



Universidade Federal de Uberlândia  
Faculdade de Matemática

Bacharelado em Estatística

**APLICAÇÃO DE GRÁFICOS DE  
CONTROLE EM DADOS  
AUTOCORRELACIONADOS DO  
CONSUMO DE ENERGIA ELÉTRICA EM  
MINAS GERAIS**

**Dalila Fraga Silva**

Uberlândia-MG

2021



Dalila Fraga Silva

**APLICAÇÃO DE GRÁFICOS DE  
CONTROLE EM DADOS  
AUTOCORRELACIONADOS DO  
CONSUMO DE ENERGIA ELÉTRICA EM  
MINAS GERAIS**

Trabalho de conclusão de curso apresentado à Co-ordenação do Curso de Bacharelado em Estatística como requisito parcial para obtenção do grau de Bacharel em Estatística.

Orientadora: Profa. Dra. Aurélia Aparecida de Araújo Rodrigues

**Uberlândia-MG**

**2021**





**Universidade Federal de Uberlândia  
Faculdade de Matemática**

**Coordenação do Curso de Bacharelado em Estatística**

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Bacharel em Estatística.

Uberlândia, \_\_\_\_\_ de \_\_\_\_\_ de 20\_\_\_\_\_

**BANCA EXAMINADORA**

---

Profª. Dra. Aurélia Aparecida de Araújo Rodrigues

---

Prof. Dr. José Waldemar da Silva

---

Prof. Dr. Quintiliano Siqueira Schroden Nomelini

**Uberlândia-MG  
2021**



# AGRADECIMENTOS

À Deus, pela minha vida, e por me permitir ultrapassar todos os obstáculos encontrados ao longo da realização do curso.

A minha mãe Lindalva e aos meus filhos Ana Luiza e Cicero Henry, que me incentivaram nos momentos difíceis e compreenderam a minha ausência enquanto eu me dedicava à realização deste curso.

Aos amigos, que sempre estiveram ao meu lado, pela amizade incondicional e pelo apoio demonstrado ao longo de todo o período de tempo em que me dediquei a este curso.

À professora Dra. Aurélia Aparecida de Araújo Rodrigues, por ter sido minha orientadora e ter desempenhado tal função com dedicação e amizade. E aos demais professores, por todos os conselhos, pela ajuda e pela paciência com a qual guiaram o meu aprendizado.

Às pessoas com quem convivi a longo desses anos de curso, que me incentivaram e que certamente tiveram impacto na minha formação acadêmica.

Aos meus colegas de turma, por compartilharem comigo tantos momentos de descobertas e aprendizado e por todo o companheirismo ao longo deste percurso.

Por fim, sou grata a todos que de alguma forma, direta ou indiretamente participaram dessa realização.



# RESUMO

O presente trabalho foi desenvolvido com a finalidade de aplicar gráficos de controle a dados autocorrelacionados do consumo de energia elétrica em Minas Gerais do período de jan/2015 a jun/2021, o qual foi dividido em categorias de consumo, sendo elas: consumo total, consumo residencial, consumo comercial e consumo industrial de energia elétrica.

No início do estudo foram realizadas análises descritivas dos dados para avaliar seu comportamento e testes para verificar se eles possuem distribuição normal.

Para descrever o comportamento das séries do consumo de energia elétrica, foi utilizada a função de autocorrelação FAC e FACp e o teste de Ljung-box para avaliar a presença de autocorrelação. Adicionalmente, foi avaliada a estacionariedade, com teste de Dickey-Fuller aumentado (ADF), o teste de Cox Stuart para avaliar tendência e o teste de Fisher para verificar sazonalidade.

Em seguida foi ajustado o modelo auto-regressivo, ou seja, modelo AR(1) para as séries de consumo total, residencial e comercial e AR(2) para consumo industrial com intuito de utilizar os resíduos do modelo no monitoramento do processo. Não foram avaliados nesse trabalho outros modelos, porém o modelo AR(1) e AR(2) foi suficiente pois eles apresentaram resíduos não autocorrelacionados.

Em relação aos gráficos de controle, o monitoramento do processo é dividido em duas fases, na fase 1, consiste no planejamento dos gráficos e na fase 2 consiste na validação do processo. Foram utilizados dados de janeiro/2015 a dezembro/2019 para a fase 1 e de janeiro/2020 a junho/2021 para a fase 2. Durante o planejamento (fase 1) foi escolhido os gráficos de medidas individuais e amplitude móvel (X e AM) pois o tamanho da amostra é igual a um e eles se mostraram adequados para dar continuidade na validação do processo (fase 2).

Na apresentação dos gráficos foram exibidos os gráficos de X e AM na fase 1 com os dados normais (autocorrelacionados) e com os resíduos do modelo AR(1) e AR(2), disposto a defrontá-los. E também foram exibidos os gráficos de X e AM da fase 1 agregado a fase 2.

**Palavras-chave:** Gráficos de Controle, Autocorrelação, Modelo Auto-regressivo, Monitoramento, Consumo de Energia Elétrica.



# ABSTRACT

The present academic work was developed with the purpose of applying control charts to data auto-correlated electricity consumption in Minas Gerais for the period Jan/2015 to jun/2021, which was divided into consumption categories, namely: total consumption, consumption residential, commercial consumption and industrial consumption of electricity.

At the beginning of the study, descriptive analyzes of the data were performed to assess the behavior of the data and verify whether the data have a normal distribution.

To describe the behavior of the series of electrical energy consumption, the function of auto correlation FAC and FACp to assess the presence of auto correlation and the Ljung-box test. Additionally, stationarity was evaluated, with increased Dickey-Fuller test (ADF), the Cox Stuart test to assess trends and the Fisher test to check seasonality. Then, the auto regressive model was adjusted, that is, the AR(1) model for the series of total consumption, residential and industrial and AR(2) for industrial consumption in order to use model residues in process monitoring.

Regarding the control charts, in phase 1, with data from January/2015 to December/2019, the control charts of individual measurements and moving range (X and AM) were obtained and compared for the original data and for the residuals of auto regressive models. The X and AM graphs for the residuals of the models were more adequate and were adopted for real-time monitoring (phase 2)

In this way, the X and AM control chart is viable for process monitoring of electricity consumption.

**Keywords:** Control Graphs, Auto correlation, Auto regressive Model, Monitoring, Electricity Consumption..



# SUMÁRIO

Lista de Figuras	I
Lista de Tabelas	III
Lista de Abreviações e Símbolos	V
<b>1 Introdução</b>	<b>1</b>
<b>2 Objetivos</b>	<b>3</b>
2.1 Objetivos Gerais	3
2.2 Objetivos Específicos	3
<b>3 Metodologia</b>	<b>5</b>
3.1 Dados	5
3.2 Análise Descritiva	5
3.3 Teste de Normalidade	6
3.4 Processos Estocásticos	6
3.4.1 Séries Temporais	6
3.4.2 Processos Estacionários	7
3.4.3 Autocorrelação	9
3.4.4 Função de Autocorrelação	10
3.4.5 Teste de Ljung-box	11
3.4.6 Teste de Dickey-Fuller Aumentado (ADF)	11
3.4.7 Teste de Cox Stuart	13
3.4.8 Teste de Fisher	13
3.4.9 Modelo Autoregressivo	15
3.5 Gráficos de Controle	16
3.5.1 Gráficos de medidas individuais e amplitude móvel (X e AM)	19
<b>4 Resultados</b>	<b>21</b>
4.1 Gráficos de Controle X e AM, fase 1, dados históricos	21
4.1.1 Análise Descritiva, dados da fase 1	21
4.1.2 Avaliação das pressuposições dos gráficos de controle, fase 1	22
4.1.3 Gráficos de Controle, X e AM, fase 1	30
4.1.4 Fase 2 - Monitoramento	33
<b>5 Conclusões</b>	<b>37</b>
<b>Referências Bibliográficas</b>	<b>39</b>

<b>Apêndice A Apêndice</b>	<b>41</b>
A.0.1 Software R . . . . .	41
<b>Apêndice B Tabelas e Dados</b>	<b>45</b>
B.0.1 Tabela para os gráficos de controle . . . . .	45
B.0.2 Dados . . . . .	45

# LISTA DE FIGURAS

3.1	Autocorrelação	9
4.1	Boxplot do Consumo de Energia Elétrica	22
4.2	QQplot do Consumo de Energia Elétrica	22
4.3	Função de autocorrelação FAC e FACp do consumo total de energia elétrica de 2015 a 2019 em Minas Gerais.	23
4.4	Função de autocorrelação FAC e FACp do consumo residencial de energia elétrica de 2015 a 2019 em Minas Gerais.	24
4.5	Função de autocorrelação FAC e FACp do consumo comercial de energia elétrica de 2015 a 2019 em Minas Gerais.	24
4.6	Função de autocorrelação FAC e FACp do consumo industrial de energia elétrica de 2015 a 2019 em Minas Gerais,	25
4.7	FAC e FACp dos resíduos do modelo AR(1) no consumo total de energia elétrica	27
4.8	FAC e FACp dos resíduos do modelo AR(1) no consumo residencial de energia elétrica	27
4.9	FAC e FACp dos resíduos do modelo AR(1) no consumo comercial de energia elétrica	28
4.10	FAC e FACp dos resíduos do modelo AR(2) consumo industrial de energia elétrica	28
4.11	Gráficos de controle X e AM consumo total de energia elétrica e dos resíduos do modelo AR(1)	31
4.12	Gráficos de controle X e AM consumo residencial de energia elétrica e resíduos do modelo AR(1)	32
4.13	Gráficos de controle X e AM do consumo comercial de energia elétrica e resíduos do modelo AR(1)	32
4.14	Gráficos de controle X e AM do consumo industrial de energia elétrica e resíduos do modelo AR(2)	33
4.15	Gráfico de controle X e AM para os resíduos do consumo total de energia elétrica	34
4.16	Gráfico de controle X e AM para os resíduos do consumo residencial de energia elétrica	35
4.17	Gráfico de controle X e AM para os resíduos do consumo comercial de energia elétrica	35
4.18	Gráfico de controle X e AM para os resíduos do consumo industrial de energia elétrica	36



---

# LISTA DE TABELAS

3.1	Síntese dos gráficos de controle aplicados a processos (adaptado por Spiegel e Stephens em 2008) . . . . .	17
4.1	Análise descritiva do consumo de energia elétrica em Minas Gerais, 2015 a 2019 em MWh. . . . .	21
4.2	Resultados do teste de normalidade de Shapiro-Wilk . . . . .	23
4.3	Resultados do Teste de Ljung-Box . . . . .	25
4.4	Resultados do Teste de Estacionariedade de Dickey-Fuller Aumentado (ADF) . . . . .	26
4.5	Resultados do Teste de tendência de Cox-Stuart . . . . .	26
4.6	Resultados do Teste de Fisher . . . . .	26
4.7	Resultados do Teste de Ljung-Box dos resíduos do modelo . . . . .	29
B.1	Fonte Montgomery [14] . . . . .	45
B.2	Consumo de energia elétrica, em milhões de MWh, de jan/2015 a dez/2019 em Minas Gerais. . . . .	45
B.3	Consumo de energia elétrica em Minas Gerais de Janeiro de 2020 a junho de 2021 (EPE, 2021) . . . . .	47



---

# LISTA DE ABREVIações E SÍMBOLOS

## LISTA DE ABREVIações

$H_0$	Hipótese Nula
$H_1$	Hipótese Alternativa
$\bar{X}$	Média
ADF	Dickey-Fuller aumentado
Ap	Amplitude
AR	Modelo Autoregressivo
ARIMA	Modelo Autoregressivo Integrado a Médias Móveis
CEMIG	Companhia Energética de Minas Gerais S.A
CEQ	Controle Estatístico de Qualidade
CUSUM	Gráfico de Controle de Soma Cumulativa
CV	Coefficiente de Variação
Dp( )	Desvio Padrão
EPE	Empresa de Pesquisa Energética
EWMA	Gráfico de Controle de Medidas Individuais e média móvel
FAC	Função de Autocorrelação
FACp	Função de Autocorrelação parcial
IID	Independente e Identicamente Distribuídos
LC	Limite de Controle
LIC	Limite Inferior de Controle
LSC	Limite Superior de Controle
MA	Modelo de Médias Móveis
MQO	Mínimos Quadrados Ordinários
Var( )	Variância
X e AM	Gráfico de Medidas Individuais e Amplitude Móvel



# 1. INTRODUÇÃO

Qualidade sempre foi parte integrante de praticamente todos os produtos e serviços. No entanto, a conscientização de sua importância e a introdução de métodos formais para o controle e melhoria da qualidade têm tido um desenvolvimento evolutivo. Para Juran qualidade significa adequação ao uso [10]. Segundo Taguchi, a produção, o uso e o descarte de um produto sempre acarretam prejuízos ("perdas") para a sociedade; quanto menor for o prejuízo, melhor será a qualidade do produto [25]. Para Crosby, qualidade significa estar dentro das especificações [4]. De acordo com Deming, qualidade significa atender e, se possível, exceder as expectativas do consumidor [5]. Já Montgomery define que qualidade é inversamente proporcional a variabilidade, e sua melhoria é a redução dessa variabilidade nos processos e produtos [14].

No Controle Estatístico de Qualidade (CEQ), todas as decisões são tomadas com base em estudos de fatos e dados. Para aproveitar ao máximo esses dados, é preciso fazer a escolha das ferramentas mais adequadas para que haja eficácia nas análises que serão realizadas. As principais ferramentas de CEQ são aquelas usadas para o monitoramento do processo: histograma, folha de verificação, gráfico de Pareto, diagrama de causa-e-efeitos, diagrama de dispersão e gráfico de controle.

Em 1924, Walter A. Shewhart, desenvolveu o conceito estatístico de gráfico de controle, que é considerado, em geral, como o começo formal do Controle Estatístico de Qualidade (CEQ) [22]. Os gráficos de controle fazem a distinção dos tipos de causas que estão atuando no processo e por isso é praticamente uma regra utilizá-los no controle da qualidade. O processo de aplicação dos gráficos é dividido em duas fases, a primeira consiste na construção do gráfico e a segunda consiste no monitoramento do processo. Quando o monitoramento de características da qualidade mensuráveis (variáveis), por exemplo o consumo de energia elétrica, e o tamanho da amostra é igual a um, os gráficos de controle que podem ser utilizados é o gráfico de medidas individuais e o de amplitude móvel (X e AM).

Em alguns casos os gráficos de controle podem sinalizar o processo fora de controle mesmo que não esteja, e o aparecimento desses alarmes falsos é comum quando as observações são autocorrelacionadas, ou seja, quando as observações não são independentes [11]. De acordo com Ide na ausência de independência ocorre o aumento de alarme falso, decorrente da subestimação da variável aleatória de  $Y$  (dados das características de qualidade estudada), pois esta não se comportará de forma aleatória em torno da média, já que os dados passaram a ser dependentes das observações anteriores, assim a largura dos limites de controle fica estreita, uma vez que

ela depende da variável aleatória  $Y$  [9].

Nos casos em que as observações são autocorrelacionadas é utilizado a teoria de series temporais que definido por Ehlers, é uma coleção de observações feitas sequencialmente ao longo do tempo e a característica mais importante dessa coleção de observações é que as observações vizinhas são dependentes e o interesse é analisar e modelar esta dependência [6].

Na modelagem de series temporais para modelar a autocorrelação dos dados, é comumente a utilização dos modelos de natureza estocástica, a base destes modelos é a estacionariedade, isto é, as características do processo estocástico não se alteram com o tempo [17]. Na literatura os modelos autorregressivos (AR), os de médias móveis (MA), e o modelos de Box e Jenkins (1970) autorregressivo integrado de médias móveis (ARIMA) é bastante utilizado para esse tipo de análise.

Nesse trabalho será realizado a aplicação dos gráficos de controle, a fase 1 (planejamento) e a fase 2 (validação), com dados do consumo de energia elétrica no estado de Minas Gerais durante o período de jan/2015 a dez/2019 para a fase 1 e de jan/2020 a jun/2021 para a fase 2, os dados foram fornecidos pelo site da Empresa de Pesquisa Energética (EPE), e estão estratificados em consumo de energia elétrica total, residencial, comercial, industrial e outros [7].

De acordo com Reis, a energia desempenha um papel fundamental na vida humana: ao lado de transportes, telecomunicações e águas e saneamento, compõe a infraestrutura necessária para incorporar o ser humano ao denominado modelo de desenvolvimento vigente [19]. Por essa razão o tratamento dos temas energéticos no seio dessa infraestrutura será de maior importância para que caminhe na busca de um desenvolvimento sustentável. E como a energia elétrica é uma das formas mais utilizadas de energia é muito importante integrar seu enfoque no contexto de desenvolvimento sustentável.

Um sistema de energia elétrica visa garantir que a demanda por eletricidade possa ser atendida a qualquer momento, com requisitos mínimos de qualidade, confiabilidade, economia e segurança, atendendo a restrições ambientais para assegurar a proteção ao meio ambiente [3].

O uso consciente de energia elétrica traz benefícios para o meio ambiente e para os consumidores, seja residencial, comercial ou industrial. Com a redução dos desperdícios, há redução nos custos e nos investimentos no setor elétrico.

O emprego de gráficos de controle no monitoramento do processo de consumo de energia elétrica permite identificar o aumento excessivo e descontrolado na demanda de energia elétrica a tempo. Se o aumento do consumo de energia elétrica for devido a causas especiais, recomenda-se que intervenções sejam feitas pra trazer o processo de volta aos níveis estatísticos aceitáveis. As causas especiais atuantes no processo, podem ser: desperdício por falta de manutenção na rede de distribuição, desperdício por falta de conscientização dos consumidores entre outros.

Algumas medidas devem ser tomadas para reduzir ou eliminar as causas especiais, são elas: campanhas de conscientização, implantação de auditorias, aplicação de multas se identificado negligências do consumidor e outros.

## 2. OBJETIVOS

### 2.1 OBJETIVOS GERAIS

O estudo tem como objetivo geral utilizar gráficos de controle no monitoramento do consumo de energia elétrica residencial, comercial, industrial e total em Minas Gerais de jan/2015 a jun/2021 .

### 2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos que este trabalho visa atingir são:

- Descrever o comportamento do consumo de energia elétrica total, residencial, comercial e industrial em Minas Gerais de jan/2015 a dez/2019.
- Utilizar métodos de séries temporais para investigar a autocorrelação dos dados.
- Utilizar o gráfico de controle para medidas individuais e amplitude móvel (X e AM).
- Investigar, monitorar e validar se o processo está fora de controle estatístico.



## 3. METODOLOGIA

Neste trabalho foi utilizado o software R. A seguir apresenta-se a fundamentação teórica da metodologia estudada.

### 3.1 DADOS

Inicialmente, para a fase 1 foram coletados os dados do consumo de energia elétrica do estado de Minas Gerais de jan/2015 a dez/2019 no site da EPE. Eles foram organizados de acordo com a categoria de consumo, ou seja consumo total, residencial, comercial, industrial e outros na escala de Megawatt-hora (MWh) [7]. Para facilitar o estudo a escala foi convertida da seguinte forma:

$$\frac{1(MWh)}{1000000} \quad (3.1)$$

### 3.2 ANÁLISE DESCRITIVA

Utiliza-se métodos de Estatística Descritiva para organizar, resumir e descrever os aspectos importantes de um conjunto de características observadas ou comparar tais características entre dois ou mais conjuntos. A descrição dos dados também tem como objetivo identificar anomalias, até mesmo resultante do registro incorreto de valores, e dados dispersos, aqueles que não seguem a tendência geral do restante do conjunto [18].

A tendência central da distribuição de frequências de uma variável em um conjunto de dados é caracterizada pelo valor típico dessa variável. Essa é uma maneira de resumir a informação contida nos dados, pois seleciona um valor para representar todos os outros, a média aritmética e a mediana são medidas de tendência central [18].

As medidas de variabilidade explicam dispersão de um conjunto de dados essas medidas podem ser encontradas através do cálculo da variância, desvio padrão, coeficiente de variação e amplitude [15].

As medidas de posição tem como objetivo indicar a localização dos dados em geral. O quartil é uma dessas medidas onde divide o conjunto de dados em 4 partes iguais, tal que 25% dos dados correspondem ao primeiro quartil superior a esquerda, de 25% a 50% dos dados

correspondem ao segundo quartil superior a direita, de 50% a 75% dos dados correspondem ao terceiro quartil inferior a esquerda e de 75% a 100% dos dados correspondem ao quarto quartil inferior a direita.

O Boxplot é um gráfico proposto para a detecção de valores discrepantes (outliers), que são aqueles valores muito diferentes do restante do conjunto de dados. Esses valores discrepantes podem representar erros no processo de coleta ou de processamento dos dados, e nesse caso, devem ser corrigidos ou excluídos do banco de dados. No entanto, os outliers podem ser valores corretos, que por alguma razão, são muito diferentes dos demais valores. Nesse caso, a análise desses dados deve ser cuidadosa, pois, algumas estatísticas descritivas como a média e o desvio-padrão, são influenciadas por valores extremos [18]. Além disso, o boxplot mostra a distribuição dos dados.

### 3.3 TESTE DE NORMALIDADE

O teste de normalidade permite avaliar se um conjunto de observações são modelados através de uma distribuição normal, existem alguns testes como o de Shapiro-Wilk que permite verificar a normalidade dos dados e é possível também verificar essa condição através de gráficos como o QQplot.

### 3.4 PROCESSOS ESTOCÁSTICOS

Um processo estocástico pode ser definido como uma coleção de variáveis aleatórias  $Y$  ordenadas no tempo e definidas em um conjunto de pontos  $T$ , ou seja  $\{Y(t), t \in T\} \forall t \in T$ . Para cada  $t$ ,  $Y(t)$  tem uma distribuição de probabilidade que podem ser iguais ou não.

A série temporal é um processo estocástico onde  $\{Y(t), t \in T\}$  é chamado de espaço estados e  $Y(t)$  são chamados de estados [6].

#### 3.4.1 SÉRIES TEMPORAIS

Uma série temporal é o conjunto de observações  $Y(t)$ ,  $t \in T$  onde  $Y$  é variável de interesse e  $T$ : conjunto de índices. Para Ehlers, alguns dos objetivos de uma série temporal são [6]:

- Na análise descritiva, verificar as propriedades da série, padrão de tendência, estacionaridade, sazonalidade, outliers e alterações estruturais.
- Na explicação, usar a variação em uma série para explicar a variação em outra série,
- Na previsão, prever valores futuros com base em valores passados
- No controle de processos, utilizar modelagens que auxiliam no controle estatístico de qualidade. Neste trabalho, o objetivo da utilização de séries temporais será esse, modelar os dados da série que estão autocorrelacionados.

### 3.4.2 PROCESSOS ESTACIONÁRIOS

Um processo estocástico cujas características estatísticas (média, variância, autocorrelação) não variam com o tempo é classificado como um processo estocástico estacionário e pode-se dizer que uma mudança da origem de tempo será impossível de detectar, sendo assim não significativa de modo que o processo continuará o mesmo. Logo, no processo estacionário,  $f_{X_1}(x_1; t_1)$  e  $f_{X_2}(x_2; t_2)$  precisam ser idênticas, isso será possível se  $f_{X_1}(x_1; t_1)$  for independente de  $t$ . Logo a densidade de primeira ordem de um processo estacionário pode ser expressa como  $f_X(x; t) = f_X(x)$  [26].

Para Fischer um processo estacionário pode ser definido como um processo onde a distribuição conjunta e a distribuição condicional são invariantes aos deslocamentos no tempo, ou seja,  $P(x_1, x_2, \dots, x_n) = P(x_{1+k}, x_{2+k}, \dots, x_{n+k})$  que é igual a uma contante e consequentemente satisfazem as seguintes condições [8]:

- O valor médio mantém-se invariante no tempo, ou seja:

$$E(x_t) = E(x_{t+k}) = \mu \text{ (constante), para qualquer } t \in T; \quad (3.2)$$

Uma vez que:

$$E(x_t) = \int_{-\infty}^{\infty} x f(x_t) dx = \int_{-\infty}^{\infty} x f(x_{t+k}) dx = E(x_{t+k}) = \mu \quad (3.3)$$

Logo, pode-se dizer que a série não apresenta uma componente tendencial.

- A variância também se mantém-se invariante no tempo, ou seja:

$$Var(X_t) = \gamma_0 \text{ (constante), para qualquer } t \in T; \quad (3.4)$$

Uma vez que:

$$Var(X_t) = E(X_t - \mu)^2 = E(X_{t+k} - \mu)^2 = Var(X_{t+k}) \quad (3.5)$$

pois,

$$\begin{aligned} Var(X_t) &= E(X_t^2) - E(X_t)^2 = \\ &= \int_{-\infty}^{\infty} x^2 f(x_t) dx - \mu^2 = \int_{-\infty}^{\infty} x^2 f(x_{t+k}) dx - \mu^2 = \\ &= E(X_{t+k}^2) - E(X_{t+k})^2 = Var(X_{t+k}) \end{aligned} \quad (3.6)$$

- A covariância é definida em termos de valores defasados (autocovariância) e mede o grau de variação de segunda ordem entre dois elementos em tempos diferentes. A autocovariância entre  $Y_t$  e  $Y_{t-1}$  para o processo estocástico  $\{Y_t\}$  é:

$$\begin{aligned} Cov(X_t, X_{t-1}) &= \gamma_1 t = E[(Y_t - \mu_t)(Y_{t-1} - \mu_{t-1})] = E(Y_t Y_{t-1}) - \mu_t \mu_{t-1} = \\ &= \int_{-\infty}^{\infty} (y_t - \mu_t)(y_{t-1} - \mu_{t-1}) f_{Y_t Y_{t-1}}(Y_t, Y_{t-1}) dy_t dy_{t-1} \end{aligned} \quad (3.7)$$

E considerando a relação  $Y_t$  com as  $k$  mais recentes observações  $\{Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}\}$ , a  $k$ -ésima autocorrelação de  $Y_t$  é:

$$\begin{aligned} \gamma_k t &= E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})] = E(Y_t Y_{t-k}) - \mu_t \mu_{t-k} = \\ &= \int_{-\infty}^{\infty} (y_t - \mu_t)(y_{t-k} - \mu_{t-k}) f_{Y_t Y_{t-k}}(Y_t, Y_{t-k}) dy_t dy_{t-k} \end{aligned} \quad (3.8)$$

Fisher acredita que a importância de saber se uma série é estacionária reside ao fato de que, trabalhar com uma série que está em presença de uma função amostral do processo e que tem a mesma forma em todos os instantes de tempo, a obtenção das estimativas do processo se tornam bastante simples, pois na prática elas podem ser obtidas através da média, variância e covariância amostral [8].

### 3.4.3 AUTOCORRELAÇÃO

As variáveis que apresentam correlação ao longo do tempo, são ditas autocorrelacionadas, a correlação permite verificar se duas variáveis independentes estão associadas uma com a outra e é permitido dizer que resíduos que atendem os pressupostos de independência e normalidade são não correlacionados. Logo a autocorrelação significa a dependência temporal dos valores sucessivos dos resíduos.

Para Maia, as principais causas da autocorrelação são a inércia, quando a série apresenta ciclos de crescimento e decaimento e quando esse comportamento reflete nos fatores não observados ( $e_t$ ), é comum que mudanças na tendência ocorram lentamente, a falha na especificação ou ausência de um regressor, onde os erros expressariam um padrão sistemático devido a ausência dessas informações e as defasagens, pois em um período  $t$  geralmente dependem de informações defasadas do período  $t - 1$  [12]. Como consequência da autocorrelação nos resíduos, os estimadores de MQO continuam sendo não viesados e consistentes, mas deixam de ser eficientes, no contexto matemático, seja um estimador de MQO  $\hat{\beta}$ , então existe outro estimador  $\hat{\beta}^*$  tal que:

$$Var(\hat{\beta}^*) < Var(\hat{\beta}) \quad (3.9)$$

Análise gráfica para detectar a autocorrelação:

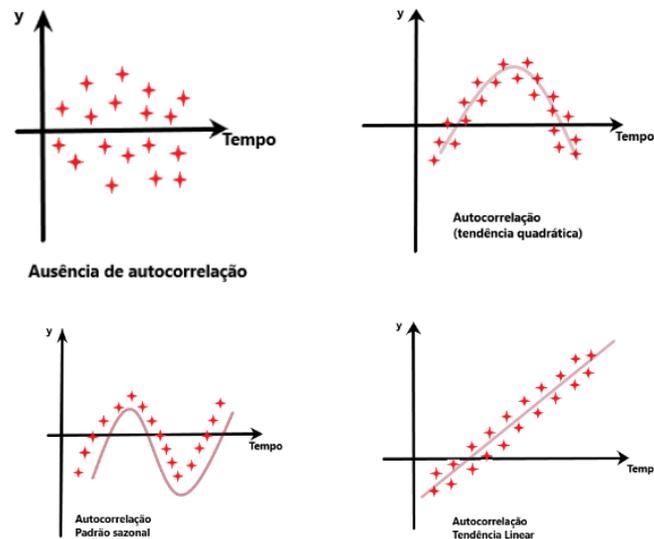


Figura 3.1: Autocorrelação

Alguns dos meios possíveis para detectar autocorrelação em uma série é pela representação do autocorrelograma com base no cálculo da função de autocorrelação FAC ou da autocorrelação parcial FACp 3.4.4. Alguns testes auxiliam na avaliação da autocorrelação, como o teste de Ljung-Box 3.4.5, o teste de Dickey Fuller Aumentado ADF 3.4.6, o teste de cox stuart 3.4.7 e o teste de fisher 3.4.8.

### 3.4.4 FUNÇÃO DE AUTOCORRELAÇÃO

Antes de conceituar a função de autocorrelação, é importante mencionar a definição de uma função de autocovariância, é basicamente a covariância entre  $Z_t$  e  $Z_{t-k}$  separados por um intervalo de tempo  $k$ .

$$\gamma_k = Cov[Z_t, Z_{t-k}] = E([Z_t - \mu][Z_{t-k} - \mu]) \quad (3.10)$$

Uma função de autocorrelação é uma função de autocovariância padronizada, que usamos para mensurar o comprimento e a memória de um processo. Para Ehlers ela pode ser definida como a razão entre a autocovariância e a variância de um conjunto de dados observados [6].

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{Cov[Z_t, Z_{t-k}]}{\sqrt{Var(Z_t) Var(Z_{t-k})}}, \quad \forall k = 1, 2, \dots \quad (3.11)$$

De fato para  $k = 0$ ,  $\rho_0 = 1$ .

Propriedades:

- A correlação entre  $Z_t$  e  $Z_{t+k}$  será a mesma entre  $Z_t$  e  $Z_{t-k}$ , ou seja,  $\rho_k = \rho_{-k}$
- $-1 < \rho_k < 1$
- Embora um processo estocástico tenha uma estrutura de autocovariância única o contrário não é verdadeiro em geral. É possível encontrar vários processos com a mesma função de autocorrelação.

A função de autocorrelação é uma ferramenta de grande importância para avaliar uma série temporal, sua representação gráfica da função de autocorrelação é chamada de autocorrelograma ou correlograma.

Em uma função de autocorrelação podemos dizer que:

- Conforme o tempo aumenta em uma série entre as observações é esperado um declínio na função de autocorrelação.
- As observações vizinhas de uma série guardam mais semelhanças entre si que as observações mais distantes então a função de autocorrelação é uma medida de interdependência das observações.
- É possível verificar a estacionariedade da série através do autocorrelograma, se o coeficiente não se aproximar de zero rapidamente conforme  $k$  cresce, podemos dizer que a série não é estacionária.

### 3.4.5 TESTE DE LJUNG-BOX

O teste de Ljung Box geralmente é empregado para testar a hipótese de que não há autocorrelação, ou seja os retornos são uma série de variáveis aleatórias independentes e identicamente distribuídas (IID).

Em 1970 Box e Pierce sugeriram um teste para diagnosticar autocorrelações que apresentavam valores elevados em diferentes defasagens. Em 1978 Ljung e Box, em vez de testar autocorrelação em cada defasagem, testa a autocorrelação global de um bloco definido. O teste de Ljung-Box permite detectar quebras específicas no comportamento aleatório [1].

A estatística é calculada da seguinte forma:

Primeiro é realizado o cálculo das estimativas de autocorrelação,

$$\hat{\rho}_k = \frac{\sum_{k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_1^n \hat{a}_t^2} \quad (3.12)$$

Em seguida é realizado o cálculo da estatística de teste,

$$Q^* = T(T-2) \sum_k^n \frac{\hat{\rho}_k^2}{T-k} \quad (3.13)$$

Onde:

$\hat{\rho}_k^2$  é a autocorrelação estimada dos resíduos na defasagem de "k";  
n é número de lags testados.

As hipóteses testadas são:

$H_0$ : As observações são IID

$H_1$ : As observações não são IID

Se  $Q_* > \chi_{1-\alpha; q}^2$ , sendo  $\chi_{1-\alpha; q}^2$  uma distribuição qui-quadrado com "q" graus de liberdade. Não rejeita  $H_0$

### 3.4.6 TESTE DE DICKEY-FULLER AUMENTADO (ADF)

Na equação 3.3 o valor médio mantém-se invariante no tempo, logo não apresenta tendência, essa é uma das condições para que uma série seja estacionária, portanto se uma série é tendenciosa podemos dizer que ela não é estacionária, porém se a série não apresenta tendência é necessário utilizar outros métodos para avaliar a estacionariedade da série.

Para Mattos, o teste de Dickey Fuller é o mais antigo e famoso método formal para verificar se uma série temporal é estacionária [13]. Ele foi introduzido em uma versão básica pelos trabalhos de Fuller (1976), Dickey(1976) e Dickey e Fuller (1979). Posteriormente, foi objeto de uma generalização no trabalho de Dickey e Fuller (1981). A versão generalizada é conhecida como teste de Dickey Fuller Aumentado (ADF). O desenvolvimento do teste foi motivado pela necessidade de verificar se uma série precisa ser diferenciada para se tornar estacionária.

A versão aumentada do teste de Dickey- Fuller consiste no estudo de uma regressão, e se difere do teste de Dickey-Fuller pois ele considera a existência de alguma estrutura de autocorrelação para os erros da equação de teste, se a estrutura da autocorrelação não for considerada, há perda de eficiência no estimador de mínimos quadrados ordinários (MQO) para  $\delta$ . A versão aumentada permite considerar termos defasados de  $\Delta Y_i$ . [13].

As hipóteses são:

$H_0 : \delta = 0$ , não é estacionária

$H_1 : \delta < 0$ , é estacionária

Para Dickey-Fuller, existem três abordagens para a realização do teste onde a regressão não possui intercepto e nem um termo de tendência determinística, que é comum utilizar quando a verificação gráfica não apresenta tendência determinística e é representada por:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + \varepsilon_t \quad (3.14)$$

Outra abordagem acontece também quando a verificação gráfica não apresenta tendência determinística, porém é considerado apenas o intercepto na regressão que é representada por:

$$\Delta y_t = \beta_0 + \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + \varepsilon_t \quad (3.15)$$

Na última abordagem é considerado na reta de regressão tanto o intercepto quanto o termo de tendência determinística, é ideal utilizar quando a verificação gráfica aponta a existência de tendência determinística ou quando se tem dúvidas sobre a existência dela, pois essa abordagem engloba as demais. A regressão é representada por:

$$\Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + \varepsilon_t \quad (3.16)$$

É dito que :

$\beta_0$  é o intercepto ou drift da série;

$\beta_1$  é o coeficiente de tendência;

$\delta$  é o coeficiente de presença de raiz unitária;

$m$  é o número de defasagens tomadas na série.;

$\varepsilon_t$  é um processo ruído branco;

$\Delta Y_{t-i}$  é termo de diferenças defasadas, para  $i = 1$ ,  $\Delta Y_1 = (y_{t-1} - y_{t-2})$ , para  $i = 2$ ,  $\Delta Y_2 = (y_{t-2} - y_{t-3})$ ... Esses termos de diferenças defasadas devem ser incluídos  $m$  vezes até que o erro não apresente autocorrelação.

A estatística de interesse é:

$$T = \frac{\widehat{\delta}}{se(\widehat{\delta})} \quad (3.17)$$

Onde:

$\widehat{\delta}$  é um estimador para  $\delta$ ;

$se(\widehat{\delta})$  é um estimador para desvio padrão do erro de  $\delta$ ;

Os valores críticos da estatística  $T$  foram tabelados por Dickey e Fuller (1981) através de simulação Monte Carlo, elas variam quando está na presença somente do intercepto, ou na presença somente da tendência ou na presença de ambos.

Portanto se  $T < \tau$  (valor crítico) há evidências suficientes para rejeitar a hipótese nula, ou seja, a série é estacionária.

### 3.4.7 TESTE DE COX STUART

O teste de Cox Stuart ou também conhecido como teste do sinal, consegue avaliar se uma série temporal é tendenciosa ou não. Suas hipóteses são:

$H_0 : P(Z_i < Z_{i+c}) = P(Z_i > Z_{i+c}), \forall_i$ : Não existe tendência

$H_1 : P(Z_i < Z_{i+c}) \neq P(Z_i > Z_{i+c}), \forall_i$ : Existe tendência

Primeiramente para realização do teste é necessário agrupar as observações em pares sendo eles :  $(Z_1 < Z_{1+c}), (Z_2 < Z_{2+c}), \dots, (Z_{N-c} < Z_N)$ , onde  $\frac{N}{2}$  para  $N$  par e  $\frac{N+1}{2}$  para  $N$  ímpar. Logo após associa-se um sinal de  $+$  se  $Z_i < Z_{i+c}$ , sinal de  $-$  caso  $Z_i > Z_{i+c}$  e caso  $Z_i = Z_{i+c}$  o par é desconsiderado. Coloca-se  $T$  como o número de pares com sinal  $+$  e  $t$  um valor tabelado, para  $n \leq 20$  utiliza-se a tabela a distribuição binomial com parâmetros  $p = \frac{1}{2}$  e  $n$  e para  $n > 20$  utiliza-se a tabela da distribuição normal com a média  $np$  e variância  $npq$ . Conclui-se que  $T \geq n - t$  rejeita  $H_0$  e a série é tendenciosa [16].

### 3.4.8 TESTE DE FISHER

O teste de Fisher para Morettin e Toloí pode ser utilizado para verificar a presença de sazonalidade ou periodicidade em séries temporais [16]. Para Borgatto os periodogramas servem

para determinar se as periodicidades são significativas [2]. Para aplicá-lo é recomendável que a série temporal tenha um número grande de observações, e para plotar o periodograma utiliza a frequência padrão que é dada por:

$$w_p = \frac{2\pi p}{N} \dots p = 0, 1, \dots, \frac{N}{2} \quad (3.18)$$

Onde  $N$  é o número de observações da série temporal.

As hipóteses são:

$H_0$ : não existe sazonalidade

$H_1$ : existe sazonalidade

Através do periodograma é possível construir uma estatística de teste para a maior periodicidade encontrada no periodograma que é dada por:

$$g = \frac{\max[I_p]}{\sum_{p=1}^{\frac{N}{2}} I_p} \quad (3.19)$$

Onde  $I_p$  é o valor do periodograma no período  $p$ .

O teste de fisher seria sob a hipótese nula,

$$z_\alpha = \left(\frac{\alpha}{n}\right)^{\frac{1}{(n-1)}} + 1 \quad (3.20)$$

Onde  $\alpha$  é a o nível de significância e  $n = \frac{N}{2}$ .

A regra de decisão é  $g > z$ , rejeita a hipótese nula. Morettin e Toloí apresentou uma aproximação dada por [16]:

$$P(g > \alpha) \equiv n(1 - \alpha)_{-1}^n \quad (3.21)$$

### 3.4.9 MODELO AUTOREGRESSIVO

Um modelo de regressão em séries temporais consiste em relacionar uma série temporal  $Z_t$  com funções do tempo e são mais úteis quando os parâmetros que descrevem a série permanecem constantes no tempo, ou seja quando são estacionárias. Os pressupostos de um modelo de regressão em séries temporais são a suposição de normalidade, independência e variância constante.

Um processo autoregressivo de ordem  $p$ , AR( $p$ ) para Ehlers consiste na suposição que  $\{\varepsilon_t\}$  seja um processo puramente aleatório com média zero, variância igual á  $\sigma_\varepsilon^2$  e com  $Z_t$  representado da seguinte forma [6]:

$$Z_t = \alpha_0 + \alpha_1 Z_{t-1} + \dots + \alpha_p Z_{t-p} + \varepsilon_t \quad (3.22)$$

onde,

$\alpha_1, \dots, \alpha_p$ , são parâmetros reais autoregressivos  
 $Z_{t-1}, \dots, Z_{t-p}$ , é série de tempo desfasado um período  
 $\varepsilon_t$ , é o termo do erro do modelo

O modelo é similar ao modelo de regressão múltipla, onde os valores passados de  $Z_t$  fazem o papel de regressoras.

Um processo autoregressivo de ordem um AR(1) é representado da seguinte forma:

$$Z_t = \alpha_0 + \alpha_1 Z_{t-1} + \varepsilon_t \quad (3.23)$$

Seja uma função polinomial de uma série constante,

$$Z_t = \mu_t + u_t, \quad t = 1, 2, \dots, T \quad (3.24)$$

onde  $u_t$  é um ruído branco e  $\mu_t$  é a tendência no tempo  $t$ .

É possível estimar um modelo de regressão representado por:

$$Z_t = \alpha_0 + \alpha_1 Z_{t-1} + u_t \quad (3.25)$$

Analogamente para o processo autoregressivo de ordem dois AR(2), que é representado por:

$$Z_t = \alpha_0 + \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + u_t \quad (3.26)$$

Logo após a aquisição do modelo, é necessário conferir as pressuposições de normalidade, por meios de testes ou, como a pressuposição de variância constante, por meios visuais e a

pressuposição de independência.

Nos casos em que é violada a pressuposição de independência é possível fazer o ajuste do modelo autoregressivo AR(p) utilizando os resíduos do modelo ( $u_t$ ), visto que está sendo considerado como um ruído branco, ou seja, o modelo relaciona  $u_t$ , o termo de erro no período  $t$ , com  $u_{t-1}$ , o termo de erro no período  $t - 1$  dessa forma:

$$u_t = \phi_1 u_{t-1} + a_t \quad (3.27)$$

Onde  $\phi$  define-se como o coeficiente de correlação entre  $u_t$  e  $u_{t-1}$ . Logo o modelo ajustado é representado da seguinte forma:

$$Z_t = \alpha_0 + \alpha_1 Z_{t-1} + \phi_1 u_{t-1} + a_t \quad (3.28)$$

Após o ajuste é interessante realizar novamente alguns testes para verificar se a série não é mais autocorrelacionada.

### 3.5 GRÁFICOS DE CONTROLE

Para Montgomery os principais objetivos do controle estatístico de qualidade CEQ é manter o processo operando em condição estável durante a maior parte do tempo e reduzir a variabilidade, os gráficos de controle são ferramentas que auxiliam atingir esses objetivos [14]. Além disso também auxiliam a reconhecer desvios em relação ao comportamento normal, a identificar oportunidades de melhoria e a verificar a eficácia das ações tomadas para corrigir desvios.

Os gráficos de controle baseiam-se em uma disposição gráfica gerada a partir de estatísticas como a média amostral, o desvio padrão amostral e outras estatísticas de uma característica da qualidade, é composto por uma linha média (LM), que corresponde o valor médio da amostra, por uma linha superior e outra inferior que são chamadas de limite superior de controle (LSC) e limite inferior de controle (LIC) gerados também através de cálculos estatísticos de acordo com o tipo de gráfico mais adequado para o modelo.

No teste de hipótese do gráfico de controle de acordo com Taconeli, se considerar o monitoramento da média do processo, as hipóteses são dadas por [24]:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Estatística de teste:

$$LSC = \mu_0 + L \frac{\sigma_0}{\sqrt{n}} \quad (3.29)$$

$$LM = \mu_0 \quad (3.30)$$

$$LIC = \mu_0 - L \frac{\sigma_0}{\sqrt{n}} \quad (3.31)$$

Onde,

$\mu_0$  e  $\sigma_0$  é a média e o desvio padrão do processo.

L é a amplitude dos limites de controle

Critério de decisão:

Se  $LIC < \bar{x} < LSC$ , sendo  $\bar{x}$  a média das amostras selecionadas periodicamente, não rejeita  $H_0$  e o processo está dentro das especificações de controle.

Existem vários tipos de gráficos, para escolher o mais adequado leva-se em consideração o tamanho da amostra, a periodicidade da coleta dos dados, o tipo variável a ser analisada, a escala das variáveis se contínuas ou discretas, a sensibilidade desejada para detectar as variações do gráfico e também a presença de autocorrelação das variáveis.

Para Spiegel os gráficos de controle são classificados em dois grupos que são os gráficos de controle para variáveis (característica da qualidade é mensurada em uma escala quantitativa) e gráficos de controle para atributos (característica da qualidade é mensurada em uma escala qualitativa), os termos variáveis e atributos estão associados a característica dos dados coletados no processo [23]. Alguns desses gráficos são:

Tabela 3.1: Síntese dos gráficos de controle aplicados a processos (adaptado por Spiegel e Stephens em 2008)

<b>Gráfico de Controle para Variáveis</b>	
<b>Gráficos</b>	<b>Estática de Interesse</b>
$\bar{X}$	Média das medidas em um subconjunto
R	Amplitude das médias em um subconjunto
S	Desvio Padrão das médias em um subconjunto
X e AM	Medidas individuais e amplitude móvel
CUSUM	Somas cumulativas dos desvios dos valores amostrais em relação ao valor alvo
EWMA	Média móvel exponencialmente ponderável
<b>Gráfico de Controle para Atributos</b>	
P	Proporção de unidades defeituosas em um subgrupo
NP	Total de unidades defeituosas em um subgrupo
U	Número médio de defeitos por unidade em um subgrupo

Para Montgomery os gráficos de  $\bar{X}$  e S são os mais utilizados quando o tamanho da amostra é moderadamente grande ( $n > 10$ ) ou o tamanho da amostra é variável [14]. Os gráficos de  $\bar{X}$  e R quando há vários dados em um período pequeno de tempo. Os gráficos de X e AM

quando o período da coleta dos dados é grande, ou quando é preciso investigar dado a dado. Os gráficos de CUSUM quando existe a necessidade de detectar pequenas variações no processo. Os gráficos de EWMA detectam pequenas mudanças de deslocamento na média do processo, esse gráfico atribui maior peso na observação mais recente ou na média do subgrupo, mesmo sendo mais utilizado com observações individuais, e assim diminui o peso das observações mais antigas.

Os gráficos de  $\bar{X}$  e AM (medidas individuais e Amplitude Móvel) são também utilizados quando não é possível calcular a amplitude em um subgrupo como por exemplo quando a amostra possui apenas um elemento que faz com que a variabilidade da amostra seja igual a zero [21].

Dos gráficos por atributos, os gráficos de np é utilizado para monitorar a quantidade de itens que não estão em conformidade em uma amostra de tamanho fixo, o gráfico de p é equivalente ao gráfico de np, porém o gráfico de p é utilizado quando o tamanho da amostra não pode ser mantido constante e o gráfico de U é utilizado para monitorar não conformidades por unidade de inspeção.

O desempenho do gráficos de controle é dado pela probabilidade de alarmes falsos e verdadeiros. Os alarmes falsos podem ser através dos erros do tipo I que ocorrem quando o gráfico aponta que o processo está fora de controle, porém não está e sua probabilidade é denominada por  $\alpha$  ou através dos erros do tipo II que ocorrem quando não é detectado que o processo está fora de controle, porém está sua probabilidade e é denominada por  $\beta$ . Logo quando não ocorrem alarmes falsos e o processo está sob controle, a probabilidade é  $1 - \beta$  e é denominado como poder do gráfico (Pd) [20].

Além dos pontos fora dos limites de especificação de controle, existem alguns padrões não aleatórios que podem indicar também que o processo não está sob controle. As 4 regras sugeridas por Western Electric são:

- 1 ou mais pontos fora dos limites de controle
- 2 em 3 pontos consecutivos fora dos limites de alerta ( $2\sigma$ ), mas ainda dentro dos limites de controle.
- 4 de 5 pontos consecutivos além dos limites de 1.
- Uma sequência mínima de 8 pontos consecutivos de um mesmo lado da linha central

A aplicação dos gráficos de controle são divididas em duas fases:

## FASE I

A construção dos gráficos de controle recomenda-se utilizar pelo menos 20 amostras e tem caráter exploratório que consiste em:

Se o processo está sob controle:

- Medir a capacidade do processo que tem como objetivo principal avaliar o potencial da produção em atender às especificações que são impostas (Não será abordado nesse trabalho)[24]. .
- Estimar os parâmetros do processo;
- Estabelecer limites de controle;

Se o processo não estiver sob controle é necessário intervir no processo identificar e eliminar todas as possíveis causas atribuíveis de variação, coletar novas amostras e repetir a fase 1 até que obtenha um processo sob controle para seguir para a fase 2.

## FASE 2

Na segunda fase da construção dos gráficos de controle tem o objetivo de monitorar e validar o processo. Com os limites estabelecidos e o comportamento do processo sob controle, a segunda fase consiste em analisar de acordo com os resultados obtidos na primeira fase novas amostras coletadas periodicamente.

### 3.5.1 GRÁFICOS DE MEDIDAS INDIVIDUAIS E AMPLITUDE MÓVEL (X E AM)

Há muitas situações onde o tamanho amostral usado para o controle do processo é  $n = 1$ . Isso acontece geralmente quando não é possível obter amostras de tamanho maior que um (1). Nesses casos não é possível estimar a variabilidade e o desvio padrão através da amplitude pois esses resultados não podem ser definidos para amostras de tamanho (1), é utilizado como estimativa da variabilidade da amplitude móvel duas, ou mais, observações sucessivas. O cálculo do LSC, LM e LIC são expressos da seguinte forma:

$$LSC = \bar{X} + \left( \frac{3 * \overline{AM}}{d_2} \right) = \bar{X} + (E_2 * \overline{AM}) \quad (3.32)$$

$$LM = \bar{X} \quad (3.33)$$

$$LIC = \bar{X} - \left( \frac{3 * \overline{AM}}{d_2} \right) = \bar{X} - (E_2 * \overline{AM}) \quad (3.34)$$

Onde,

$\overline{AM} = |x_1 - X_{i-1}|$  para  $i = 1, 2, \dots, n$  amostras da amplitude móvel;

$E_2 = \frac{3}{d_2}$ ,  $d_2$  é uma constante tabelada.

O gráfico X é sempre acompanhado a outro gráfico de amplitudes móveis AM para monitorar a variabilidade, o cálculo do LSC, LM e LIC são expressos da seguinte forma:

$$LSC = D4 * \overline{AM} \quad (3.35)$$

$$LM = \overline{AM} \quad (3.36)$$

$$LSC = D3 * \overline{AM} \quad (3.37)$$

$$\overline{AM} = \frac{1}{n} \sum_{i=1}^n AM_i \quad (3.38)$$

Onde,

$AM$  = Amplitudes Móveis

## 4. RESULTADOS

### 4.1 GRÁFICOS DE CONTROLE X E AM, FASE 1, DADOS HISTÓRICOS

Os dados da fase 1 foram apresentados na tabela B.2 (apêndice B)

#### 4.1.1 ANÁLISE DESCRITIVA, DADOS DA FASE 1

Na Tabela 4.1 são apresentados as principais medidas de estatísticas descritivas para o consumo de energia elétrica total, residencial, comercial, industrial e outros, em Minas Gerais, jan/ 2015 a dez/2019 em MWh.

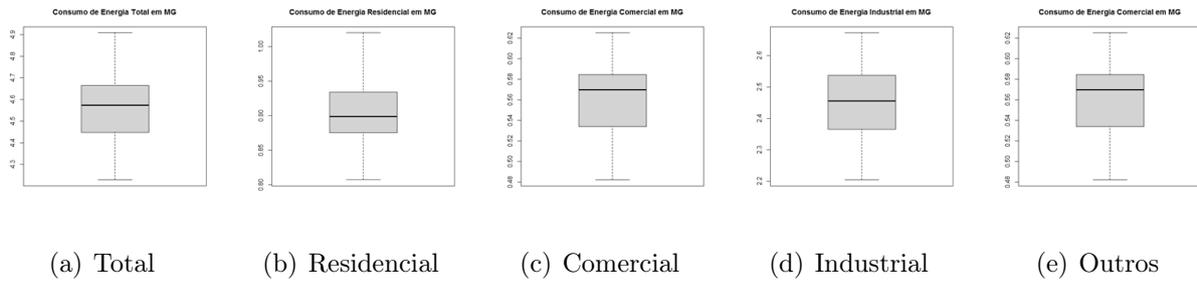
Tabela 4.1: Análise descritiva do consumo de energia elétrica em Minas Gerais, 2015 a 2019 em MWh.

Consumo	Total	Residencial	Comercial	Industrial	Outros
Média ( $\bar{x}$ )	4,5589	0,9029	0,5595	2,4574	0,6390
Mediana	4,5740	0,8990	0,5695	2,4555	0,6470
Variância ( $s^2$ )	0,0245	0,0019	0,0012	0,0119	0,0023
Desvio-Padrão (s)	0,1566	0,0444	0,0354	0,1091	0,0487
Coefficiente de Variação	3,4342	4,9208	6,3289	4,4402	7,6254
Amplitude	0,6800	0,2130	0,1430	0,4670	0,1810
1° quartil ( $Q_1$ )	4,4500	0,8755	0,5345	2,3690	0,6048
2° quartil ( $Q_2$ )	4,5740	0,8990	0,5695	2,4550	0,6470
3° quartil ( $Q_3$ )	4,6630	0,9343	0,5843	2,5350	0,6777

De acordo com os resultados da tabela 4.1, verifica-se que entre os consumos residencial, comercial, industrial e outros. Em média a categoria industrial apresenta o maior consumo (2,46 milhões de MWh mensais), e o consumo comercial apresenta o menor (0,56 milhões de MWh mensais). Analisando as medidas de variabilidade pelo coeficiente de variação percebe-se que a variabilidade dos dados em torno da média é baixa, logo pode-se dizer que a média é representativa.

Na Figura 4.1 não foram encontrados outliers nas observações de nenhum dos segmentos de consumo de energia elétrica de acordo com os gráficos de boxplot.

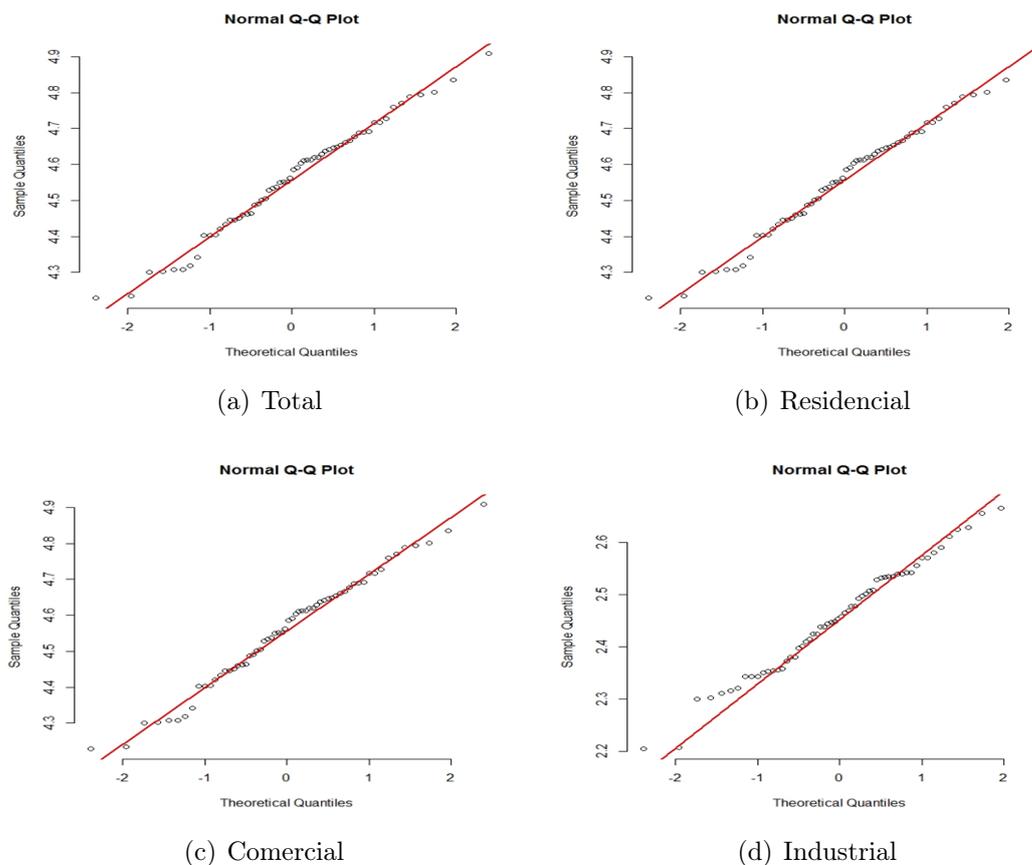
Figura 4.1: Boxplot do Consumo de Energia Elétrica



#### 4.1.2 AVALIAÇÃO DAS PRESSUPOSIÇÕES DOS GRÁFICOS DE CONTROLE, FASE 1

Para que os gráficos de controle sejam eficientes, os dados devem ser normalmente distribuídos e independente e identicamente distribuídos (sem autocorrelação)[14]. Na figura 4.2, os gráficos QQplot dos segmentos consumo de energia elétrica total, residencial, comercial e industrial indicam que os dados estão normalmente distribuídos.

Figura 4.2: QQplot do Consumo de Energia Elétrica



Na Tabela 4.2, o teste de Shapiro wilk do consumo total, residencial, comercial e industrial de energia elétrica confirmaram a apresentação gráfica do QQplot, pois há evidências suficientes,  $valor - p > 0,05$ , que as observações de cada segmento seguem a distribuição normal.

Tabela 4.2: Resultados do teste de normalidade de Shapiro-Wilk

Consumo	Total	Residencial	Comercial	Industrial
valor-p	0,6622	0,8706	0,1452	0,6051

De acordo com os gráficos da série, da função de autocorrelação FAC e FACp, nas Figuras 4.3, 4.4, 4.5 e 4.6, há indicativos de que o consumo de energia elétrica é autocorrelacionado (ou não é IID), pois na observação dos gráficos é possível detectar algumas tendências tanto linear quanto cíclica em todas as categorias de consumo.

Figura 4.3: Função de autocorrelação FAC e FACp do consumo total de energia elétrica de 2015 a 2019 em Minas Gerais.

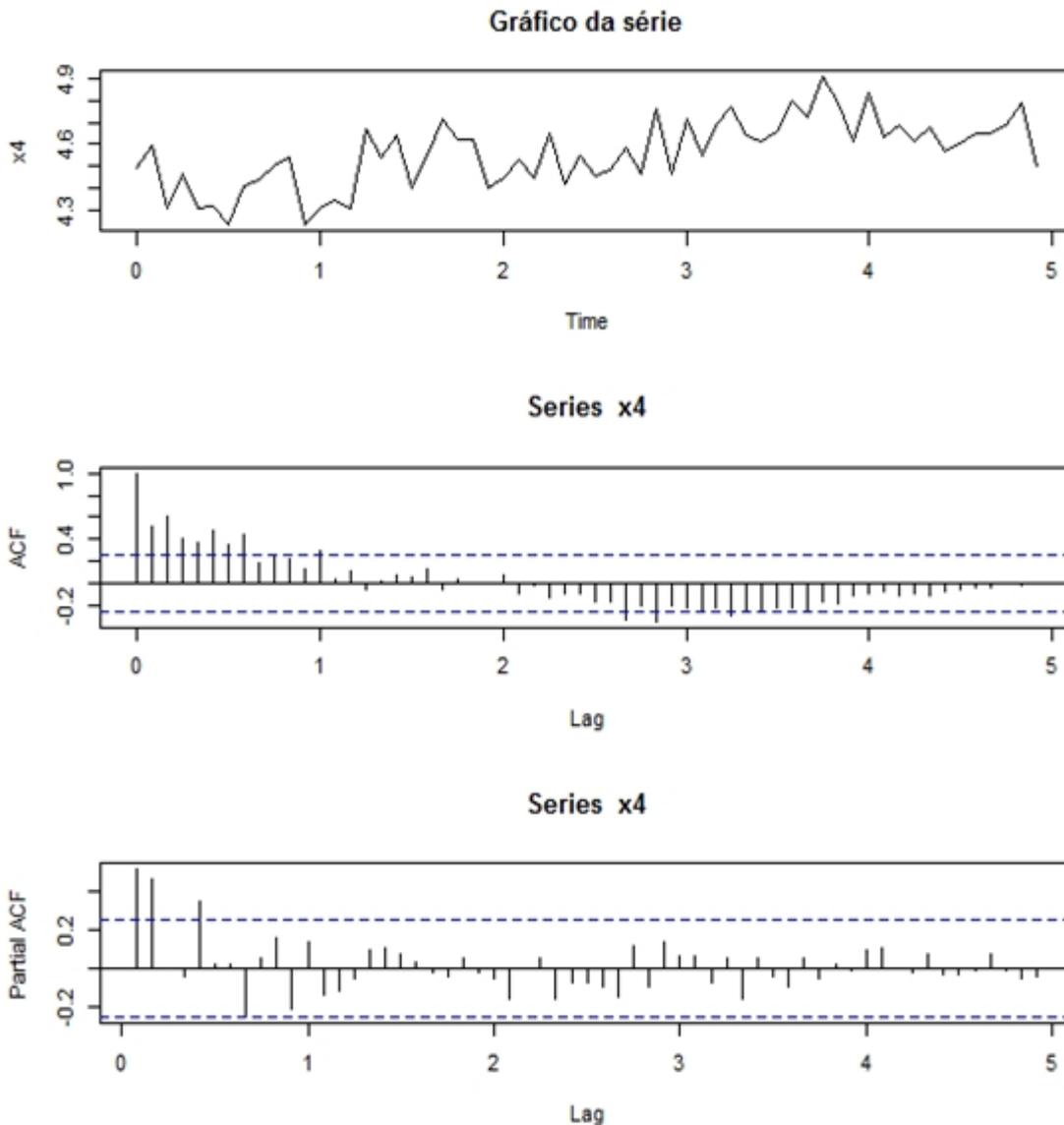


Figura 4.4: Função de autocorrelação FAC e FACp do consumo residencial de energia elétrica de 2015 a 2019 em Minas Gerais.

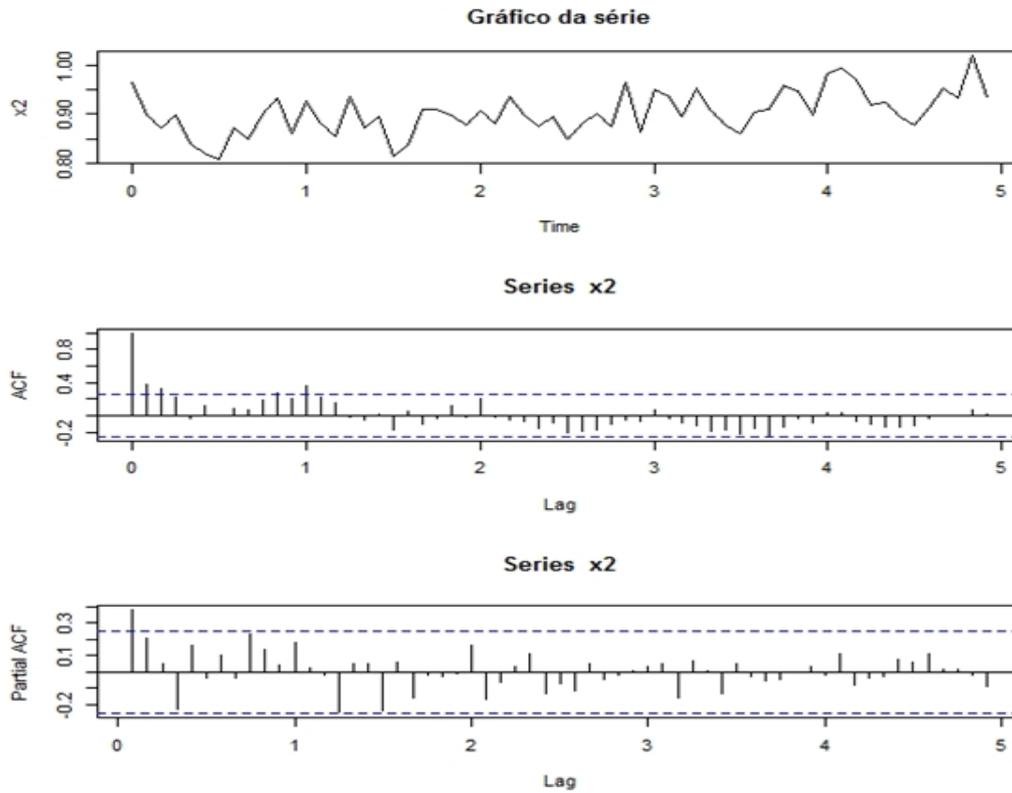


Figura 4.5: Função de autocorrelação FAC e FACp do consumo comercial de energia elétrica de 2015 a 2019 em Minas Gerais.

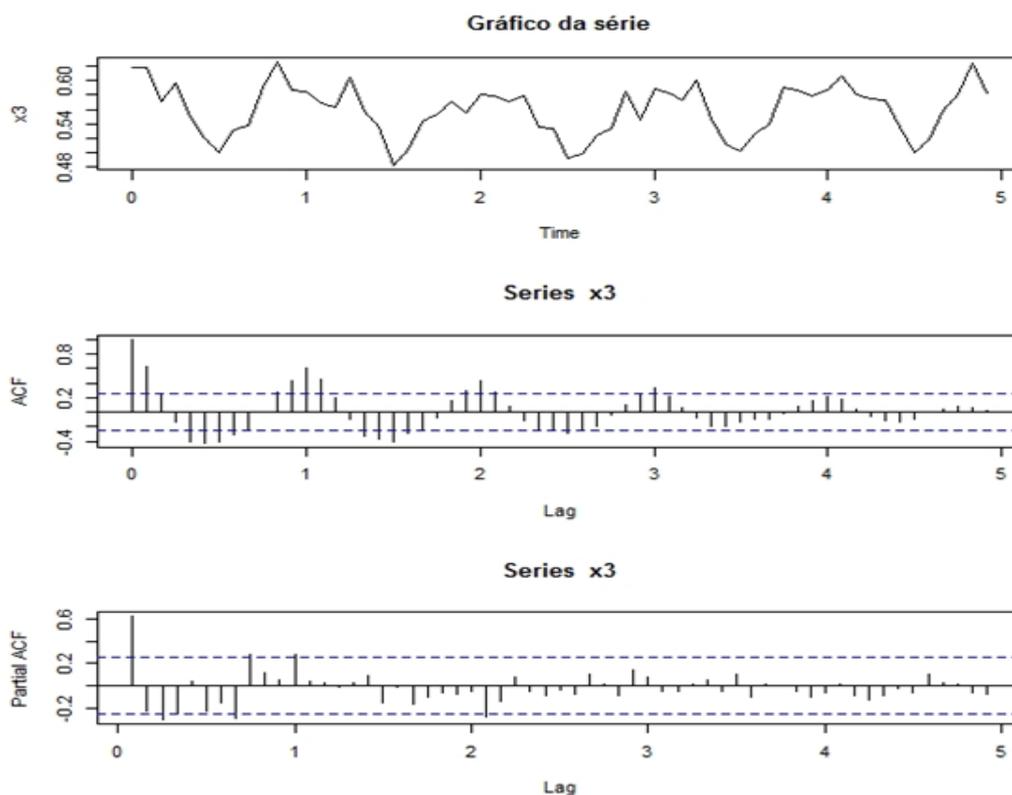
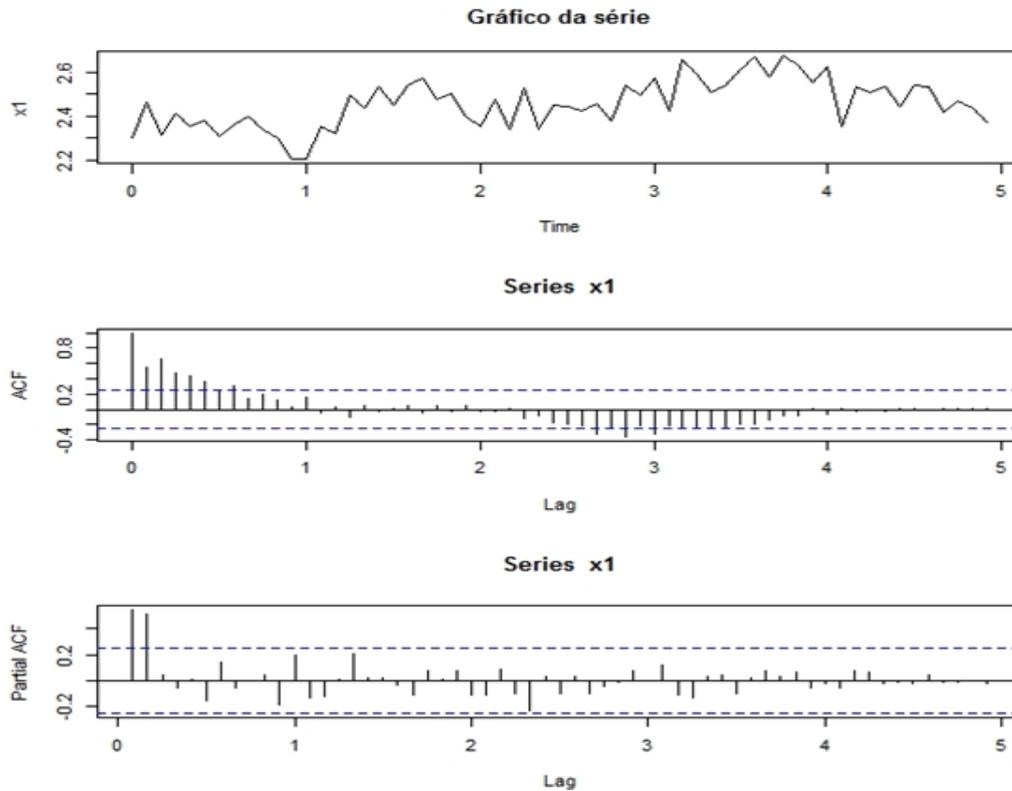


Figura 4.6: Função de autocorrelação FAC e FACp do consumo industrial de energia elétrica de 2015 a 2019 em Minas Gerais,



Para complementar a avaliação da autocorrelação, foi utilizado o teste de Ljung-Box. Conforme os resultados obtidos na Tabela 4.3, a hipótese de não autocorrelação foi rejeitada para todas categorias de consumo, logo o consumo ao longo do tempo são dependentes (autocorrelacionados).

Tabela 4.3: Resultados do Teste de Ljung-Box

Consumo	valor-p	Decisão
Total	0,0000	Rejeita-se $H_0$ , existe autocorrelação
Residencial	0,0027	Rejeita-se $H_0$ , existe autocorrelação
Comercial	0,0000	Rejeita-se $H_0$ , existe autocorrelação
Industrial	0,0000	Rejeita-se $H_0$ , existe autocorrelação

Adicionalmente foi utilizada o teste de Dickey-Fuller Aumentado (ADF) para verificar estacionaridade de uma série um vez que, um série dita estacionária não deve apresentar tendência e uma série independente (não autocorrelacionada) também não deve apresentar tendência. Conforme a Tabela 4.4 com valor crítico de 5% ( $\tau = -3.45$ ), tem-se que o consumo de energia total e industrial não são estacionários e o consumo de energia residencial e comercial são estacionários.

Tabela 4.4: Resultados do Teste de Estacionaridade de Dickey-Fuller Aumentado (ADF)

Consumo	Lag-máx	Estatística T	tau	Decisão
Total	4	-2,26	-3,45	$T > \tau$ , a série não é estacionária
Residencial	3	-5,50	-3,45	$T < \tau$ a série é estacionária
Comercial	5	-4,42	-3,45	$T < \tau$ , a série é estacionária
Industrial	1	-2,10	-3,45	$T > \tau$ a série não é estacionária

Para complementar a avaliação da autocorrelação, foram investigadas a tendência (teste de Cox-Stuart) e a sazonalidade (teste de Fisher) do consumo de energia elétrica.

Na realização do teste de Cox-Stuart, nota-se que o consumo de energia elétrica total, residencial e industrial apresentaram evidências suficientes,  $valor - p < 0,05$ , de que as séries possuem tendência.

Tabela 4.5: Resultados do Teste de tendência de Cox-Stuart

Consumo	p-valor	Decisão
Total	0,0003	Rejeita $H_0$ , apresenta tendência
Residencial	0,0253	Rejeita $H_0$ , apresenta tendência
Comercial	0,6547	Não Rejeita $H_0$ , não apresenta tendência
Industrial	0,0073	Rejeita $H_0$ , apresenta tendência

No teste fisher para detectar sazonalidade, tabela 4.6, onde rejeita-se a hipótese nula se  $valor-p < 0,05$ , os resultados obtidos foram:

Tabela 4.6: Resultados do Teste de Fisher

Consumo	p-valor	Decisão
Total	0,0000	Rejeita-se $H_0$ existe sazonalidade na série
Residencial	0,0092	Rejeita-se $H_0$ existe sazonalidade na série
Comercial	0,0000	Rejeita-se $H_0$ existe sazonalidade na série
Industrial	0,0000	Rejeita-se $H_0$ existe sazonalidade na série

Logo, o consumo de energia elétrica total, residencial, comercial e industrial possui sazonalidade. Foi observado que o teste de Dickey Fuller não foi eficiente ao detectar as tendências vistas no FAC e FACp, no teste de Cox Stuart e também no teste de Fisher nos consumos residencial e comercial.

## RESULTADOS APÓS AJUSTE DO MODELO AUTOREGRESSIVO

Quando os dados do monitoramento são autocorrelacionados, uma opção é construir o gráfico de controle para os resíduos do modelo de séries temporais [14]. Os modelos autoregressivos, AR(p), modelos de médias móveis (moving average), MA(p), modelos autoregressivos integrados de médias móveis, ARIMA(p, d, q) e outros podem ser usados. Desde

que ao modelar tenhamos um modelo cujo os resíduos sejam independentes (não autocorrelacionados). Neste trabalho, ajustou-se o modelo auto-regressivo de ordem 1, AR(1), para o consumo total, residencial e comercial e o modelo auto-regressivo de ordem 2, AR(2), para o consumo de energia industrial.

Nas Figuras 4.7, 4.8, 4.9 e 4.10 será apresentado a avaliação da autocorrelação dos resíduos do modelo ajustado.

Figura 4.7: FAC e FACp dos resíduos do modelo AR(1) no consumo total de energia elétrica

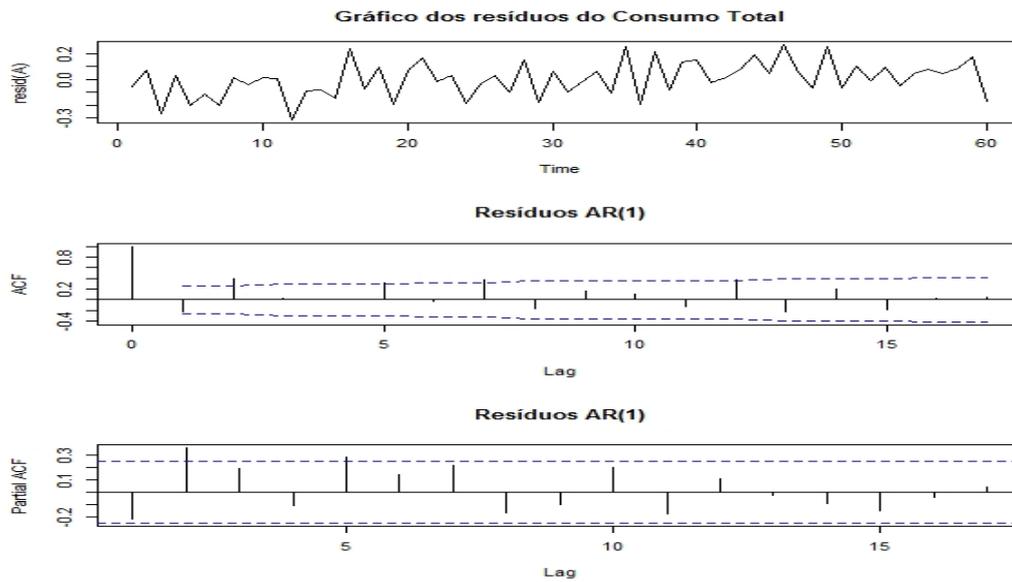


Figura 4.8: FAC e FACp dos resíduos do modelo AR(1) no consumo residencial de energia elétrica

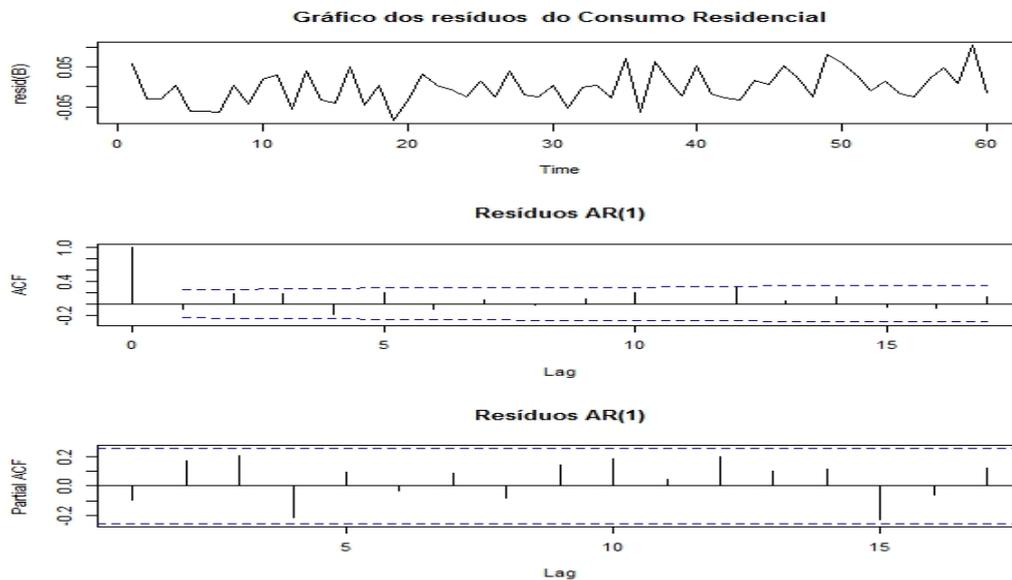


Figura 4.9: FAC e FACp dos resíduos do modelo AR(1) no consumo comercial de energia elétrica

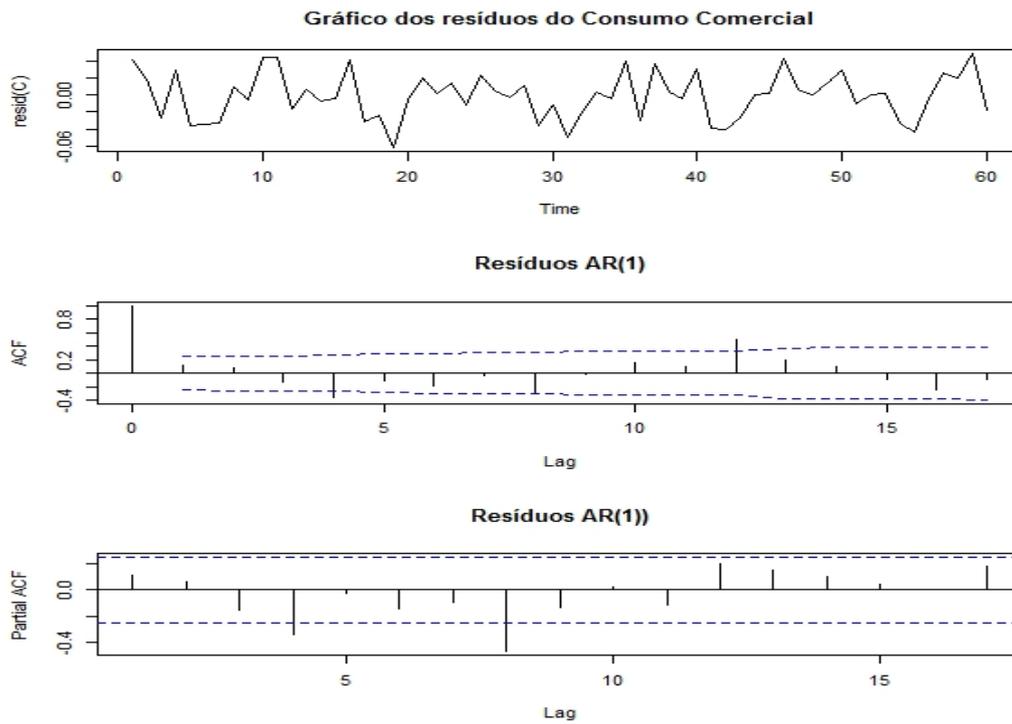
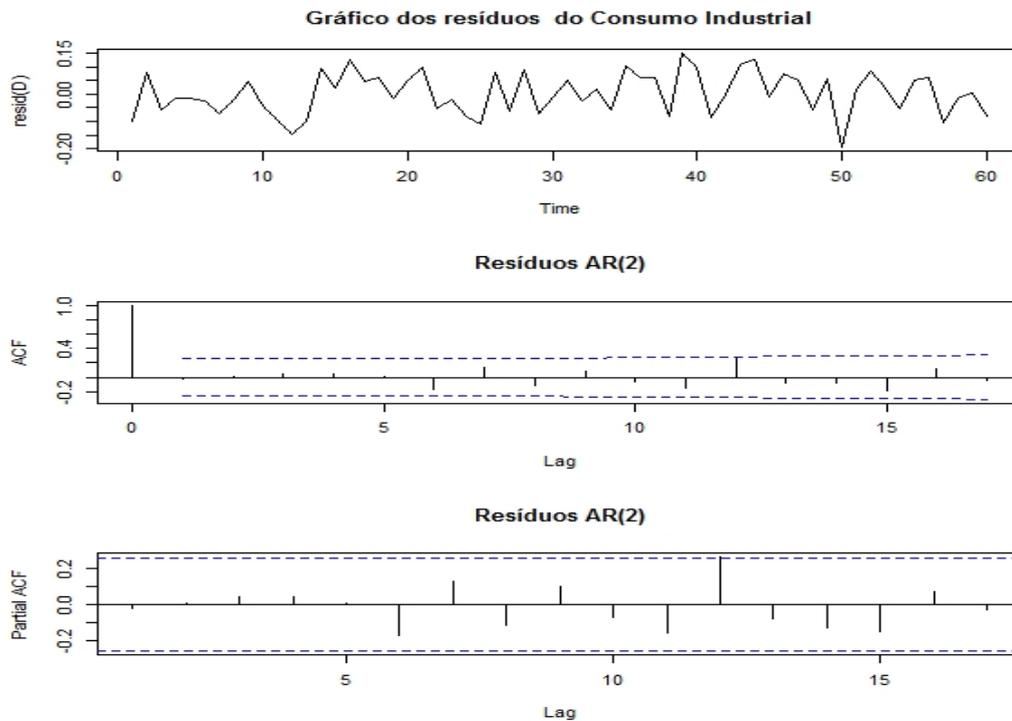


Figura 4.10: FAC e FACp dos resíduos do modelo AR(2) consumo industrial de energia elétrica



No teste de Ljung-Box após o ajuste AR(1) e AR(2), confirmou que os resíduos do modelos são não autocorrelacionados, então pode-se dizer que o modelo é adequado, talvez não o modelo que melhor ajusta os dados, mas para a necessidade do trabalho (trabalhar com dados não autocorrelacionados) o modelo se mostrou eficiente.

Tabela 4.7: Resultados do Teste de Ljung-Box dos resíduos do modelo

<b>Consumo</b>	<b>valor-p</b>	<b>Hipótese</b>
Total	0,0773	Os resíduos do modelo AR(1) não são autocorrelacionados
Residencial	0,4625	Os resíduos do modelo AR(1) não são autocorrelacionados
Comercial	0,3566	Os resíduos do modelo AR(1) não são autocorrelacionados
Industrial	0,0178	Os resíduos do modelo AR(1) são autocorrelacionados
Industrial	0,8656	Os resíduos do modelo AR(2) não são autocorrelacionados

Além disso, de acordo com teste de Shapiro-Wilk, os resíduos do modelo AR(1) do consumo de energia total ( $valor - p = 0,797$ ), residencial ( $valor - p = 0,645$ ) e comercial ( $valor - p = 0,201$ ) possuem distribuição normal. Os resíduos do modelo AR(2) do consumo de energia industrial ( $valor - p = 0,378$ ) possuem distribuição normal.

Portanto atingindo os requisitos para eficiência dos gráficos, dados normalmente distribuídos e independentes e igualmente distribuídos (sem autocorrelação), o próximo passo é a construção dos gráficos.

### 4.1.3 GRÁFICOS DE CONTROLE, X E AM, FASE 1

Das análises acima, conclui-se que o consumo de energia elétrica é autocorrelacionado, apresentou tendência e sazonalidade. Esse comportamento prejudica o desempenho dos gráficos de controle, ou seja, aumenta a frequência de alarmes falsos. No entanto, a abordagem de modelos de séries temporais tem se mostrado útil nas situações em que é necessária a modelagem da autocorrelação dos dados. Neste caso, o gráfico de controle é aplicado aos resíduos do modelo de séries temporais. Geralmente, os modelos autorregressivos de primeira e segunda ordem são suficientes para eliminar a autocorrelação dos dados.

Neste trabalho, foi ajustado o modelo AR(1) para o consumo de energia elétrica total, residencial e comercial e o modelo AR(2) para o consumo de energia elétrica industrial. Veja os gráficos de controle para os resíduos de AR(1) e AR(2) a seguir nas Figuras 4.11 a 4.14.

Os gráficos de controle de Shewhart gerados pelo software R utilizam automaticamente as regras suplementares, as quais são úteis para acelerar detecção de alterações no processo [14]. Essas regras são os pontos em amarelo nos gráficos de controle. No entanto, o uso dessas regras levam ao aumento de alarmes falsos. Por isso, essas regras não serão adotadas neste aqui.

Na Figura 4.12(a), os gráficos de X e AM, para dados originais do consumo total apresentam pontos fora dos limites, indicando que o processo está fora de controle. No entanto, essa análise não é confiável, pois os dados estão autocorrelacionados e isso pode aumentar a ocorrência de alarmes falsos. Na Figura 4.12(b), os gráficos de X e AM, para os resíduos do modelo AR(1), apresentam pontos dentro dos limites, indicando que o processo está em controle. Portanto, conclui-se que este processo está em controle.

Logo, para o monitoramento em tempo real, fase 2, na Figura 4.15, consumo total, serão utilizados os resíduos de AR(1) e os limites serão:

Limites	Gráficos de X	Gráficos de AM
LIC	-0,4660	0
LM	0,0007	0,1754
LSC	0,4674	0,5733

Na figura 4.16, o consumo residencial, AR(1), os limites serão:

Limites	Gráficos de X	Gráficos de AM
LIC	-0,1319	0
LM	-0,0004	0,0494
LSC	0,1310	0,1615

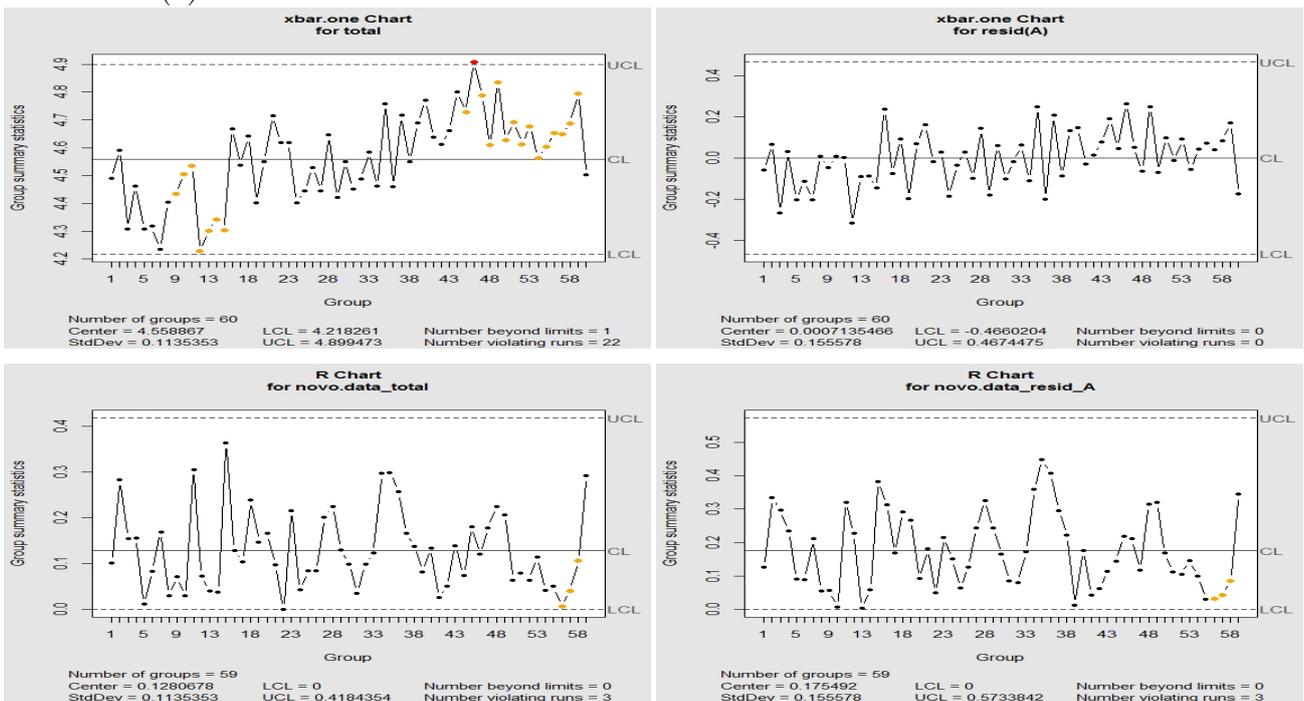
Na figura 4.17, o consumo comercial, AR(1), os limites serão:

Limites	Gráficos de X	Gráficos de AM
LIC	-0,0778	0
LM	-0,0008	0,0289
LSC	0,0762	0,1615

Na figura 4.18, o consumo industrial, AR(2), os limites serão:

Limites	Gráficos de X	Gráficos de AM
LIC	-0,2379	0
LM	-0,0027	0,0905
LSC	0,2435	0,2957

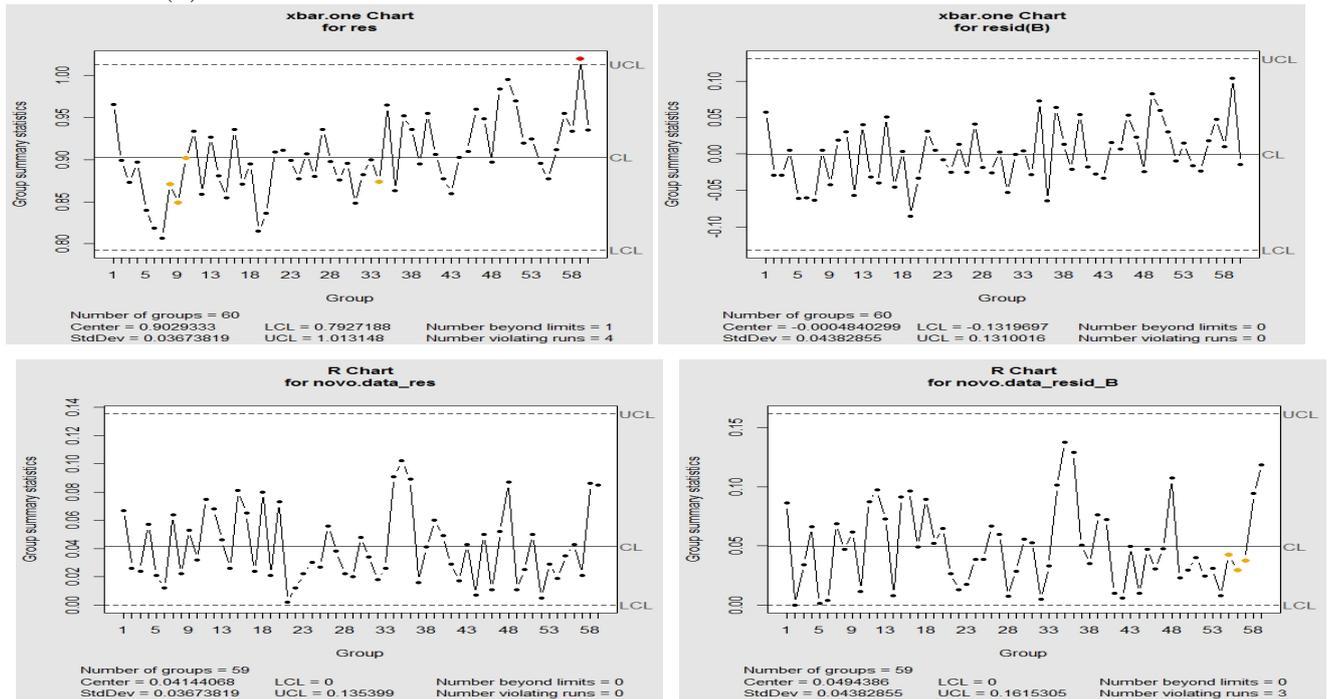
Figura 4.11: Gráficos de controle X e AM consumo total de energia elétrica e dos resíduos do modelo AR(1)



(a) Gráficos de X e AM do Consumo Total

(b) Gráficos de X e AM dos resíduos do modelo AR(1) Consumo Total

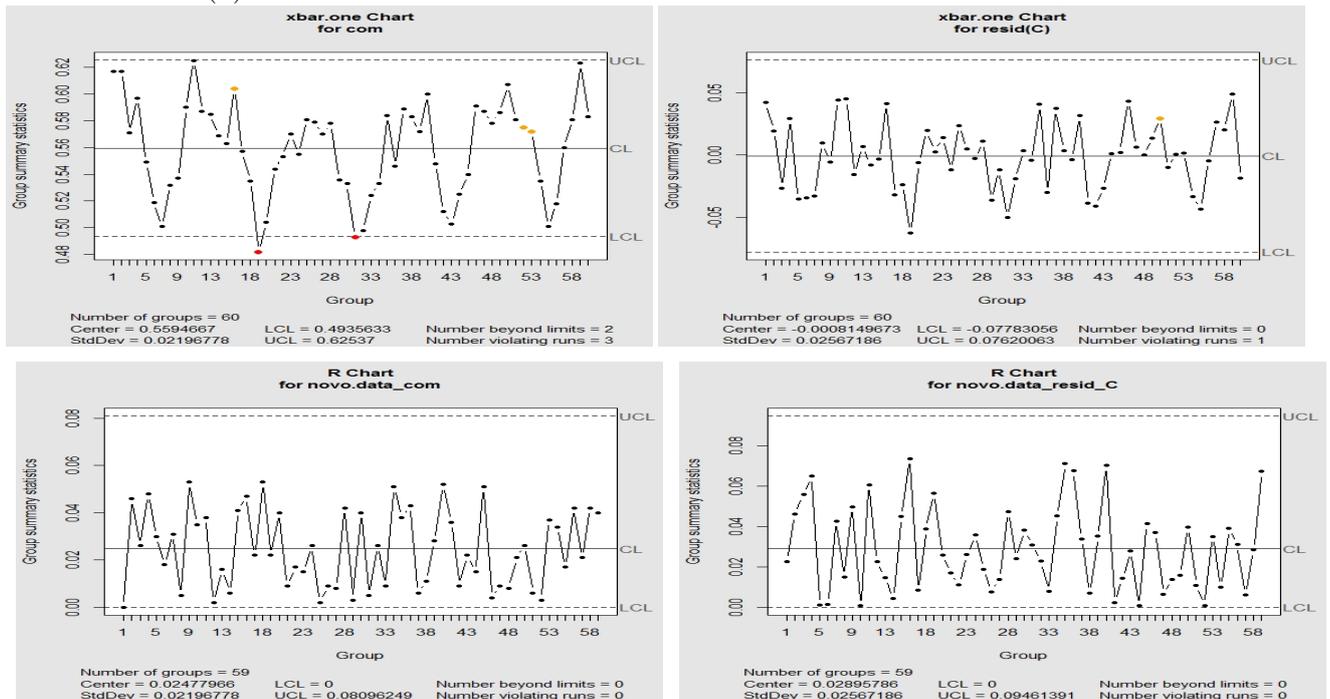
Figura 4.12: Gráficos de controle X e AM consumo residencial de energia elétrica e resíduos do modelo AR(1)



(a) Gráficos de X e AM Consumo Residencial

(b) Gráficos de X e AM dos resíduos do modelo AR(1) Consumo Residencial

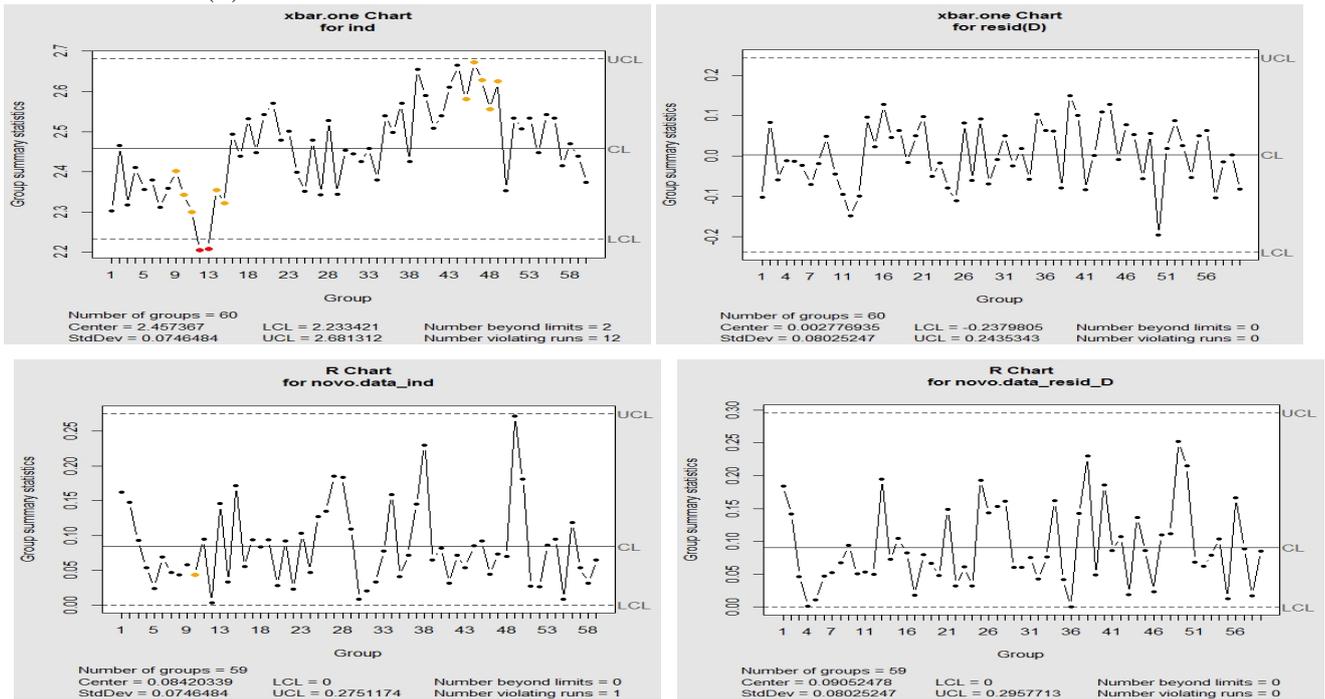
Figura 4.13: Gráficos de controle X e AM do consumo comercial de energia elétrica e resíduos do modelo AR(1)



(a) Gráficos X e AM Consumo Comercial

(b) Gráficos X e AM dos resíduos do modelo AR(1) Consumo Comercial

Figura 4.14: Gráficos de controle X e AM do consumo industrial de energia elétrica e resíduos do modelo AR(2)



(a) Gráficos de X e AM Consumo Industrial

(b) Gráficos de X e AM dos resíduos do modelo AR(2) Consumo Industrial

### 4.1.4 FASE 2 - MONITORAMENTO

#### DADOS

Os dados assim como na fase 1, são do consumo de energia elétrica mensal residencial, comercial, industrial e total porém na fase 2 de monitoramento será analisado do período de janeiro de 2020 a julho de 2021 na escala de Megawatt-hora (MWh), a escala também será reduzida para manter a lógica do estudo da seguinte forma:

$$\frac{1(MWh)}{1000000} \tag{4.1}$$

Os dados utilizados no monitoramento em tempo real (fase 2) estão na tabela B.3 (apêndice B).

## GRÁFICOS DE CONTROLE - FASE 2 - MONITORAMENTO

A seguir será exibido os gráficos de controle durante a fase de monitoramento, onde as observações do 1 até a 60 (linha pontilhada) pertencem a fase 1 do gráfico e da observação 60 até a 78 pertencem a fase 2 (validação).

Figura 4.15: Gráfico de controle X e AM para os resíduos do consumo total de energia elétrica

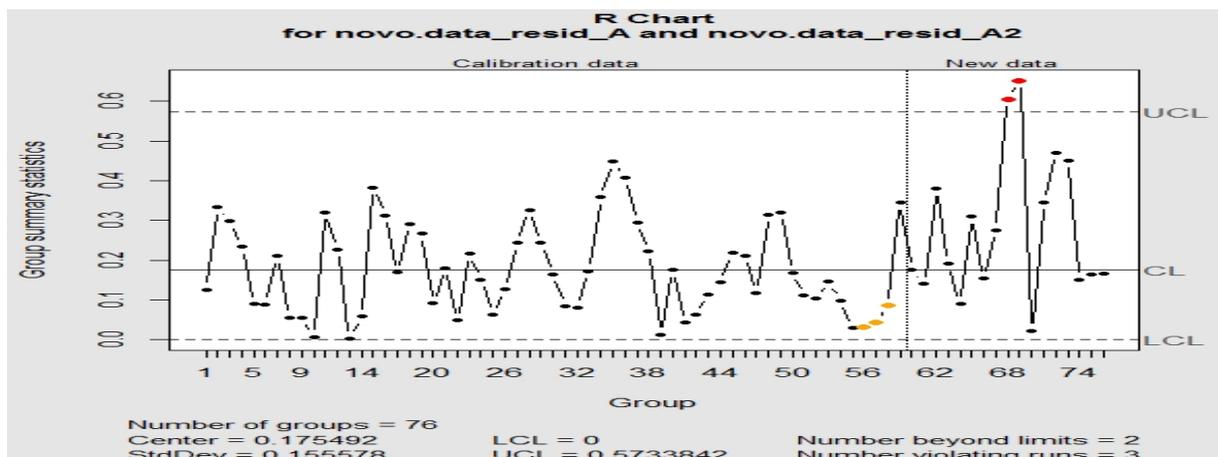
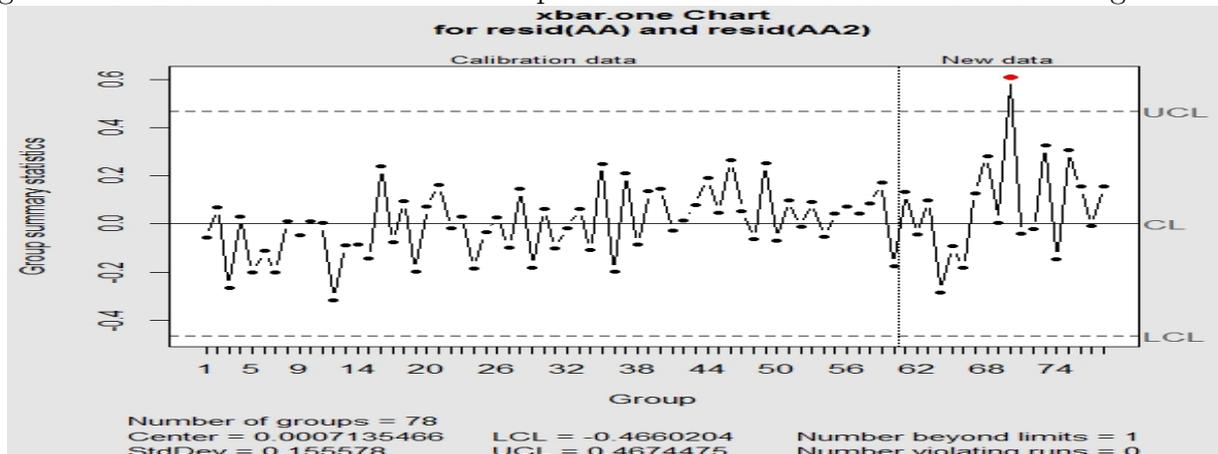


Figura 4.16: Gráfico de controle X e AM para os resíduos do consumo residencial de energia elétrica

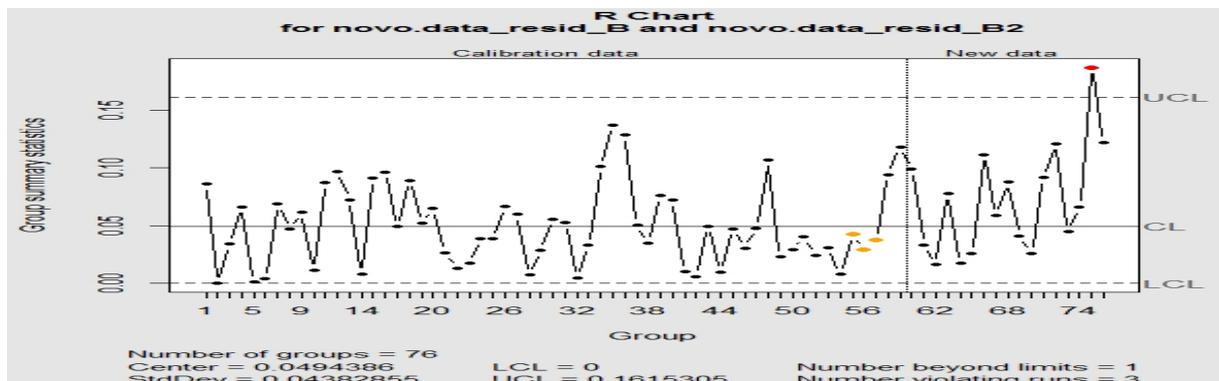
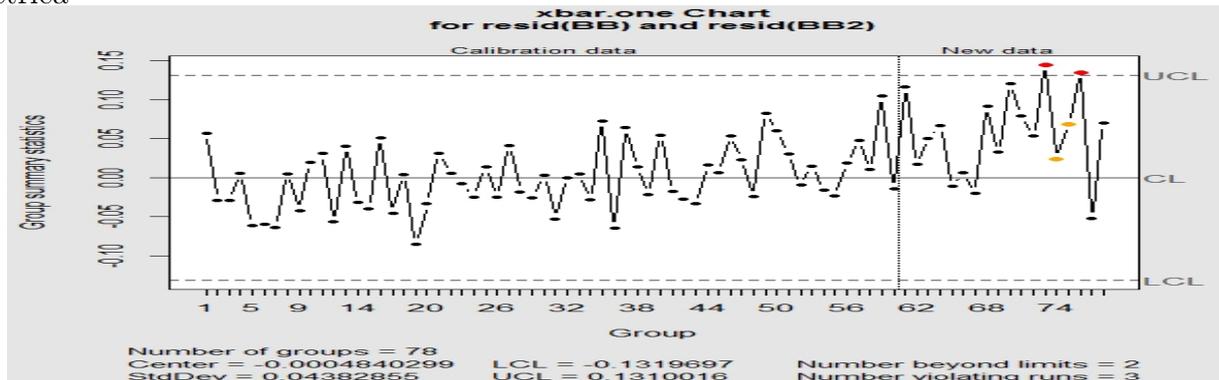


Figura 4.17: Gráfico de controle X e AM para os resíduos do consumo comercial de energia elétrica

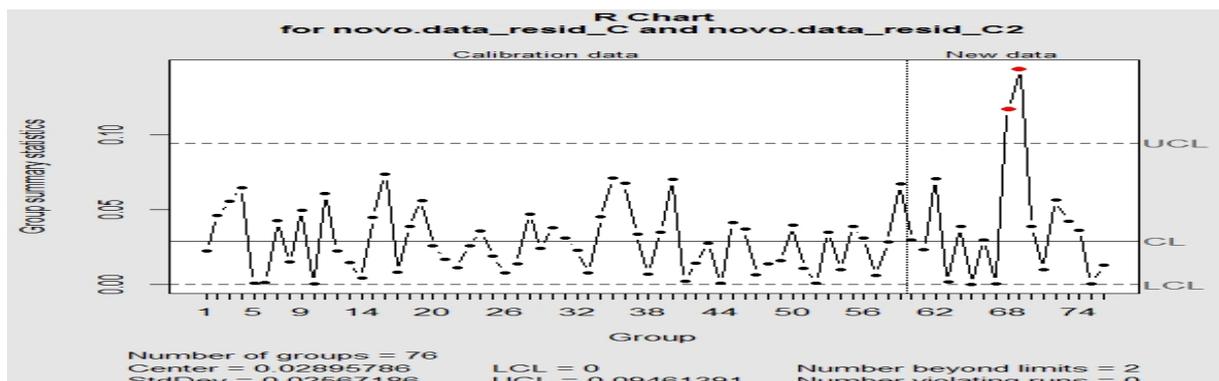
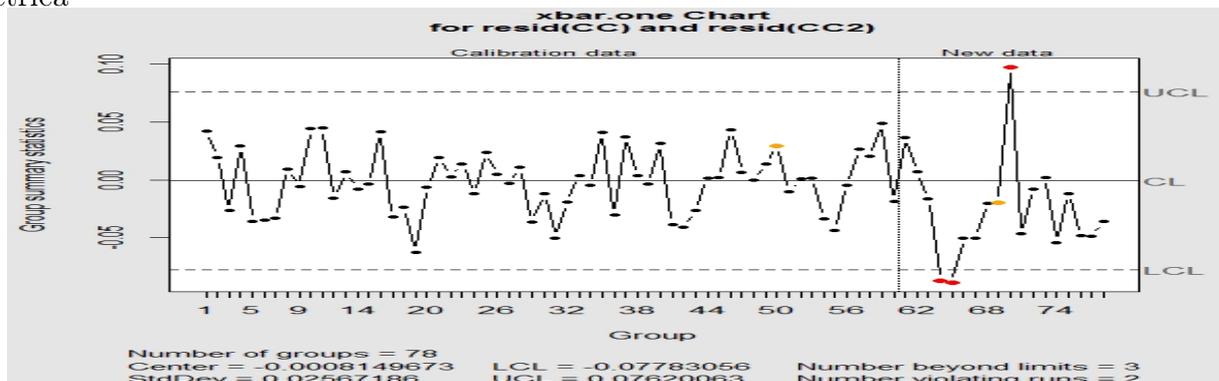
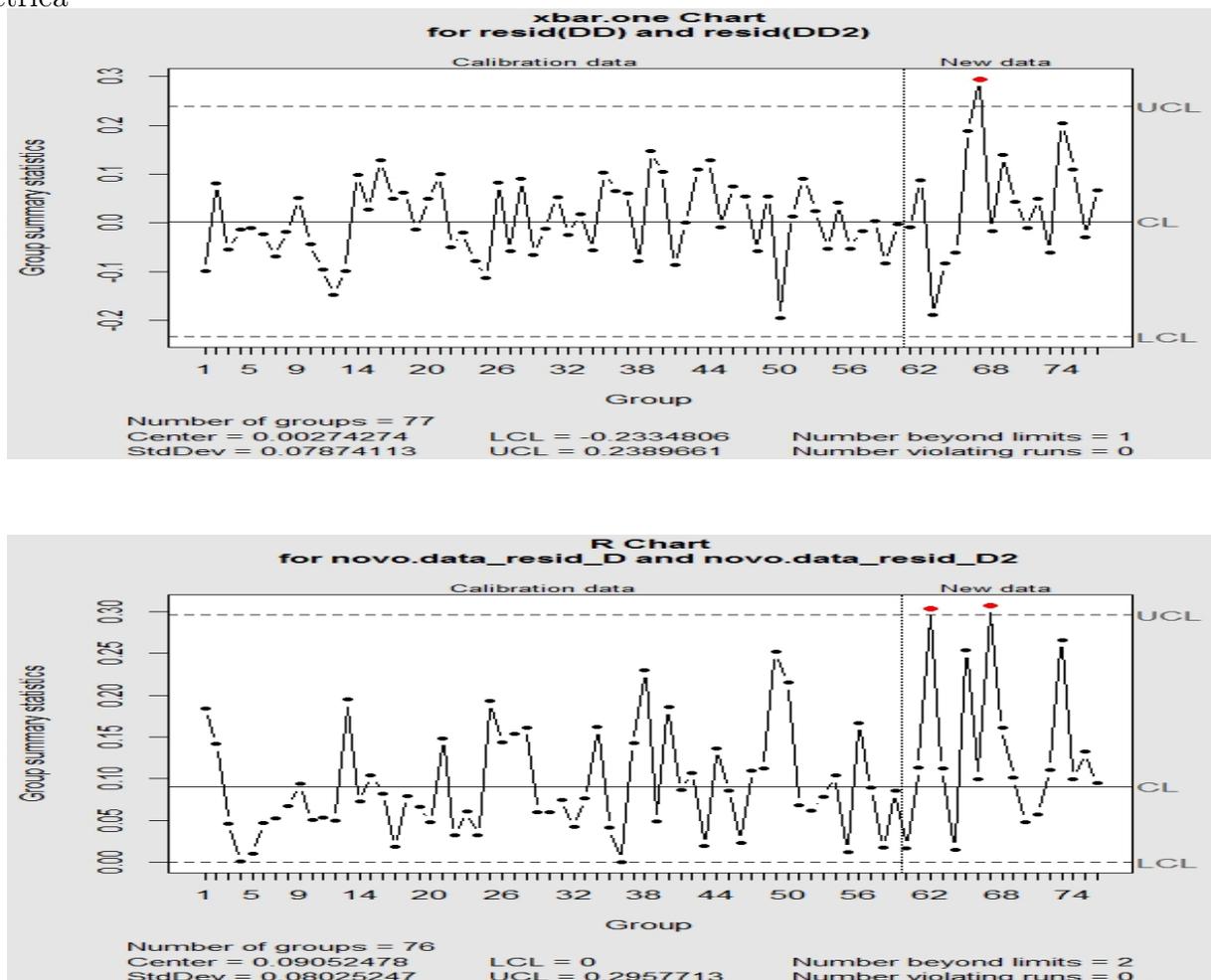


Figura 4.18: Gráfico de controle X e AM para os resíduos do consumo industrial de energia elétrica



Na figura 4.15, o gráfico X dos resíduos do consumo total de energia elétrica emitiu sinal de alerta no mês de outubro de 2020 e o gráfico de AM nos meses de outubro e novembro de 2020.

Na figura 4.16 o gráficos X dos resíduos do consumo de energia elétrica residencial emitiu sinal de alerta nos meses de janeiro e abril de 2021 e o gráfico de AM em maio de 2021.

Na figura 4.17 o gráficos X dos resíduos do consumo de energia elétrica comercial emitiu sinal de alerta nos meses de abril, maio, outubro de 2020 e o gráfico de AM em outubro e novembro de 2020.

Na figura 4.18 o gráficos X dos resíduos do consumo de energia elétrica industrial emitiu sinal de alerta nos meses de agosto de 2020 e o gráfico de AM em abril e setembro de 2020.

Na prática, um sinal de alerta no monitoramento em tempo real, orienta para a adoção de medidas que impeçam a continuidade do aumento do consumo de energia elétrica. Por exemplo campanhas para o uso consciente de energia elétrica, implantação de auditorias, aplicação de multas se identificado negligências do consumidor e outros.

## 5. CONCLUSÕES

A utilização de gráficos de controle compreende duas fases. Neste trabalho, na fase 1, foram utilizados os dados históricos, janeiro/2015 a dezembro/2019. Da avaliação inicial, na fase 1, obteve-se que o consumo de energia elétrica possui distribuição normal conforme apontado pelo teste Shapiro-wilk, e é autocorrelacionado, conforme foi indicado pelos gráficos da função de autocorrelação (FAC e FACp) e pelo teste Ljung-Box. Adicionalmente, o teste Dickey-Fuller aumentado verificou que a série não é estacionária para o consumo total e industrial, já as séries de consumo residencial e comercial são estacionárias. Por outro lado, para complementar a avaliação da estacionariedade, o teste Cox-Stuart indicou a presença de tendência para o consumo total, residencial e industrial. E o teste de Fisher indicou a presença de sazonalidade para o consumo total, residencial, comercial e industrial.

Ainda, na fase 1, os gráficos de controle de X e AM para os dados originais do consumo de energia elétrica apresentaram pontos fora dos limites indicando o processo estava fora de controle. Esse resultado não é confiável, pois os dados originais são autocorrelacionados e leva à ocorrência de alarmes falsos. Então, ajustou-se o modelo auto-regressivo, AR(1) para consumo de energia total, residencial e comercial, e o AR(2) para o consumo de energia industrial. Com os gráficos de controle de X e AM para os resíduos dos modelos auto-regressivos não houve pontos fora dos limites, portanto, o processo estava em controle. Logo, esses limites podem ser utilizados na fase 2 para o monitoramento do processo.

Na fase 2 de monitoramento, no período janeiro 2020 a junho/2021, todos os gráficos de controle para o resíduo do modelo auto-regressivo do consumo de energia elétrica emitiram sinal de alerta, sinalizaram o processo fora de controle. Ou seja, pontos acima do limite superior indicam que o aumento do consumo é expressivo, maior do que o estatisticamente aceitável. Esses sinais, quando são alarmes verdadeiros, indicam que causas especiais estão atuando no processo, nesse momento da sinalização o ideal é que ações sejam tomadas para que se evite, de forma preventiva, que o consumo de energia elétrica continue aumentando.

Os alarmes falsos são indesejáveis, pois os impactos de alarme falsos podem ser prejudiciais ao processo de produção e fornecimento de energia elétrica. Se o alarme falso, for referente ao ponto acima do limite superior, indicando que o consumo de energia elétrica está alto, então pode ser que ocorram ações desnecessárias, como por exemplo o aumento investimentos para aumentar a produção ou a taxaço da energia para os consumidores e aumento do preço.

Se o alarme falso for pontos abaixo o limite inferior, indicando que o consumo de energia elétrica está baixo, então pode ser que ocorram ações equivocadas, como a interrupção de

investimentos no setor para aumentar a produção.

Lembrando que as possíveis causas de um processo fora de controle devem ser analisadas cautelosamente por um especialista no assunto (energia elétrica), e existem algumas ferramentas na qualidade que podem auxiliar na identificação dessas possíveis causas como o diagrama de Pareto e o diagrama de Ishikawa (diagrama de causa e efeito). Os gráficos de controle (a ferramenta estudada) apenas sinalizam se o processo está em controle ou fora de controle.

Aqui, foram utilizados os modelos auto-regressivos de ordem 1 e ordem 2, mesmo que eles não sejam o modelo que melhor se ajusta aos dados, os resíduos deles foram suficientes para modelar a autocorrelação do processo.

Portanto, o uso de gráficos de controle X e AM para os resíduos dos modelos auto-regressivos é um método adequado para o monitoramento de energia elétrica com dados autocorrelacionados.

### **Trabalhos futuros**

Em continuidade ao estudo apresentado neste trabalho, propõe-se:

1. Recalcular os limites dos gráficos de controle, para monitoramento em tempo real, de tempos em tempos, utilizando os dados atualizados do processo.
2. Propor outros modelos de séries temporais para eliminar a autocorrelação dos dados do processo, por exemplo, o ARMA e o ARIMA.
3. Aplicar outros tipos de gráficos, por exemplo o gráficos de somas acumuladas (CUSUM) e gráfico da média móvel ponderada exponencialmente (EWMA)
4. Utilizar outras ferramentas da CEQ para identificar possíveis causas do processo fora de controle.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AMARAL, M. V. S. G.: *Ajuste de modelos e comparação de séries temporais para dados de vazão específica em microbacias pareadas*. Tese de Mestrado, 2014. Escola Superior de Agricultura Luiz de Queiroz.
- [2] BORGATTO, A. F.: *Análise de intervenção em séries temporais: aplicações em transporte urbano*. Tese de Mestrado, 2000. Universidade Federal de Lavras.
- [3] CEMIG: *Companhia Energética de Minas Gerais S.A.*, 2021. <https://www.cemig.com.br/>, maio 2021.
- [4] CROSBY, P. B.: *Quality without tears: the art of hassle free management*. Mc Graw-Hill, New York, 1995.
- [5] DEMING, W. E.: *Out of crisis*. Cambridge: MIT Press, Massachusetts, 2000.
- [6] EHLERS, R. S.: *Análise de Séries Temporais*. Departamento de Estatística, 2007. <http://leg.est.ufpr.br/~ehlers/notas>, acesso: maio de 2021.
- [7] EPE: *Empresa de Pesquisa Energética*, 2021. <https://www.epe.gov.br/>, maio 2021.
- [8] FISCHER, S.: *Séries Univariantes de Tempo- Metodologia de Box Jenkins*. Porto Alegre, 1982.
- [9] IDE, M. S., RIBEIRO JUNIOR, J. I. e SOUZA, L. M.: *Gráficos de Controle com uma abordagem de séries temporais*. Revista eletrônica Produção e Engenharia, 2:69–84, jan/jul 2009. <https://periodicos.ufjf.br/index.php/producaoengenharia/article/view/28862/19739>, acesso: maio de 2021.
- [10] JURAN, J. M. e GODFREY, A. B.: *Juran's quality handbook*. Mc Graw-Hill, New York, 1999.
- [11] LIMA, M. W. P. e CHARNET, R.: *Modelos de Previsão Aplicados no Controle Estatístico do Processo na Presença de dados Autocorrelacionados*, 2008. <https://www.ime.unicamp.br/sites/default/files/pesquisa/relatorios/rp-2008-2.pdf>, acesso: maio de 2021.
- [12] MAIA, A. G.: *Econometria: conceitos e aplicações*. 2017. Cap. 13.

- [13] MATTOS, R. S.: *Tendências e Raízes Unitárias*. Departamento de Economia-UFJF, maio, 2018.
- [14] MONTGOMERY, D. C.: *Introduction to statistical quality control*. New York, 7ª ed., 2013.
- [15] MORETTIN, P. A. e BUSSAB, W. O.: *Estatística Básica*. São Paulo, 8ª ed., 2013.
- [16] MORETTIN, P. A. e TOLOI, C. M. C.: *Análise de Séries Temporais*. ABE- Projeto Fisher e Editora Blucher, 2ª ed., 2006.
- [17] RATNIEKS, I.: *Métodos de Previsão: Aplicação da Metodologia de Box e Jenkins ao Varejo Brasileiro - o Caso das Lojas Americanas S.A.* Trabalho de Conclusão de Curso, 2010. Universidade Federal do Rio Grande do Sul.
- [18] REIS, E. A. e REIS, I. A.: *Análise Descritiva de Dados*. Relatório Técnico do Departamento de Estatística da UFMG, 2002. [www.est.ufmg.br](http://www.est.ufmg.br).
- [19] REIS, L. B.: *GERAÇÃO DE ENERGIA ELÉTRICA*. Barueri-SP, 2ª ed., 2015.
- [20] ROSA, L. M. F.: *Probabilidade do alarme falso no gráfico de controle EWMA para o monitoramento de processo autocorrelacionado*. Magister Scientiae, 2012. Universidade Federal de Viçosa-MG.
- [21] SAMOHYL, R. W.: *Controle Estatístico da Qualidade*. Rio de Janeiro, 2009.
- [22] SHEWHART, W.: *Statistical method: from the viewpoint of quality control*. Washington: Dover, 1986.
- [23] SPIEGEL, M. R. e STEPHENS, L. J.: *Schaum's outline of theory and problems of statistics*. New York: McGraw-Hill, 4th ed., 2008.
- [24] TACONELI, C. e ZEVIANI, W.: *Gráficos de Controle*, 2019. <http://leg.ufpr.br/~walmes/ensino/CEQ/slides/03-grafico-de-controle.pdf>, Laboratório de Estatística e Geoinformação UFPR. Acesso: maio 2020.
- [25] TAGUCHI, G., TAGUCHI, S. e CHOWDHURY, S.: *Robust engineering*. Mc Graw-Hill, New York, 1999.
- [26] YNOGUTI, C. A.: *Probabilidade, Estatística e Processos Estocásticos*, 2011.

# A. APÊNDICE

## A.0.1 SOFTWARE R

**#Análise descritiva**

```
summary(dados)
```

**#boxplot**

```
boxplot(dados)
```

**#teste Shapiro-Wilk**

```
shapiro.test(dados)
```

**#Gráficos QQplot**

```
qqnorm(dados)
```

```
qqline(dados)
```

```
ts.plot(dados)
```

```
g=acf(dados, lag.max=1000)
```

**#Teste da raiz unitária de Dickey-Fuller (ADF)**

```
library(urca)
```

```
library(forecast)
```

```
x0<-ur.df(dados,type=c("none","drift","trend"),lags=1,selectlags=c("Fixed","AIC","BIC"))
```

```
plot(x0)
```

```
summary(x0)
```

**##Obs.:**

"dados" refere-se aos dados;

"type" especifica o tipo de teste: "none" sem intercepto e sem tendência determinística, "drift" considera apenas o intercepto e "trend" engloba as outras e considera o intercepto e a tendência determinística;

"Lag" é a quantidade máxima de parâmetros que um modelo poderá possuir.

"selectlags" para especificar o critério que será utilizado.

# "Value of test-statistic is:" indica o valor da estatística T, e "tau1" representa os valores

críticos  $\tau$ . Se  $T < \tau$  a série é estacionária.

### **#Ljung-Box antes do ajuste**

```
Box.test(dados, lag = 1, type = c ("Ljung-Box"), fitdf = 0)
```

### **#Teste Cox-Stuart**

```
library(trend)
cs.test(dados)
```

### **# teste de Fisher**

```
library(GeneCycle)
fisher.g.test(dados)
```

### **# Modelo AR(1)**

```
A<-arima(dados, order=c(1,0,0))
```

### **# Modelo AR(2)**

```
A<-arima(dados, order=c(2,0,0))
```

### **# FAC e FACp após ajuste AR(p)**

```
par(mfrow=c(3,1))
plot(resid(A), main=c("Gráfico dos resíduos"))
acf(resid(A), main='Resíduos AR(1)', ci.type='ma')
pacf(resid(A), main='Resíduos AR(1)')
```

### **# Teste Ljung Box depois do ajuste**

```
Box.test(resid(A), lag = 1, type = c ("Ljung-Box"), fitdf = 0)
```

### **#gráficos de controle X e AM**

#### **#fase 1**

#### **# sem ajuste**

```
library(qcc)
```

#### **#X**

```
X1 = qcc(dados, type="xbar.one")
```

#### **#AM**

```
novo.data = matrix(cbind(dados[1:length(dados)-1], dados[2:length(dados)]), ncol=2)
```

```
plot.amp.movel = qcc(novo.data, type="R", plot = TRUE)
```

```
# Após ajuste #X
```

```
qcc(resid(A), type="xbar.one")
```

```
#AM novo.dataresidA = matrix(cbind(resid(A)[1 : length(resid(A)) - 1], resid(A)[2 :  
length(resid(A))]), ncol = 2)plot.amp.movelresidA = qcc(novo.dataresidA, type = "R", plot =  
TRUE)
```

```
# fase 2
```

```
#Gráfico X
```

```
AA<- Arima(window(dados1,end=60),order=c(1,0,0)) #fase 1(1a60) (AR(1))
```

```
AA<- Arima(window(dados1,end=60),order=c(2,0,0)) #fase 1(1a60) (AR(2))
```

```
resid(AA)
```

```
AA2 <- Arima(window(dados1,start=61),model=AA)fase 2(61a78), com modelo da fase1
```

```
graf<-qcc(resid(AA), type="xbar.one")
```

```
graf<-qcc(resid(AA), type="xbar.one", newdata=resid(AA2))
```

```
#Gráfico AM
```

```
novo.dataresidA = matrix(cbind(resid(AA)[1 : length(resid(AA)) - 1], resid(AA)[2 :  
length(resid(AA))]), ncol = 2)
```

```
novo.dataresidA2 = matrix(cbind(resid(AA2)[1 : length(resid(AA2)) - 1], resid(AA2)[2 :  
length(resid(AA2))]), ncol = 2)
```

```
plot.amp.movelresidA = qcc(novo.dataresidA, type = "R", plot = TRUE)
```

```
plot.amp.movelresidAf2 = qcc(novo.dataresidA, type = "R", newdata = novo.dataresidA2)
```



## B. TABELAS E DADOS

### B.0.1 TABELA PARA OS GRÁFICOS DE CONTROLE

n	$A_2$	$D_3$	$D_4$	$d_2$
2	1,880	0,000	3,267	1,128
3	1,023	0,000	2,574	1,163
4	0,729	0,000	2,282	2,059
5	0,577	0,000	2,114	2,326
6	0,483	0,000	2,004	2,534
7	0,419	0,076	1,924	2,704
8	0,373	0,136	1,864	2,847
9	0,337	0,184	1,816	2,970
10	0,308	0,223	1,777	3,078

Tabela B.1: Fonte Montgomery [14]

### B.0.2 DADOS

#### FASE 1 (MONITORAMENTO)

Tabela B.2: Consumo de energia elétrica, em milhões de MWh, de jan/2015 a dez/2019 em Minas Gerais.

mês/ano	Total	Residencial	Industrial	Comercial	outros
jan/2015	4,49	0,96	2,30	0,62	0,60
fev/2015	4,59	0,90	2,46	0,62	0,61
mar/2015	4,31	0,87	2,32	0,57	0,55
abr/2015	4,46	0,90	2,41	0,60	0,56
mai/2015	4,31	0,84	2,36	0,55	0,56
jun/2015	4,32	0,82	2,38	0,52	0,60
jul/2015	4,23	0,81	2,31	0,50	0,67
ago/2015	4,40	0,87	2,36	0,53	0,64
set/2015	4,43	0,85	2,40	0,54	0,65
out/2015	4,50	0,90	2,34	0,59	0,67

nov/2015	4,53	0,93	2,30	0,62	0,67
dez/2015	4,23	0,86	2,20	0,59	0,58
jan/2016	4,30	0,93	2,21	0,58	0,58
fev/2016	4,34	0,88	2,35	0,57	0,54
mar/2016	4,30	0,85	2,32	0,56	0,56
abr/2016	4,67	0,94	2,49	0,60	0,63
mai/2016	4,54	0,87	2,44	0,56	0,67
jun/2016	4,64	0,89	2,53	0,53	0,68
jul/2016	4,40	0,81	2,45	0,48	0,66
ago/2016	4,55	0,84	2,54	0,50	0,67
set/2016	4,72	0,91	2,57	0,54	0,69
out/2016	4,62	0,91	2,48	0,55	0,68
nov/2016	4,62	0,90	2,50	0,57	0,65
dez/2016	4,40	0,88	2,40	0,55	0,57
jan/2017	4,44	0,91	2,35	0,58	0,60
fev/2017	4,53	0,88	2,48	0,58	0,59
mar/2017	4,44	0,94	2,34	0,57	0,60
abr/2017	4,65	0,90	2,53	0,58	0,64
mai/2017	4,42	0,88	2,34	0,54	0,66
jun/2017	4,55	0,90	2,45	0,53	0,67
jul/2017	4,45	0,85	2,44	0,49	0,67
ago/2017	4,49	0,88	2,42	0,50	0,68
set/2017	4,58	0,90	2,46	0,52	0,70
out/2017	4,46	0,87	2,38	0,53	0,67
nov/2017	4,76	0,965	2,539	0,584	0,670
dez/2017	4,46	0,86	2,50	0,55	0,55
jan/2018	4,72	0,95	2,57	0,59	0,61
fev/2018	4,55	0,94	2,42	0,58	0,61
mar/2018	4,69	0,89	2,65	0,57	0,57
abr/2018	4,77	0,95	2,59	0,60	0,63
mai/2018	4,64	0,91	2,51	0,53	0,67
jun/2018	4,61	0,88	2,54	0,51	0,68
jul/2018	4,66	0,86	2,61	0,50	0,69
ago/2018	4,80	0,90	2,66	0,52	0,71
set/2018	4,73	0,91	2,58	0,54	0,70
out/2018	4,91	0,96	2,67	0,59	0,68
nov/2018	4,79	0,95	2,63	0,59	0,62
dez/2018	4,61	0,90	2,55	0,58	0,58
jan/2019	4,83	0,98	2,62	0,59	0,64
fev/2019	4,63	0,99	2,35	0,61	0,67

mar/2019	4,69	0,97	2,53	0,58	0,61
abr/2019	4,61	0,92	2,51	0,57	0,61
mai/2019	4,68	0,92	2,53	0,57	0,65
jun/2019	4,56	0,90	2,45	0,53	0,68
jul/2019	4,60	0,88	2,54	0,50	0,68
ago/2019	4,65	0,91	2,53	0,52	0,69
set/2019	4,65	0,95	2,41	0,56	0,72
out/2019	4,69	0,93	2,47	0,58	0,70
nov/2019	4,79	1,02	2,44	0,62	0,71
dez/2019	4,50	0,93	2,37	0,58	0,61

## DADOS FASE 2

Tabela B.3: Consumo de energia elétrica em Minas Gerais de Janeiro de 2020 a junho de 2021 (EPE, 2021)

mês/ano	Total	Residencial	Industrial	Comercial
Jan/2020	4,71	1,03	2,42	0,61
Fev/2020	4,59	0,97	2,43	0,60
Mar/2020	4,67	0,98	2,54	0,57
Abr/2020	4,33	1,00	2,27	0,48
Mai/2020	4,35	0,93	2,37	0,42
Jun/2020	4,27	0,92	2,27	0,42
Jul/2020	4,54	0,89	2,55	0,42
Ago/2020	4,83	0,99	2,67	0,45
Set/2020	4,70	0,97	2,54	0,47
Out/2020	5,24	1,05	2,73	0,60
Nov/2020	4,86	1,04	2,61	0,54
Dez/2020	4,69	1,01	2,63	0,54
Jan/2021	4,95	1,09	2,63	0,55
Fev/2021	4,61	1,00	2,53	0,50
Mar/2021	4,89	1,01	2,77	0,51
Abr/2021	4,88	1,08	2,68	0,48
Mai/2021	4,71	0,92	2,65	0,46
,Jun/2021	4,79	0,98	2,69	0,46