

FUNDAÇÃO GETÚLIO VARGAS, FGV. **Índices de preços**, 2018. Disponível em: <https://portalibre.fgv.br>. Acesso em: jun de 2018.

GOMES, A. C.; ABARCA, C. D. G.; FARIA, E. A. S. T.; FERNANDES, H. H. O. O setor elétrico. **BNDES. BNDES**, v. 50, p. 1964-1973, 2002. Disponível em: http://www.bndes.gov.br/SiteBNDES/export/sites/default/bndes_pt/Galerias/Arquivos/conhecimento/livro_setorial/setorial14.pdf. Acesso em: dez de 2017.

GRETTL - GNU Regression, Econometrics and Timeseries Library, version 2018d, 2018. Disponível em: <http://gretl.sourceforge.net/>. Acesso em: abr de 2018

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. Porto Alegre: Editora AMGH, 2011.

HALVORSEN, B.; LARSEN, B. M. The flexibility of household electricity demand over time. **Resource and Energy Economics**, v. 23, n. 1, p. 1-18, 2001. Disponível em: <http://www.sciencedirect.com/science/article/pii/S092876550000035X>. Acesso em: jan de 2018.

HANNAN, E. J.; QUINN, B. G. The determination of the order of an autoregression. **Journal of the Royal Statistical Society. Series B**, v.41, p.190–195, 1979.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA, IPEA. **Macroencômico**, 2018. Disponível em: <http://www.ipeadata.gov.br/Default.aspx>. Acesso em: fev. de 2018.

IRFFI, G.; CASTELAR, I.; SIQUEIRA, M.L.; LINHARES, F.C. Previsão da demanda por energia elétrica para classes de consumo na região Nordeste, usando OLS dinâmico e mudança de regime. **Economia Aplicada**, v. 13, n. 1, p. 69-98, 2009. Disponível em: <http://www.scielo.br/pdf/ecoa/v13n1/v13n1a04.pdf>. Acesso em: jan. de 2018.

JOHANSEN, S. Statistical analysis of cointegration vectors. **Journal of economic dynamics and control**, v.12, n.2-3, p.231–254, 1988. Disponível em: <http://www.sciencedirect.com/science/article/pii/0165188988900413>. Acesso em: jan de 2018.

KAVANAGH, K. Short Term Demand Forecasting for the Integrated Electricity Market. **Student Journal of Energy Research**, v. 2, n. 1, art. 1. 2017. Disponível em: <https://arrow.dit.ie/cgi/viewcontent.cgi?article=1005&context=sjer>. Acesso em: maio de 2018.

KWIATKOWSKI, D.; PHILIPS, P. C.B.; SCHIMDT, P.; SCHIN, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root. **Journal of econometrics**, v. 54, n. 1-3, p. 159-178, 1992. Disponível em: <http://www.sciencedirect.com/science/article/pii/030440769290104Y> Acesso em: jan de 2018.

LEITE, S. Q. **Projeções para a demanda por energia elétrica no Brasil, 2006-2015**, 2006. Dissertação de Mestrado Profissionalizante, RJ: Faculdades Ibemec.

- LIMA, H. C. G. **A demanda setorial de energia elétrica em Pernambuco**, 2011. Dissertação de Mestrado, PE: Universidade Federal de Pernambuco. Disponível em: <http://repositorio.ufpe.br/handle/123456789/1240>. Acesso em: jan. de 2018.
- MATTOS, L. B. **Demanda de energia elétrica no estado de Minas Gerais: 1970-2002**, 2004. Dissertação de Mestrado, MG: Universidade Federal de Viçosa.
- MATTOS, L. B.; LIMA, J. E. de. Demanda residencial de energia elétrica em Minas Gerais: 1970-2002. **Nova Economia**. v.15, n.3, p.31–52, 2005. Disponível em: <http://www.scielo.br/pdf/neco/v15n3/v15n3a02.pdf>. Acesso em: jan de 2018.
- MODIANO, E. M. Elasticidade-renda e preços da demanda de energia elétrica no Brasil. Rio de Janeiro PUC/RJ, 1984. (Texto para discussão, 68).
- MORETTIN, P. A. **Econometria Financeira: Um curso em Séries Temporais Financeiras**. São Paulo: Editora Edgard Blucher, 2006.
- NUNES, Vinícius Pinheiro. Modelagem E Previsão Do Índice Nacional De Preços Ao Consumidor: O Uso Da Metodologia De Box-Jenkins. In: Conferência Sul Em Modelagem Computacional, 7º ed., 2016, Rio Grande. **Anais**. Rio Grande: Universidade Federal do Rio Grande (FURG), 2016. 806 pp. 776 – 782.
- R CORE TEAM. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <https://www.R-project.org>. Acesso em: 05 abr. 2018.
- SCHMIDT, C. A. J.; LIMA, M. A. A demanda por energia elétrica no Brasil. **Revista Brasileira de Economia**, v.58, n.1, p.67–98, 2004.
- SCHWARZ, G. Estimating the dimension of a model. **Annals of Statistics**, v.6, n.2, p.461–464, 1978. Disponível em: <http://qwone.com/~jason/trg/papers/schwarz-dimension-78.pdf>. Acesso em: dez. de 2017.
- SCHWERT, G. William. 1989. “Test for Unit Roots: A Monte Carlo Investigation.” **Journal of Business and Economic Statistics** 7 (2), pp 147-159.
- SECRETARIA DE PLANEJAMENTO, GOVERNANÇA E GESTÃO, “**Atlas Socioeconômico do Rio Grande do Sul**”, 2017. Disponível em: <https://atlassocioeconomico.rs.gov.br/geracao-e-transmissao-de-energia>. Acesso em: jan de 2019.
- SILK, J. I.; JOUTZ, F. L. Short and long-run elasticities in US residential electricity demand: a co-integration approach. **Energy Economics**, v. 19, n. 4, p. 493-513, 1997. Disponível em: <http://www.sciencedirect.com/science/article/pii/S014098839701027X>. Acesso em: jan de 2018.
- SILVA, Vinícius Montenegro, 2016. Avaliação Do Método Dos Mínimos Quadrados No Ajuste De Um Modelo De Regressão Linear. In: **Conferência Sul Em Modelagem Computacional**, 7º ed., 2016, Rio Grande. **Anais**. Rio Grande: Universidade Federal do Rio Grande (FURG), 2016. 806 pp. 558 – 565.

SILVEIRA, A. G. **Estudo da demanda energia elétrica no Brasil**, 2017. Dissertação de Mestrado, RS: Universidade Federal de Rio Grande (FURG). Disponível em: <https://ppgmc.furg.br/publica>. Acesso em: jan 2017.

SIQUEIRA, M. L.; CORDEIRO JUNIOR, H. H.; CASTELAR, I. A demanda por energia elétrica no Nordeste brasileiro após o racionamento de 2001-2002: previsões de longo prazo. **Pesquisa e Planejamento Econômico**. v. 36, n. 1., p. 137-178, 2006. Disponível em: <http://ppe.ipea.gov.br/index.php/ppe/article/viewFile/26/6>. Acesso em: fev. de 2018.

VASCONCELLOS, M. A. S.; ALVES, D. **Manual de econometria: nível intermediário**. 1ª edição. São Paulo: Atlas, 2000.

VIANA, G. I. M. N. **Um modelo para projeções para demanda por energia elétrica, 2009-2017 e a evolução do custo social e tarifa ótima para o Brasil**, 2010. Dissertação de Mestrado, AL: Universidade Federal de Alagoas.

VIANA, G. I. M. N.; SILVA, A. L. M. Um modelo para projeções para demanda por energia elétrica, 2009-2017 para o setor residencial no Brasil. **Revista Brasileira de Energia**, v.20, n.1, p.107-126, 2014. Disponível em: <http://new.sbpe.org.br/artigo/327/> Acesso em: fev. de 2018.

WESTLEY, G. D. **New directions in econometric modeling of energy demand: with applications to Latin America**. USA: Johns Hopkins University Press, 1992.

APÊNDICE 01 – CÓDIGO UTILIZADO NA VERIFICAÇÃO DOS MODELOS E DA GERAÇÃO DA ESTATÍSTICA U DE THEIL

```

library(tseries)
library(forecast)
library(fpp)
library(aTSA)
library(urca)
setwd("C:\\Users\\USUARIO\\Documents\\Mestrado\\FURG\\Dissertação")
ytr <- read.table("ytr.txt", header = TRUE)
ytr <- ts(ytr,start = c(1971), end=c(2010), frequency = 1)
ytr
lytr<-log(ytr)
lytr
modelo1<-auto.arima(lytr, d=1, start.q = 0, max.q = 2, start.p = 0, max.p = 2,
seasonal = FALSE)
modelo1
summary(lytr)
ytc <- read.table("ytc.txt", header = TRUE)
ytc <- ts(ytc,start = c(1971), end=c(2010), frequency = 1)
ytc
lytc <-log(ytc)
lytc
modelo2<-auto.arima(lytc, start.p = 0,max.p = 2, d=1, start.q = 0, max.q = 2,
seasonal = FALSE)
modelo2
summary(lytc)
yti <- read.table("yti.txt", header = TRUE)
yti <- ts(yti,start = c(1971), end=c(2010), frequency = 1)
yti
lyti <-log(yti)
lyti

```

```

modelo3<-auto.arima(lyti, start.p = 0,max.p = 3, d=1, start.q = 0, max.q = 3,
seasonal = FALSE)
modelo3
summary(lyti)
setwd("C:\\Users\\USUARIO\\Documents\\Mestrado\\FURG\\Dissertação\\Capít
ulos\\u de theil 2011 2017")
“Residencial”
ytr <- read.table("ytr.txt", header = TRUE)
ytr <- ts(ytr,start = c(2011), end=c(2017), frequency = 1)
ytr
y111 <- read.table("y111.txt", header = TRUE)
y111 <- ts(y111,start = c(2011), end=c(2017), frequency = 1)
y111
y210 <- read.table("y210.txt", header = TRUE)
y210 <- ts(y210,start = c(2011), end=c(2017), frequency = 1)
y210
y110 <- read.table("y110.txt", header = TRUE)
y110 <- ts(y110,start = c(2011), end=c(2017), frequency = 1)
y110
yvar <- read.table("yvar.txt", header = TRUE)
yvar <- ts(yvar,start = c(2011), end=c(2017), frequency = 1)
yvar
yvar_vec <- read.table("yvar_vec.txt", header = TRUE)
yvar_vec <- ts(yvar_vec,start = c(2011), end=c(2017), frequency = 1)
yvar_vec
TheilU(ytr, y111, type = c(2, 1), na.rm = FALSE)
TheilU(ytr, y210, type = c(2, 1), na.rm = FALSE)
TheilU(ytr, y110, type = c(2, 1), na.rm = FALSE)
TheilU(ytr, yvar, type = c(2, 1), na.rm = FALSE)
TheilU(ytr, yvar_vec, type = c(2, 1), na.rm = FALSE)

“Comercial”
ytc <- read.table("ytc.txt", header = TRUE)
ytc <- ts(ytc,start = c(2011), end=c(2017), frequency = 1)

```

```

ytc
y210 <- read.table("y210.txt", header = TRUE)
y210 <- ts(y210,start = c(2011), end=c(2017), frequency = 1)
y210
y111 <- read.table("y111.txt", header = TRUE)
y111 <- ts(y111,start = c(2011), end=c(2017), frequency = 1)
y111
yvar <- read.table("yvar.txt", header = TRUE)
yvar <- ts(yvar,start = c(2011), end=c(2017), frequency = 1)
yvar
yvar_vec <- read.table("yvar_vec.txt", header = TRUE)
yvar_vec <- ts(yvar_vec,start = c(2011), end=c(2017), frequency = 1)
yvar_vec
TheilU(ytc, y210, type = c(2, 1), na.rm = FALSE)
TheilU(ytc, y111, type = c(2, 1), na.rm = FALSE)
TheilU(ytc, yvar, type = c(2, 1), na.rm = FALSE)
TheilU(ytc, yvar_vec, type = c(2, 1), na.rm = FALSE)

“Industrial”
yti <- read.table("ytc.txt", header = TRUE)
yti <- ts(ytc,start = c(2011), end=c(2017), frequency = 1)
yti
y112 <- read.table("y112.txt", header = TRUE)
y112 <- ts(y112,start = c(2011), end=c(2017), frequency = 1)
y112
y111 <- read.table("y111.txt", header = TRUE)
y111 <- ts(y111,start = c(2011), end=c(2017), frequency = 1)
y111
yvar <- read.table("yvar.txt", header = TRUE)
yvar <- ts(yvar,start = c(2011), end=c(2017), frequency = 1)
yvar
yvar_vec <- read.table("yvar_vec.txt", header = TRUE)
yvar_vec <- ts(yvar_vec,start = c(2011), end=c(2017), frequency = 1)
yvar_vec

```

```
TheilU(yti, y112, type = c(2, 1), na.rm = FALSE)
```

```
TheilU(yti, y111, type = c(2, 1), na.rm = FALSE)
```

```
TheilU(yti, yvar, type = c(2, 1), na.rm = FALSE)
```

```
TheilU(yti, yvar_vec, type = c(2, 1), na.rm = FALSE)
```