

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO

JICELENE ELISSA DA SILVA

APLICAÇÃO DE ARRANJOS PLACKETT-BURMAN NA
MODELAGEM DE SÉRIES TEMPORAIS LINEARES

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Produção como parte dos requisitos para obtenção do título de ***Mestre em Ciências em Engenharia de Produção***

Área de Concentração: Engenharia de Produção

Orientador: Prof. Dr. Rafael Coradi Leme

Co-orientador: Prof. PhD. Sebastião Carlos da Costa

Itajubá, MG

2013

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO

APLICAÇÃO DE ARRANJOS PLACKETT-BURMAN NA
MODELAGEM DE SÉRIES TEMPORAIS LINEARES

JICELENE ELISSA DA SILVA

Itajubá, MG

2013

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO

JICELENE ELISSA DA SILVA

APLICAÇÃO DE ARRANJOS PLACKETT-BURMAN NA
MODELAGEM DE SÉRIES TEMPORAIS LINEARES

Dissertação aprovada por banca examinadora em 17 de maio de 2013, conferindo ao autor título de ***Mestre em Ciências em Engenharia de Produção.***

Banca Examinadora:

Prof. Dr. Rafael Coradi Leme (Orientador)

Prof. Dr. Sebastião Carlos da Costa (Co-orientador)

Prof. Dr. Anderson Paulo de Paiva (UNIFEI)

Prof. Dr. Ronã Rinston Amaury Mendes (IFSULDEMINAS)

Itajubá, MG

2013

Ficha catalográfica elaborada pela Biblioteca Mauá
Bibliotecária Jacqueline Rodrigues de Oliveira Balducci- CRB_6/1698

S586

Silva, Jicelene Elissa da

Aplicação de arranjos Plackett-Burman na modelagem de séries temporais lineares / Jicelene Elissa da Silva. – Itajubá, (MG) : [s.n.], 2013.

81 p. : il.

Orientador: Prof. Dr. Rafael Coradi Leme.

Co-orientador: Prof. Dr. Sebastião Carlos da Costa.

Dissertação (Mestrado) – Universidade Federal de Itajubá.

1. Delineamento e análise de experimentos. 2. Séries temporais lineares. 3. Máxima Verossimilhança. I. Leme, Rafael Coradi, orient. II. Costa, Sebastião Carlos da, co-orient. III. Universidade Federal de Itajubá. IV. Título.

Agradecimentos

Agradeço primeiramente a Deus, que me dá forças e coragem para vencer todas as etapas do meu dia-a-dia.

Aos meus pais José Carlos e Cacilda, pelo exemplo de vida e apoio durante toda a minha vida.

Às minhas irmãs Jaqueline, Juçara e Jessica por fazerem parte de minha vida.

Ao professor e amigo Rafael Coradi Leme, pela amizade, orientação e paciência durante o desenvolvimento desse trabalho.

Ao meu namorado Ricardo Xavier pela dedicação, apoio e compreensão.

Aos professores Sebastião Carlos da Costa, Anderson Paiva e Pedro P. Balestrassi pela ajuda em momentos importantes do trabalho.

A todos os professores do Programa de Mestrado em Engenharia de Produção da UNIFEI, pela transferência de conhecimento e troca de experiências.

Por fim, a todos os meus amigos e familiares, que tiveram uma contribuição direta ou indireta para a realização deste trabalho.

*"Don't never prophesy: If you prophesies right, ain't nobody going to remember and if
you prophesies wrong, ain't nobody going to let you forget"*

Mak Twain

Resumo

Previsões baseadas em séries temporais são as formas mais simples e usuais de se fazer previsão. Neste sentido, o interesse nos estudos de modelagem de séries temporais tem crescido cada vez mais em diversas áreas de atividades, tais como industrial, comercial, econômica, saúde, financeiro, etc. Além da previsão, o estudo de séries temporais também permite descrever os padrões de uma resposta de interesse, tais como tendência e sazonalidade. Esta análise permite a construção de modelos que explicam o comportamento da série, e até mesmo permite o controle do processo em geral. Apesar de muitas aplicações, a identificação de modelos não é simples, e se faz necessário a utilização de várias técnicas estatísticas que, muitas vezes, acabam se baseando em tentativas e erros, e até mesmo em uma busca exaustiva. Com essa motivação, esse estudo propõe uma abordagem alternativa baseada em Delineamento e análise de Experimentos (DOE) para identificar a influência de cada parâmetro de um modelo AR, MA e ARMA em algumas respostas de desempenho de previsão de interesse, como (incondicional) variância, o erro médio quadrado, ou qualquer outra estatística. Nesse trabalho, é apresentada a metodologia proposta aplicada a doze séries temporais lineares genéricas, construídas com a intenção de representar os diferentes comportamentos e tipos de modelos existentes e também a aplicação para três processos reais. Os parâmetros dos modelos estudados foram estimados através do método de máxima verossimilhança e comparados com os valores reais. Para todas as séries, principalmente para as séries reais a metodologia proposta apresentou bons resultados, porém alguns pontos de melhorias são sugeridos.

Palavras chave: Delineamento e análise de Experimentos, Séries temporais lineares, Máxima Verossimilhança.

Abstract

Predictions based on time series are the simplest forms and usual to make prediction. In this sense the interest in the modeling studies of time series has grown increasingly in diverse fields of activities, such as industrial, commercial, economic, financial, health, etc. Besides forecasting, the study of time series also allows to describe the patterns of a response of interest, such as trend and seasonality. This analysis permits the construction of models that explain the behavior of the series, and even process control in general. In spite of many applications, the models identification is not straightforward, it's necessary to use several statistical techniques, and often ends up being trial and error, or even exhaustive search. With the motivation this study proposes an alternative approach based in Design of Experiments (DOE) to identify the impact of each parameter of an AR, MA and ARMA model on some forecasting performance responses of interest, such as (unconditional) variance, mean square error, or any other statistic. In this framework is presented the proposal methodology applied for twelve linear time series built in order to represent the different kind of models and the application for three real processes. The models parameters were estimated trough the maximum likelihood method and compared with the real values. For all series, especially real series, the methodology had good results, however some improvements are suggested.

Keywords: Design of Experiment, Linear Time Series, Maximum Likelihood.

Lista de Figuras

1	Métodos de Previsão. Fonte: Adaptado de MILESKI, 2007.	2
2	Comportamento da corrente de um processo de soldagem MIG/MAG por Curto-Circuito.	8
3	Representação das componentes de uma série temporal. Adaptado (MOREIRA et al., 2010)	10
4	Visão da Fig. 2 amplificada, mostrando o padrão sazonal da série. . . .	13
5	Correlograma (ou gráfico ACF) para os dados da produção de bebidas (MAKRIDAKIS,1998).	17
6	Estágios do ciclo interativo para os modelos Box-Jenkins. (Adaptado: CASTRO,1995)	24
7	Comparação entre as curvas $\sin(x)$, expansão de Taylor para $\sin(x)$ e AR(2) para $\sin(x)$	26
8	- Funções de autocorrelação ACF (à esquerda) e PACF (à direita) de um modelo AR(1).	28
9	Funções de autocorrelação FAC (à direita) e FACP (à esquerda) de um modelo MA(1)	30
10	Representação dos testes bilateral e unilateral à direita	40
11	Pseudo-código para cálculo das respostas para análise do DOE	52
12	Autocorrelação Parcial dos modelos AR	53
13	Diagrama de pareto e efeitos principais para os modelos AR	58
14	Validação dos modelos AR	59
15	Autocorrelação dos modelos MA	60
16	Validação dos modelos MA	61
17	Diagrama de pareto e efeitos principais para os modelos ARMA	62

Lista de Figuras

18	Validação dos modelos ARMA	64
19	Autocorrelação da série M	66
20	Autocorrelação da série M após a diferenciação	66
21	Autocorrelação das séries de pressão e temperatura de um poço de petróleo	67
22	Autocorrelação das séries de tensão e corrente de um processo de soldagem MIG/MAG por curto-circuito	68
23	Diagrama de pareto e efeitos principais para os modelos das séries reais	70
24	Validação do modelo de previsão da Série M	72
25	Análise dos resíduos do modelo da série M	72
26	Validação do modelo para os dados de Pressão da série N	72
27	Validação do modelo para os dados de Temperatura da Série N	73
28	Análise de resíduos para o modelo da série N	73
29	Validação do modelo para os dados de tensão da série O	73
30	Validação do modelo para os dados de corrente da Série O	74
31	Análise de resíduos para Série O	74

Lista de Tabelas

1	Comportamento das funções ACF e PACF para os modelos AR(p), MA(q) e ARMA(p,q)	31
2	Erro do Tipo I e Tipo II	40
3	Principais Técnicas de Análise de Experimentos e características fundamentais (Adaptado NILO JÚNIOR, 2003)	44
4	Arranjo Plackett-Burman padrão para $k = 11$ $n = 12$	46
5	Séries Temporais Lineares Genéricas	49
6	Matriz experimental para modelos AR	50
7	Matriz experimental para modelos MA	50
8	Matriz experimental para modelos ARMA	51
9	Exemplo de um modelo AR representado pela linha da matriz experimental	51
10	Respostas obtidas para os modelos AR	53
11	Efeitos significativos para os modelos AR	54
12	Coeficientes estimados para os modelos AR	55
13	Respostas obtidas para os modelos MA	56
14	Coeficientes estimados para os modelos MA	57
15	Respostas obtidas para os modelos ARMA	60
16	Coeficientes estimados para os modelos ARMA	63
17	Respostas obtidas para os modelos das séries reais	69
18	Coeficientes estimados para os modelos das séries reais	71

Lista de Acrônimos

ACF	<i>Autocorrelation Function</i>
ANOVA	<i>Analysis of variance</i>
AR	<i>Autoregressive Model</i>
ARCH	<i>Autoregressive Conditional Heteroscedastic Model</i>
ARIMA	<i>Autoregressive Integrated Moving Average Model</i>
ARMA	<i>Autoregressive Moving Average Model</i>
DOE	<i>Design and Analysis of Experiments</i>
GARCH	<i>Generalized Autoregressive Conditional Heteroscedastic Model</i>
IGCM-C	Índice Geral do Mercado Imobiliário - Comercial
MA	<i>Moving Average model</i>
MAG	<i>Metal Active Gas</i>
MDN	<i>Mixture Density Network</i>
MIG	<i>Metal Inert Gas</i>
PACF	<i>Partial Autocorrelation Function</i>
PCA	<i>Principal Component Analysis</i>
PIB	Produto Interno Bruto
RNA	Redes Neurais
SARIMA	<i>Sazonal Autoregressive Integrated Moving Average Model</i>

Sumário

1	Introdução	1
1.1	Considerações Iniciais	1
1.2	Objetivos	4
1.3	Justificativa	4
1.4	Limitação do Trabalho	5
1.5	Estrutura do Trabalho	5
2	Fundamentação Teórica	7
2.1	Séries Temporais	7
2.2	Processos Estocásticos Estacionários	10
2.3	Transformações	12
2.4	Função Autocorrelação e Autocovariância	15
2.5	Ruído Branco	19
2.6	Análise de Regressão	20
2.7	Modelos de Previsão Box-Jenkins	23
3	Técnicas Estatísticas	34
3.1	Análise das Componentes Principais	34
3.2	Método Mínimos Quadrados Ordinários	35
3.3	Método da Máxima Verossimilhança	36
3.4	Teste de Hipótese	37
3.5	Delineamento e Análise de Experimentos	42
3.6	Análise de Variância	46

4	Desenvolvimento	48
4.1	Metodologia Proposta	48
4.2	Resultados Obtidos	65
5	Conclusões	75
5.1	Considerações Iniciais	75
5.2	Conclusões	75
5.3	Recomendações Para Trabalhos Futuros	76

1 *Introdução*

Este capítulo teve como objetivo retratar a importância de uma boa previsão em vários ramos de atividades. Além disso, familiarizar o leitor com o tema proposto, descrever a proposta do trabalho, justificar sua importância e apresentar suas limitações.

1.1 **Considerações Iniciais**

Uma boa previsão pode ser essencial para a tomada de decisão por ser um suporte importante para um planejamento efetivo e eficiente. Mas tentar prever o futuro envolve incertezas, por isso o interesse no estudo de métodos que possibilitem encontrar um bom modelo, ou seja, um modelo que retorne uma resposta com mínimo erro possível, vem crescendo cada vez mais em diversos ramos de atividades, como industrial, comercial, econômico-financeiro, saúde, etc.

Apesar das inúmeras situações que requerem previsões, existem apenas duas possibilidades: os métodos qualitativos e os quantitativos. O primeiro, geralmente de natureza subjetiva, requer julgamentos por parte de especialistas e são frequentemente usados em casos que se tem pouca ou nenhuma informação histórica sobre a variável de interesse. De uma forma geral, os métodos quantitativos de previsão utilizam o comportamento atual e passado de uma variável para se prever seus valores no futuro. Existem vários tipos de modelos de previsão de uso geral. Os três modelos mais utilizados são os modelos de regressão, os modelos de alisamento exponencial e os modelos de séries temporais em geral (MONTGOMERY et al., 2008) .

Os métodos de previsão baseados em séries temporais são os métodos mais simples e usuais de previsão (TUBINO, 2000) .

Uma série temporal é uma sequência cronológica de observações de uma variável de interesse, geralmente igualmente espaçadas (MONTGOMERY et al., 2008). A previsão a partir dessa análise parte do pressuposto que os dados da série no passado podem trazer informações sobre o comportamento da série no futuro. Mas apenas

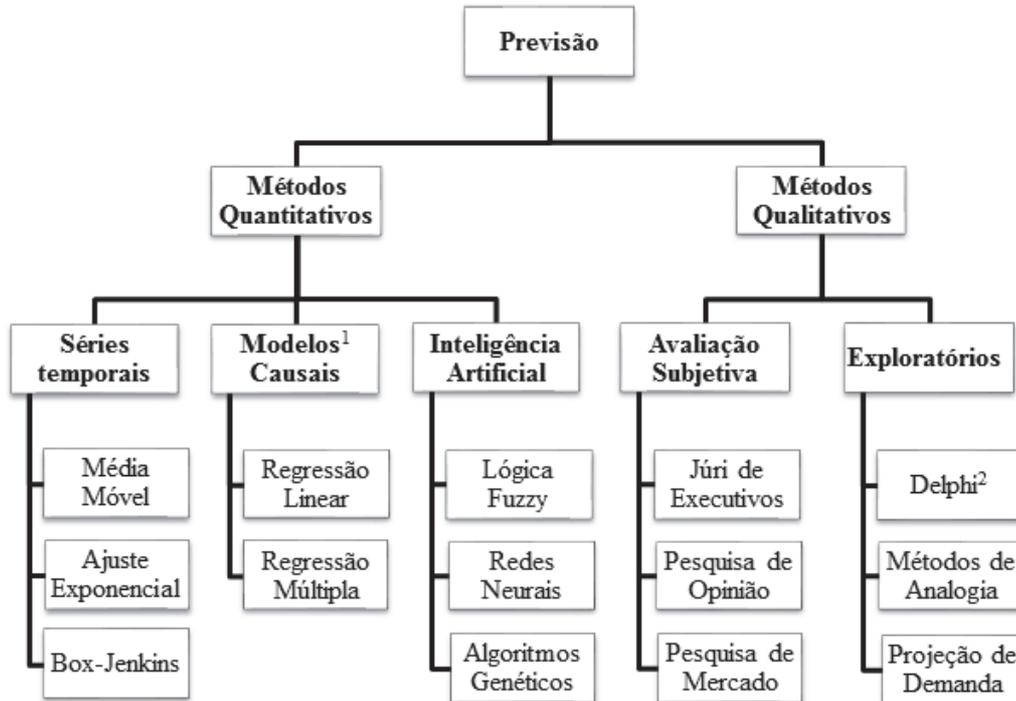


Figura 1: Métodos de Previsão. Fonte: Adaptado de MILESKI, 2007.

os dados históricos por si só podem não ser suficientes para se obter um modelo de previsão eficiente. Para isso, se faz necessário uma combinação do uso de técnicas estatísticas e algoritmos que podem ser de fáceis aplicações até métodos complexos que demandam uma grande quantidade de tempo para se chegar a uma solução¹².

Inúmeras são as metodologias para se fazer previsões, com graus variados de eficiência. Tais métodos podem ser clássicos, ou seja, simplesmente estatísticos ou até mesmo utilizar de técnicas baseadas em inteligência artificial (ZAGRAJEK e WERON, 2002).

Os métodos de previsão estão presentes em diversas áreas, sendo utilizadas como elemento fundamental para decisões de planejamento estratégico. Na gestão operacional, várias empresas utilizam a técnica para prever as vendas e a demanda por serviços e suprimentos com a finalidade de se ter um controle da sua capacidade produtiva. No marketing, a previsão é importante em vários aspectos, como na análise do retorno das vendas em relação às propagandas com o intuito de se ter um controle

¹Os métodos causais descrevem a resposta do processo (variável dependente) em função de variáveis independentes, ou seja, fatores que podem influenciar o comportamento do processo, tais como clima, inflação, etc.

²Técnica mais difundida na metodologia qualitativa. Painel de especialistas é interrogado utilizando-se uma sequência de questionários, sendo que, a partir do segundo, eles são construídos a partir das respostas do anterior e assim por diante. (SILVEIRA, 2002)

de gastos, avaliar a eficácia da campanha e verificar previamente se as metas serão atingidas. Na área de finanças e gerenciamento de riscos utiliza-se a previsão para prever o retorno de investimentos, taxa de juros, preço de fechamento de ações e taxa de câmbio. A gestão de riscos financeiros requer previsões da volatilidade dos retornos de ativos de modo que os riscos associados às carteiras de investimento possam ser avaliados e decisões possam ser tomadas de forma eficaz. Na economia o governo, bem como instituições financeiras, necessitam da previsão das principais variáveis econômicas do país tais como o PIB, taxa de desemprego, taxa de juros, inflação, etc. No controle de processos industriais, a previsão pode ser útil na avaliação prévia da qualidade do produto, apontando falhas existentes no processo, permitindo que se façam os ajustes necessários nas variáveis de importância. Na demografia, existe a necessidade de se fazer previsões sobre a população do país, tais como nascimentos, mortes e migração para auxiliar a política de planejamento e definir as ações do serviço social. Além disso, muitas empresas utilizam as previsões da faixa etária da população para fazer planos estratégicos sobre desenvolvimento de novas linhas de produtos ou serviços.

Tomé e Latorre (2001) analisaram as tendências da mortalidade infantil e suas componentes para o Município de Guarulhos, no período de 1971 a 1998 através de modelos de séries temporais. Aviv (2002) analisou os benefícios que a previsão pode trazer em processos de reposição de estoque e como algumas mudanças no mercado podem explicar a demanda futura, enfatizando que a demanda evolui de acordo com um modelo auto-regressivo de séries temporais. Cardoso (2005) propôs um modelo quantitativo para a previsão da volatilidade da demanda de energia elétrica para consumidores livres em regime de curto prazo, através da simulação de modelos GARCH para séries temporais univariadas, avaliando o comportamento de cargas industriais. Ribeiro e Paula (2000) aplicaram a metodologia de séries temporais para prever a população brasileira até 2020. Já Engles (1982), estimou a média e a variância da inflação no Reino Unido através de modelos não lineares de séries temporais.

Os modelos clássicos ou estatísticos podem ser representados pela família dos modelos lineares, sendo eles os modelos AR (*Auto Regressive*), MA (*Moving Average*), ARMA (*Auto Regressive - Moving Average*), ARIMA (*Auto Regressive Integrated Moving Average*), SARIMA (*Seasonal Auto Regressive Integrated Moving Average*) e pela família dos modelos não-lineares, os modelos ARCH (*Autoregressive Conditional Heteroscedasticity*), GARCH (*Generalized Autoregressive Conditional Heteroscedasticity*) entre outros. Já os modelos baseados em inteligência artificial podem ser representados pelos modelos obtidos através de Redes Neurais (RNA), e MDN (*Mixture*

Density Network).

1.2 Objetivos

Este trabalho tem por finalidade apresentar um método otimização de obtenção de modelos de previsão utilizando técnicas convencionais de modelagem de séries temporais combinadas com o delineamento e análise de experimentos. Tal objetivo pode ser desdobrado em três outros secundários sendo estes:

- Propor um método de identificação de modelos de séries temporais lineares utilizando a técnica de delineamento e análise de experimentos.
- Comparar a eficiência dos métodos tradicionais de identificação de modelos de séries temporais com o método proposto.
- Reduzir os erros estatísticos de tipo I, ou seja, rejeitar coeficientes de um modelo de série temporal linear sendo que o mesmo é significativo, e tipo II, aceitar coeficientes dos modelos lineares sendo que esses são insignificantes.

1.3 Justificativa

A previsão é um processo metodológico para a determinação de dados futuros baseados em modelos estatísticos, matemáticos ou econométricos, ou ainda em modelos subjetivos apoiados em uma metodologia de trabalho clara e previamente definida (MARTINS e LAUGENI, 1999).

A necessidade de se conhecer os conceitos básicos de séries temporais vem da necessidade de se fazer previsões. Conseguir prever o futuro pode ser essencial em uma tomada de decisão, além disso, pode ser um suporte importante para um planejamento eficiente. Mas como já dito anteriormente, tentar prever o futuro envolve incertezas, e previsões para horizontes maiores possuem graus de incertezas ainda maiores (CORRÊA et al., 2000) .

A modelagem de séries temporais não é trivial e, muitas vezes, podem conter muitos parâmetros, para se conseguir uma boa aproximação. Além disso, os métodos que são baseados em análises das funções de autocorrelação e autocorrelação parcial podem dificultar a identificação do modelo, quando tratamos de modelos ARMA por exemplo, pois apresentam o mesmo comportamento em ambas funções.

A busca por métodos eficientes e eficazes que resultem em bons modelos para se fazer previsões é uma questão que desperta a atenção de muitos pesquisadores. Assim, a principal justificativa para a escolha desse tema consiste na preocupação de minimizar a utilização de técnicas baseadas em tentativas e erros e otimizar o processo de modelagem de séries temporais, com a finalidade de identificar modelos matemáticos eficientes e eficazes, priorizando a qualidade da obtenção de valores futuros de uma série temporal linear.

1.4 Limitação do Trabalho

Este trabalho apresenta as seguintes limitações:

- A metodologia apresentada é válida apenas para dados de séries temporais univariadas.
- Foi utilizada a função de máxima verossimilhança incompleta.
- Entre as respostas do processo, este trabalho se propôs apenas a otimização do erro, ou seja, a diferença entre o valor previsto pelo modelo e o valor real e da variância.
- A metodologia foi proposta apenas para modelos de séries temporais lineares.
- Foram desconsideradas as interações entre os coeficientes.
- Foi utilizado apenas o arranjo Plackett-Burman para modelar o erro de previsão a partir da variação dos parâmetros.

1.5 Estrutura do Trabalho

O presente trabalho está dividido em cinco capítulos:

1. Introdução: Este capítulo teve como objetivo retratar a importância de uma boa previsão em vários ramos de atividades. Além disso, familiarizar o leitor com o tema proposto, descrever a proposta do trabalho, justificar sua importância e apresentar suas limitações.
2. Fundamentação teórica: Este capítulo introduz conceitos importantes e fundamentais na análise e previsão de séries temporais, além de apresentar alguns

modelos estatísticos de previsão lineares como AR, MA, ARMA e ARIMA e SARIMA.

3. Técnicas Estatísticas: Nesse capítulo são introduzidas algumas técnicas estatística utilizadas nesse trabalho, como Análise das Componentes Principais, o Método dos Mínimos Quadrados Ordinários, Método da Máxima Verossimilhança, teste de hipóteses, Delineamento e Análise de Experimentos e Análise de Variância.
4. Resultados: Nesse capítulo é apresentado o desenvolvimento da pesquisa e os resultados alcançados.
5. Conclusões e Recomendações: Por fim, são apresentadas as conclusões e recomendações, bem como sugestões para trabalhos futuros.

2 Fundamentação Teórica

Este capítulo introduz conceitos importantes e fundamentais sobre a análise e previsão de séries temporais, além de apresentar alguns modelos estatísticos de previsão lineares como AR, MA, ARMA e ARIMA e SARIMA.

2.1 Séries Temporais

Pode-se conceituar uma serie temporal como qualquer conjunto de observações ordenadas no tempo (MILONE e ANGELINI 1995).

Por convenção, existem dois tipos de séries temporais, as chamadas contínuas, que são séries cuja variável de interesse é mensurada continuamente, sem intervalos, definindo o conjunto $T = \{t : t_1 < t < t_2\}$, sendo denotada por $\{X(t) : t \in T\}$; e as chamadas discretas, diferentemente da anterior a variável observada é mensurada com intervalos de tempo específicos, geralmente equidistantes, definindo o conjunto $T = \{t_1, \dots, t_n\}$ sendo a série temporal denotada por $\{X_t : t \in T\}$. Para facilitar o entendimento pode-se fazer uma analogia das observações de uma série temporal qualquer, com as variáveis aleatórias $X_1, X_2, X_3, X_4, \dots, X_n$ sendo assim, defini-se que X_1 representa o valor da série para $t = 1$, X_2 o valor da observação do período seguinte, ou seja, $t = 2$, e assim por diante.

Ao se utilizar essa analogia podemos perceber que esse conjunto de variáveis aleatórias espaçadas por um período t refere-se a um processo estocástico¹ e nesse caso, como ocorre tipicamente, trata-se de uma série discreta, onde o índice $t \in I$. Geralmente, ao se analisar séries temporais, utilizam-se séries discretas, pois pouca ou nenhuma informação é perdida, desde que o intervalo de amostragem seja pequeno o suficiente, assim não sendo necessário o uso de séries contínuas.

Outra definição dada a uma série temporal é a realização de um processo esto-

¹Um processo estocástico é uma família de variáveis randômicas $\{X_t, t \in T\}$ definida no espaço de probabilidade (Ω, F, P) . (BROCKWELL e DAVIS, 1990)

cástico definido por observações sequenciais regidas por leis probabilísticas. Sendo assim, pode-se definir uma série temporal como uma amostra de um determinado processo estocástico (RIBEIRO e PAULA, 2000).

A Fig. 2 representa uma série temporal discreta. O gráfico apresentado ilustra o comportamento da corrente de um processo de soldagem MIG/MAG por curto-circuito de um experimento realizado no laboratório de Soldagem da Universidade Federal de Itajubá com frequência de amostragem 4000 Hz.

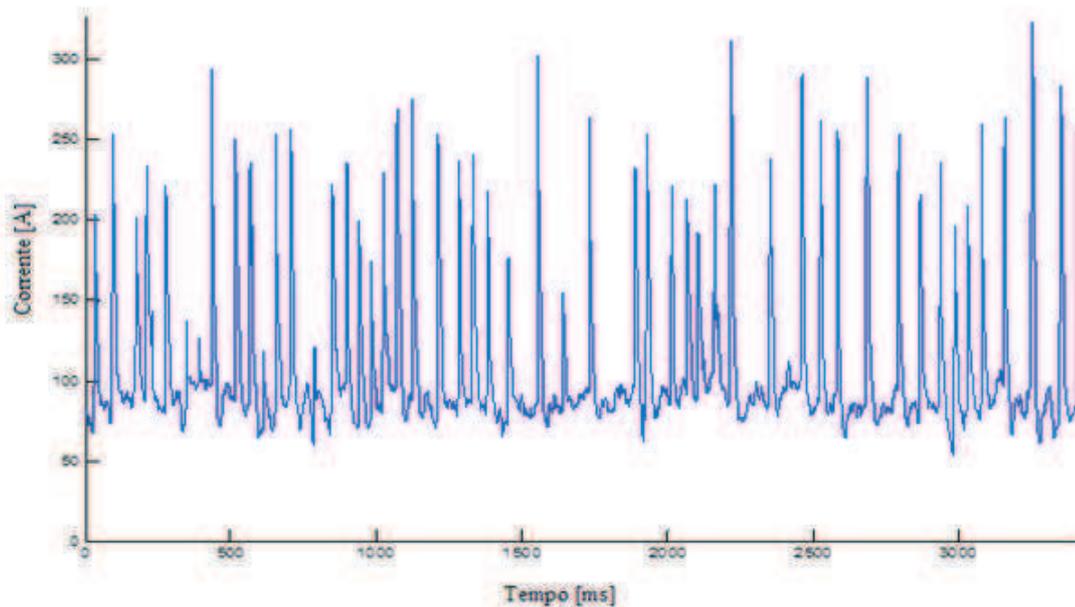


Figura 2: Comportamento da corrente de um processo de soldagem MIG/MAG por Curto-Circuito.

Uma das principais características das séries temporais é que as sucessivas observações são correlacionadas e por esse motivo deve-se levar em conta a ordem dos dados, pensando assim, pode-se entender esse tipo de série como regida por duas variáveis: a variável de interesse e o tempo (CHATFIELD, 2000).

De um modo geral, os principais objetivos de se analisar séries temporais podem ser:

- **Descrição:** Descrever a série de estudo através das características estatísticas e/ou gráficos que resumem o comportamento da série.
- **Modelagem:** Encontrar um modelo adequado que descreve o componente gerador do processo. Um modelo univariado é baseado apenas em valores passados da variável de interesse para a previsão dos valores futuros, enquanto um modelo multivariado os valores previstos dependem dos valores passados, mas

também de valores presentes e valores passados de outras variáveis, ou seja, muitas vezes, uma série pode ser utilizada para explicar outra série.

- **Previsão:** Estimar valores futuros da série estudada.
- **Controle:** Permitir tomadas de decisões a fim de se controlar o processo seja esse industrial, econômico, etc.

O gráfico das observações espaçadas sequencialmente em relação ao tempo pode auxiliar como um direcionador na análise desse tipo de série. Através dele é possível se fazer uma inspeção visual das principais propriedades de uma série tais como tendência, sazonalidade, descontinuidades, pontos de inflexão e mudanças suaves na estrutura, além disso, é vital para a escolha do método de previsão. As principais componentes de uma série temporal são:

- **Tendência:** É a componente mais relevante de uma série. Variação que tende ao crescimento ou ao decréscimo, podendo ser linear ou não-linear. Uma série que não possui tendência é chamada de estacionária. Um exemplo de série que apresenta essa componente é a do IGCM-C² índice que tem apresentado um crescimento todos os anos.
- **Sazonalidade:** Variação que ocorre em muitas séries quando um determinado padrão se repete em determinados períodos, geralmente no intervalo de um ano. Como por exemplo, o comportamento da taxa de venda de sorvetes que durante os meses de verão apresenta um aumento significativo em relação aos outros meses do ano.
- **Ciclo:** Variação cíclica que ocorre em períodos regulares durante intervalos maiores do que um ano. Uma série que apresenta um comportamento cíclico é o número anual de lincas capturados em armadilhas entre os anos de 1821 e 1934 no Canadá. Existe um padrão cíclico em torno de 10 ou 11 anos.
- **Erro aleatório:** Componente resultante quando se retira a tendência, a sazonalidade e o ciclo de uma série temporal. Essa componente geralmente tem origem totalmente randômica e de natureza imprevisível.

Enfatizando a importância de um bom gráfico “temporal”, a má escolha das escalas, muitas vezes, dificulta a interpretação dessa ferramenta o que pode atrapalhar a visualização de propriedades importantes da série, como por exemplo, a sua característica

²IGCM-C - Índice Geral do mercado imobiliário - Comercial - Indicador de rentabilidade de imóveis comerciais - escritórios, hotéis, shoppings e outros.

quanto à estacionariedade, conceito que será abordado no item a seguir. Apesar desse tipo de gráfico ter uma grande importância nas análises de séries temporais, a ordenação das observações em relação ao tempo dificulta a utilização de outros tipos que são comuns em análises estatísticas, como por exemplo o histograma, o qual é normalmente utilizado para identificar a distribuição do objeto de estudo, e pode ter seu comportamento distorcido devido à presença de sazonalidade e tendência. No caso de uma série univariada, existe apenas a relação entre as observações e o tempo, o que não traz problema na análise das interações. No entanto, quando existem, por exemplo, duas variáveis aleatórias ordenadas no tempo, na verdade estamos estudando quatro variáveis, ou seja, as duas variáveis medidas e seus respectivos valores de tempo. Nesse caso, estamos tratando de séries multivariadas que exigem um pouco mais de aprofundamento no tema e não serão abordadas nesse trabalho.

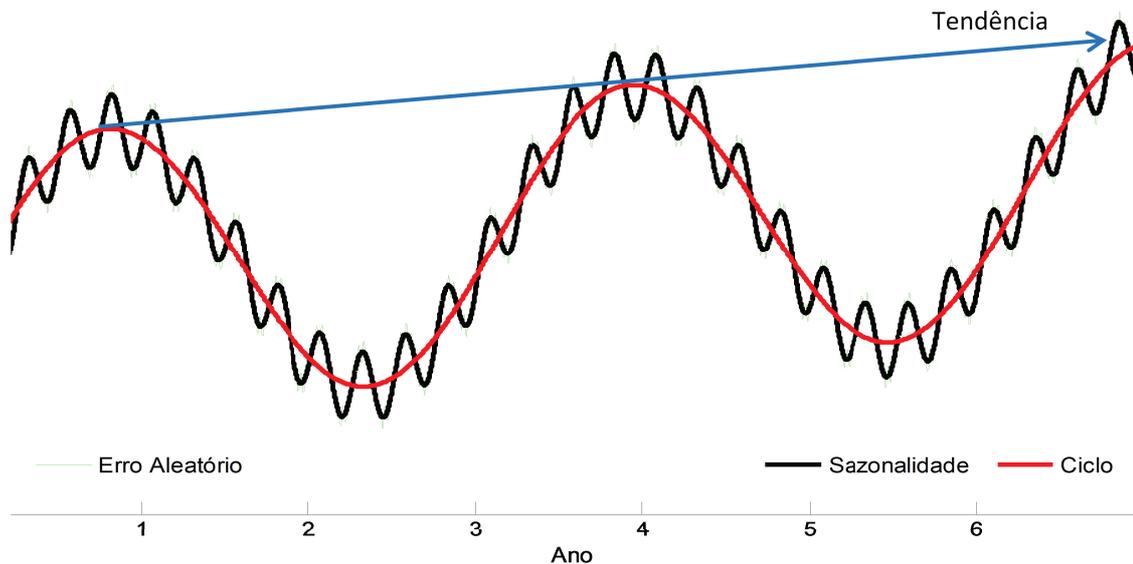


Figura 3: Representação das componentes de uma série temporal. Adaptado (MOREIRA et al., 2010)

Outro gráfico de grande importância na análise das séries temporais é o chamado correlograma, que será abordado nos próximos itens após a apresentação de alguns conceitos essenciais para seu entendimento.

2.2 Processos Estocásticos Estacionários

Um conceito muito importante em séries temporais é o da estacionariedade de uma série. Um processo estocástico é estacionário se suas propriedades estatísticas

são invariantes no tempo, ou seja, se a distribuição de probabilidade do conjunto de observações $X_t, X_{t+1}, \dots, X_{t+n}$ é similar a distribuição de probabilidade do conjunto de observações $X_{t+k}, X_{t+k+1}, \dots, X_{t+k+n}$ (MONTGOMERY et al., 2008).

Estacionariedade implica em um tipo de equilíbrio estatístico ou até mesmo em uma estabilidade dos dados. De uma forma geral, pode-se dizer que uma série estacionária tem média constante definida por:

$$\mu_X = E(X) = \int_{-X}^X X f(X) dX \quad (2.1)$$

e variância constante definida por:

$$\sigma_X^2 = Var(X) = \int_{-\infty}^{\infty} (X - \mu_X)^2 f(X) dX \quad (2.2)$$

A média e a variância de uma amostra podem ser utilizadas para estimar esses parâmetros. Considerando as observações X_1, X_2, \dots, X_t de uma série temporal, a média amostral pode ser definida por:

$$\bar{X} = \hat{\mu}_X = \frac{1}{T} \sum_{t=1}^T X_t \quad (2.3)$$

e variância amostral como:

$$s^2 = \hat{\sigma}_X^2 = \frac{1}{T} \sum_{t=1}^T (X_t - \hat{X})^2 \quad (2.4)$$

Em séries temporais é comum que se utilize apenas T como divisor da equação ao invés de $T - 1$ como é usado comumente. Isso ocorre, pois T geralmente é relativamente grande. Sendo assim a diferença entre T e $T - 1$ não é significativa (MONTGOMERY et al., 2008).

Existem duas possíveis classificações para estacionariedade de uma série temporal, a de fracamente estacionária e a de estritamente estacionária. Se assumir-se que a função distribuição de um processo estocástico é exatamente a mesma independentemente do deslocamento do tempo, o processo é dito estritamente estacionário. Já para conceituar um processo fracamente estacionário é necessários introduzir alguns conceitos de estacionariedade para momentos de um processo estocástico X_t .

1. Estacionariedade da média: Um processo tem média estacionária se $E[X_t] =$

$\mu_t = \mu$ é constante para todo t .

2. Estacionariedade da variância: Um processo tem variância estacionária se $V[X_t] = E[(X_t - \mu_t)^2] = \sigma^2$ é constante e finita para todo t .
3. Estacionariedade da covariância: Um processo tem covariância estacionária se $Cov[X_t, X_s] = E[(X_t - \mu_t)(X_s - \mu_s)] = \gamma_{(|s-t|)}$ é uma função somente da diferença do tempo entre duas variáveis aleatórias e independente do tempo atual.

Como a estacionariedade da variância resulta da estacionariedade da covariância para $s = t$, um processo é considerado fracamente estacionário quando possui média e variância estacionária (KIRCHGÄSSNER e WOLTERS, 2007).

2.3 Transformações

Muitos métodos estatísticos funcionam melhor quando os dados estão normalmente distribuídos ou, pelo menos, simetricamente distribuídos e possuem variância constante (RUPPERT, 2011).

Existem alguns tipos de transformações que são utilizadas em análises de séries temporais, geralmente com o intuito de se estabilizar a variância dos dados. As séries transformadas geralmente são menos assimétricas e mais constantes comparadas às séries originais.

A heterocedasticidade³ é uma característica muito comum em dados desse tipo de séries (MONTGOMERY et al., 2008). A Fig. 3 mostra que a série em questão apresenta um comportamento padrão sazonal na magnitude da variância. Aumentando-se o gráfico, pode-se notar que a variabilidade da série para $t = 560$ a 580 ms é menor do que a de $t = 580$ a 600 ms. Esse tipo de padrão pode ser identificado repetitivamente por toda a série.

O uso das transformações também é indicado para transformar os dados em uma série com características mais perto de uma distribuição normal, caso não as apresente. Além disso, se a sazonalidade da série apresentar efeito multiplicativo, ou seja, se o tamanho do efeito sazonal aumentar de acordo com o valor local da média, é interessante transformá-lo em um efeito aditivo⁴, pois efeito linear é geralmente mais fácil de se trabalhar (CHATFIELD, 2000).

³O comportamento não constante da variância de uma série é chamado de heterocedasticidade de uma série. Ao contrário, a homocedasticidade é a característica de a série possuir variância constante.

⁴Efeito sazonal aditivo é quando o efeito sazonal é independente do valor da local da média.

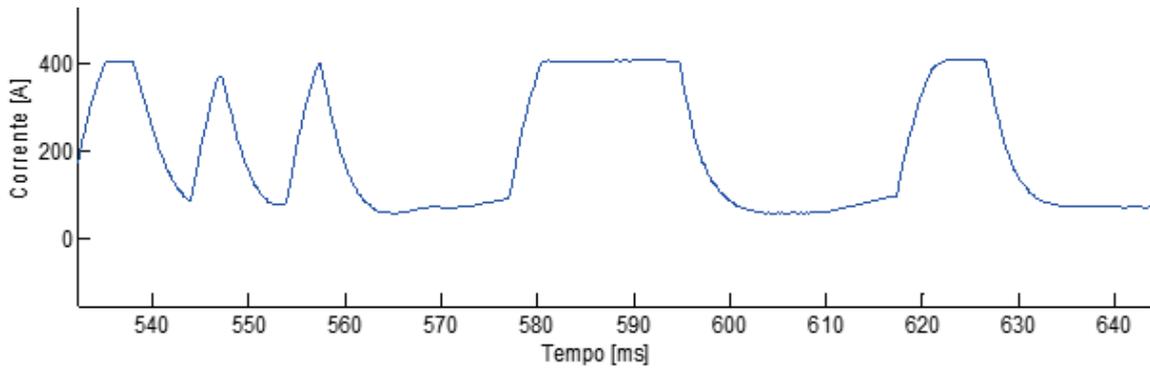


Figura 4: Visão da Fig. 2 ampliada, mostrando o padrão sazonal da série.

Existe um tipo de transformação clássica conhecida como transformação Box-Cox. Sendo X_t uma série temporal, sua transformação é dada por:

$$y_t = \begin{cases} (X_t^\lambda - 1)/\lambda & , \lambda \neq 0 \\ \log(X_t) & , \lambda = 0 \end{cases} \quad (2.5)$$

onde λ denota o parâmetro de transformação o qual geralmente é escolhido empiricamente através de ajuste do modelo para vários λ , selecionando o valor de λ que retorne o menor valor para a soma quadrática residual. Se $\lambda = 1$, então não existe transformação. Existem valores típicos para λ , sendo eles $\lambda = 0.5$ (Transformação raiz quadrado), $\lambda = 0$ (Transformação logarítmica), $\lambda = -0.5$ (Transformação recíproca) e $\lambda = -1$ (Transformação inversa).

Na Eq. (2.5), a subtração de 1 de X_t^λ e a divisão por λ não são essenciais, mas elas fazem a transformação ser contínua em $\lambda = 0$ desde que:

$$\lim_{\lambda \rightarrow 0} \frac{X_t^\lambda - 1}{\lambda} = \log(X_t) \quad (2.6)$$

A transformação logarítmica é provavelmente a transformação mais utilizada em análise de dados, sendo a raiz quadrada a segunda. O logaritmo estabiliza a variância da variável da qual o desvio padrão condicional é proporcional a sua média condicional (RUPPERT, 2011).

Apesar das transformações serem muito utilizadas em análise de séries temporais, existe pouca evidência que a utilização de uma transformação não-linear melhore as previsões (NELSON e GRANGER, 1979) e podem surgir dificuldades na interpretação de um modelo para uma série transformada.

Complementando as transformações, existem ajustes em modelagem de séries or-

denadas no tempo que são amplamente utilizadas, sendo que os ajustes de tendência e de sazonalidade são os dois que possuem maior espaço.

Uma série temporal que apresenta tendência é não estacionária. A eliminação dessa componente pode simplificar, e muito, a modelagem e previsão. Um modo comum de se fazer isso é através da descrição da tendência utilizando um modelo de regressão e eliminando-o das observações originais. Outro modo de se eliminar essa componente é através da diferenciação dos dados, que pode ser descrita como:

$$Y_t = X_t - X_{t-1} = \nabla X_t \quad (2.7)$$

onde ∇ é o operador de diferença. Outra forma de descrever a diferenciação é através do “*Backshift operator*” B definido por $BX_t = X_{t-1}$, sendo assim:

$$Y_t = (1 - B)X_t = \nabla X_t = X_t - X_{t-1} \quad (2.8)$$

onde $\nabla = (1 - B)$. As diferenciações podem ser feitas sucessivamente até que a tendência seja eliminada. Desse modo, de uma forma geral, pode-se escrever que:

$$B^d X_t = X_{t-d} \quad (2.9)$$

$$\nabla^d = (1 - B)^d \quad (2.10)$$

Sazonalidade, ou tendência e sazonalidade, são componentes comumente presentes em séries temporais. A técnica de diferenciação, também pode ser utilizada para eliminar a sazonalidade. Pode-se definir a diferenciação para a sazonalidade como:

$$\nabla_d X_t = (1 - B^d) X_t = X_t - X_{t-d} \quad (2.11)$$

Por exemplo, uma situação muito comum é uma série de dados mensurada mensalmente com padrão sazonal de um ano, ou seja, um “*lag-operador*” d igual a 12. Desse modo, a Eq. (2.10) ficará:

$$\nabla_{12} X_t = (1 - B^{12}) X_t = X_t - X_{t-12} \quad (2.12)$$

Quando uma série possui ambas componentes, tendência e sazonalidade, pode-se fazer a diferenciação sequencialmente para eliminar os dois efeitos. Primeiramente, retira-se o efeito sazonal para, então, utilizar a diferença de um ou mais vezes, utilizando o operador de diferença regular para remover a tendência.

2.4 Função Autocorrelação e Autocovariância

Sejam X e Y duas variáveis aleatórias. A covariância é uma estatística que indica a relação linear de duas variáveis aleatórias. Para um processo estacionário, essa estatística é definida por:

$$Cov_{XY} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \quad (2.13)$$

onde \bar{X} e \bar{Y} denotam as médias de X e Y respectivamente e n o número de observações de cada variável (MAKRIDAKIS, 1998).

Analisando a Eq. (2.13) pode-se notar que a unidade resultante é de difícil interpretação, o que faz perceber que a covariância não é palpável como estimador de relacionamento, podendo assumir valores de menos até mais infinito, sem ter um ponto de referência que delimite um forte grau de relacionamento de um grau fraco. O coeficiente de correlação, designado por r , é uma especial medida de correlação da covariância que trata o problema de escala mencionado. Se a covariância for dividida pelos desvios padrões das variáveis de estudo (S_X e S_Y), então as unidades no numerador e no denominador se cancelam e o valor calculado se torna um número adimensional o qual é denominado coeficiente de correlação entre as variáveis X e Y . Desse modo, a equação que denota tal estatística é definida por:

$$r_{XY} = \frac{Cov_{XY}}{S_X S_Y} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2} \sqrt{\sum (Y_i - \bar{Y})^2}} \quad (2.14)$$

Essa padronização restringe o intervalo de r_{XY} para $[-1, +1]$ independente das unidades de medidas de X e Y .

A autocovariância e a autocorrelação são medidas que possuem o mesmo propósito das estatísticas apresentadas anteriormente e são utilizadas em análises de séries temporais univariadas. Trata-se de analisar a relação de uma observação X_t com a observação do tempo anterior X_{t-1} de uma variável aleatória X_t , e então verificar

como as consecutivas observações estão relacionadas. Do mesmo modo, observações espaçadas de dois (X_t, X_{t-2}), três (X_t, X_{t-3}), ou mais períodos (X_t, X_{t-k}) também podem ter sua relação analisada. Assim, pode-se definir a função autocovariância e a função autocorrelação (ACF) respectivamente em função de k :

$$c_k = \frac{1}{n} \sum_{t=k+1}^n (X_t - \bar{X})(X_{t-k} - \bar{X}) \quad (2.15)$$

$$r_k = \frac{\sum_{t=k+1}^n (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \quad (2.16)$$

onde k denota o “lag” entre as observações, ou seja, o número de períodos entre as observações que terão sua relação analisada.

Através da análise das equações acima, podemos perceber nitidamente que em um processo estacionário as funções autocovariância e autocorrelação possuem as seguintes propriedades:

1. $c_0 = Var(X_t)$; $r_0 = 1$
2. $|c_k| \leq c_0$; $|r_k| \leq 1$
3. $c_k = c_{-k}$ e $r_k = r_{-k}$ para todo k , isto porque c_k e r_k são a mesma função e por isso são simétricas no “lag” $k = 0$ visto que o período entre X_t e X_{t+k} é o mesmo entre X_t e X_{t-k} , embora a função autocorrelação seja plotada apenas para “lags” não negativos.
4. c_k e r_k são semidefinidas positivas no sentido que:

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j c_{|t_i - t_j|} \geq 0 \quad (2.17)$$

e

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j r_{|t_i - t_j|} \geq 0 \quad (2.18)$$

para qualquer ponto t_1, t_2, \dots, t_n e qualquer número real $\alpha_1, \alpha_2, \dots, \alpha_n$. Pela definição de variável aleatória $X = \sum_{i=1}^n \alpha_i X_{t_i}$, a solução da inequação (2.17) segue de:

$$0 \leq Var(X) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j Cov(X_{t_i}, X_{t_j}) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j c_{|t_i - t_j|} \quad (2.19)$$

O resultado similar para r_k na inequação (2.18) segue da divisão da inequação (2.17) por c_o . Portanto, é importante saber que nem toda função arbitrária que satisfaça as propriedades de (1) a (3) podem ser função de autocovariância e autocorrelação para um processo. Uma condição necessária para uma função ser de autocovariância ou de autocorrelação de um processo é que a mesma seja semidefinida positiva (WEI, 2006).

Pode-se notar que ao se calcular as correlações de uma série temporal estacionária uni-variada o que tem-se como resposta é uma lista de números. Melhor do que analisar vários números seria analisar graficamente as autocorrelações da série em função dos “lags”. O gráfico das autocorrelações em função dos “lags” é conhecido como correlograma e é uma ferramenta muito importante que auxilia a visualização de forma clara e rapidamente a função autocorrelação.

O correlograma é essencial em análise de séries temporais antes da previsão. Através dessa, é possível verificar se existe sazonalidade, ciclo ou outros padrões na série analisada (MAKRIDAKIS,1998).

A Fig. (5) apresenta o correlograma da série de produção de bebidas. É possível ver claramente a presença de sazonalidade na série.

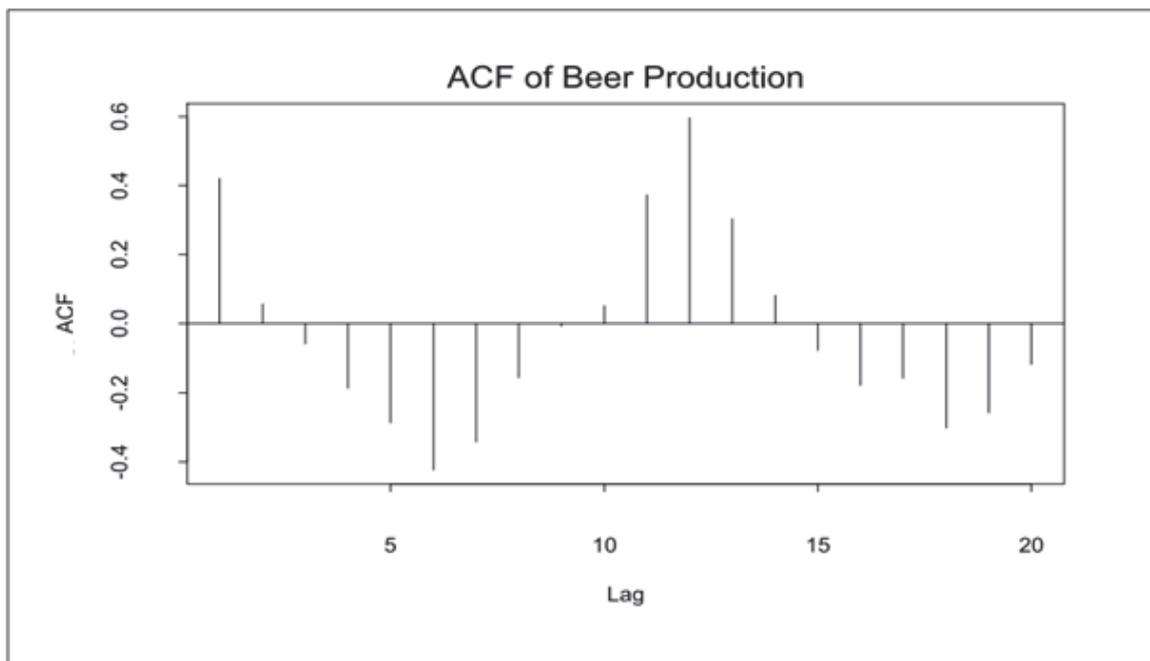


Figura 5: Correlograma (ou gráfico ACF) para os dados da produção de bebidas (MAKRIDAKIS,1998).

Complementando a função autocorrelação, para analisar a correlação existente entre X_t e X_{t+k} pode-se querer retirar os efeitos das variáveis $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$.

Esse tipo de análise é conhecida como função autocorrelação parcial (PACF) para séries estacionárias sendo X_t uma série normalmente distribuída. A correlação condicional pode ser representada por:

$$\phi_{kk} = Corr(X_t, X_{t+k} | X_{t+1}, \dots, X_{t+k-1}) \quad (2.20)$$

onde ϕ_{kk} é o coeficiente de correlação da distribuição de X_t, X_{t+k} condicional a $X_{t+1}, \dots, X_{t+k-1}$.

Para um dado k , mostra-se que ϕ_{kk} satisfaz a equação de Yule-Walker⁵:

$$r_j = \phi_{k1}r_{j-1} + \phi_{k2}r_{j-2} + \dots + \phi_{kk}r_{j-p}, \quad j = 1, 2, \dots, k \quad (2.21)$$

Mais explicitamente:

$$\begin{aligned} r_1 &= \phi_{k1} + \phi_{k2}r_1 + \dots + \phi_{kk}r_{j-1} \\ r_2 &= \phi_{k1}r_1 + \phi_{k2} + \dots + \phi_{kk}r_{k-2} \\ r_k &= \phi_{k1}r_{k-1} + \phi_{k2}r_{k-2} + \dots + \phi_{kk} \end{aligned} \quad (2.22)$$

Pode-se resolver essas equações para os “lags” $k = 1, 2, 3, \dots$, e obter ϕ_{kk} como mostrado a seguir:

$$\phi_{11} = r_1 \quad (2.23)$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & r \\ r & r_2 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 \\ r_1 & 1 \end{vmatrix}} = \frac{r_2 - r_1^2}{1 - r_1^2} \quad (2.24)$$

e para $k = 3$:

⁵Yule Walker é um dos métodos existentes para se estimar a ordem p e os coeficientes de um modelo Autoregressivo - AR(p) de uma série temporal discreta com média igual a zero. Esse tipo e outros modelos serão estudados nos capítulos a seguir.

$$\phi_{33} = \frac{\begin{vmatrix} 1 & r_1 & r_1 \\ r_1 & 1 & r_2 \\ r_2 & r_1 & r_3 \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{vmatrix}} \quad (2.25)$$

De uma forma geral, tem-se que:

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|} \quad (2.26)$$

onde P_k é a matriz de autocorrelação sendo seu determinante denotado por:

$$|P_k| = \begin{vmatrix} 1 & r_1 & r_2 & \cdots & r_{k-2} & r_{k-1} \\ r_1 & 1 & r_1 & \cdots & r_{k-3} & r_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{k-1} & r_{k-2} & r_{k-3} & \cdots & r_1 & 1 \end{vmatrix} \quad (2.27)$$

e P_k^* é a matriz P_k com a última coluna substituída pelo vetor de autocorrelação.

$$|P_k^*| = \begin{vmatrix} 1 & r_1 & r_2 & \cdots & r_{k-2} & r_1 \\ r_1 & 1 & r_1 & \cdots & r_{k-3} & r_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{k-1} & r_{k-2} & r_{k-3} & \cdots & r_1 & r_k \end{vmatrix} \quad (2.28)$$

Tanto a função autocorrelação (ACF) quanto a função autocorrelação parcial (PACF) têm um papel importante em análises de séries temporais. Essas duas funções, auxiliadas pelo correlograma, são utilizadas com na identificação da ordem dos modelos da metodologia Box-Jenkins. Essa análise será melhor apresentada na Seção 2.7.

2.5 Ruído Branco

Um processo a_t é chamado de ruído branco se a_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média e variância finitas. Em particular, se o processo a_t é normalmente distribuído com média zero e variância σ^2 , a série é conhecida como um ruído branco Gaussiano.

Para uma série de ruído branco, todas as ACF são zero. Na prática, se todas as ACF da amostra estão próximas de zero, então a série é uma série de ruído branco (TSAY, 2005).

Sendo assim, pela definição, pode-se dizer que o ruído branco é um processo estacionário e tem como função autocovariância:

$$c_k = \begin{cases} \sigma_a^2 & , k = 0 \\ 0 & , k \neq 0 \end{cases} \quad (2.29)$$

E função de autocorrelação:

$$r_k = \begin{cases} 1 & , k = 0 \\ 0 & , k \neq 0 \end{cases} \quad (2.30)$$

e função autocorrelação parcial:

$$\phi_{kk} = \begin{cases} 1 & , k = 0 \\ 0 & , k \neq 0 \end{cases} \quad (2.31)$$

O ruído branco é a confirmação do bom ajuste do modelo. Para se ter como resultado um modelo que faça previsões satisfatórias é necessário que a modelagem tenha como resíduo um ruído branco, ou seja, o erro do modelo seja uma variável aleatória independente e identicamente distribuída, com distribuição normal, média zero e variância σ^2 (SÁFADI e NOGUEIRA, 2000).

2.6 Análise de Regressão

A análise de regressão é uma técnica estatística utilizada para modelar e investigar as relações entre uma variável de resposta, conhecida como variável dependente, e uma ou mais variáveis de regressão, ou regressoras. O objetivo de uma análise de regressão, geralmente, é construir um modelo que seja capaz de prever valores futuros de uma variável de resposta, através de específicos valores de variáveis preditoras (MONTGOMERY et al., 2008).

Um modelo de regressão linear simples envolve apenas uma variável preditora, e é escrito como:

$$X_t = \beta_0 + \beta_1 Z_t + \varepsilon_t \quad (2.32)$$

onde X é a resposta, Z é variável independente, β_0 e β_1 são parâmetros desconhecidos, e ε é o termo erro.

Os modelos de regressão, geralmente incluem mais do que uma variável regressora. Sendo X_t para $t = 1, 2, \dots, n$ uma variável dependente, que sofre influência de conhecidas séries independentes, $Z_{t1}, Z_{t2}, \dots, Z_{tq}$, o modelo de regressão linear múltiplo que expressa essa relação é:

$$X_t = \beta_0 + \beta_1 Z_{t1} + \beta_2 Z_{t2} + \dots + \beta_q Z_{tq} + \varepsilon_t \quad (2.33)$$

onde $\beta_0, \beta_1, \beta_2, \dots, \beta_q$ denotam coeficientes fixos de regressão e ε_t o erro aleatório, sendo esse um ruído branco.

Os coeficientes de regressão representam os efeitos marginais dos regressores separadamente. Por exemplo, o coeficiente β_2 representa o impacto na média da variável dependente quando o segundo regressor aumenta de uma unidade enquanto os outros regressores permanecem constantes. Na linguagem matemática, essa relação pode ser expressa como $\frac{\partial X_t}{\partial Z_{t2}} = \beta_2$. O termo erro representa a parte da variável dependente que não pode ser explicada pelos regressores (HAYASHI, 2000).

Uma aplicação interessante da técnica de regressão linear para se obter um modelo de previsão é apresentado por Byron&Ashenfelter (1995) que através de uma regressão simples, um modelo para prever a qualidade do vinho Grange é construído. No entanto, é muito mais comum utilizar a modelagem de regressão para explicar variações históricas do que para fins de previsão (HYNDMAN, 2009) .

As Eqs. (2.30) e (2.31) são modelos de regressão lineares, pois são lineares nos parâmetros desconhecidos (β 's) e não porque eles descrevem necessariamente relações lineares entre a resposta e os regressores. Por exemplo, o modelo:

$$X_t = \beta_0 + \beta_1 Z_{t1} + \beta_2 Z_{t2}^2 + \varepsilon_t \quad (2.34)$$

é um modelo de regressão linear porque ele é linear nos parâmetros β_0, β_1 e β_2 , embora ele descreva uma relação quadrática entre X_t e Z_{t2} .

Outra forma de se representar um modelo de regressão linear, é utilizando a notação matricial. Seja Z_t e β vetores de dimensão n definidos por:

$$\begin{matrix} Z_t \\ (n \times 1) \end{matrix} = \begin{bmatrix} Z_{t1} \\ Z_{t2} \\ \vdots \\ Z_{tn} \end{bmatrix}, \quad \begin{matrix} \beta \\ (n \times 1) \end{matrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \quad (2.35)$$

Pela definição do produto interno⁶ entre vetores, $Z_t' \beta = \beta_1 Z_{t1} + \beta_2 Z_{t2} + \dots + \beta_n Z_{tn}$.

Desse modo, a Eq. (2.33) pode ser escrita como:

$$X_t = Z_t' \beta + \varepsilon_t \quad (2.36)$$

sendo:

$$\begin{matrix} X \\ (n \times 1) \end{matrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \quad \begin{matrix} \varepsilon \\ (n \times 1) \end{matrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \quad \begin{matrix} Z \\ (n \times k) \end{matrix} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} = \begin{bmatrix} Z_{11} & \cdots & Z_{1k} \\ \vdots & \ddots & \vdots \\ Z_{1n} & \cdots & Z_{nk} \end{bmatrix} \quad (2.37)$$

Nas matrizes e vetores o número de linhas na Eq. (2.37) é o mesmo do número de observações das séries.

Os modelos de regressão são muito utilizados em modelagem de séries temporais, mas nesse contexto, os regressores incluem os “lags” da variável dependente. Um simples modelo genérico que ilustra essa aplicação pode ser descrito como abaixo:

$$X_t = \beta_0 + \beta X_{t-1} + \varepsilon_t \quad (t = 1, 2, \dots, n) \quad (2.38)$$

Esse tipo de modelo é conhecido como Autoregressivo de primeira ordem (AR(1)), onde a X_{t-1} denota o valor da variável X em $t-1$. O modelo AR entre outros modelos que têm aplicação em séries temporais serão abordados mais profundamente nos próximos itens, assim como algumas técnicas para fazer a estimação dos parâmetros desconhecidos.

⁶Produto interno é uma função de dois vetores que satisfaz os seguintes axiomas:

- (a) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, onde a barra denota conjugação complexa
- (b) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ para todo $x, y, z \in \mathbb{R}$
- (c) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle + \langle y, z \rangle$ para todo $x, y, z \in \mathbb{R}$ e $\alpha \in \mathbb{C}$
- (d) $\langle x, x \rangle \geq 0$ para todo $x \in \mathbb{R}$
- (e) $\langle x, x \rangle = 0$ se e somente se $x = 0$, onde x, y e z são vetores de \mathbb{R} , \mathbb{R} um espaço vetorial sobre um corpo \mathbb{C} .

2.7 Modelos de Previsão Box-Jenkins

Como já foi dito anteriormente, existem dois tipos de modelos para previsão. Modelos qualitativos, que geralmente são subjetivos e necessitam da opinião de especialistas e os modelos quantitativos, ou, estatísticos.

A suposição geral para os modelos quantitativos é que quaisquer dados de séries temporais podem ser representados como a soma de duas distintas componentes: determinísticas e estocásticas (aleatórias). A primeira componente corresponde a modelagem do processo através de uma função de tempo e a última a existência de um ruído aleatório adicionado ao sinal determinístico gerando o comportamento estocástico das séries temporais (MONTGOMERY et al., 2008).

Geralmente, sucessivas observações mostram uma dependência serial. Nestas circunstâncias, alguns métodos de modelagem podem ser ineficientes e inadequados, por não tirar proveito dessa dependência de forma mais eficaz.

Uma das técnicas quantitativas mais difundidas é a metodologia de Box-Jenkins, descrita por esses autores na década de 70. Os modelos de Box-Jenkins têm como pressupostos que cada valor da série (temporal) pode ser explicado por valores históricos, a partir do uso da estrutura de correlação temporal que geralmente há entre os valores da série (WENER e RIBEIRO, 2003).

O procedimento original para aplicação da metodologia Box-Jenkins envolve três fases iterativas: a seleção do modelo, a estimação dos parâmetros e por fim, a verificação do modelo. Alguns trabalhos (e.g. Makridakis et al., 1998) adicionam uma fase preliminar, a de preparação dos dados e uma fase final, a de aplicação do modelo (ou de previsão) (HYNDMAN, 2001).

1. **Preparação dos dados** – A fase de preparação dos dados envolve o uso de transformações e diferenciações com a finalidade de estabilizar a variância em séries em que essa muda com a média e remover padrões óbvios tais como tendência ou sazonalidade.
2. **Identificação ou seleção do modelo** – A seleção do modelo consiste em descobrir qual dentre as várias versões dos modelos de Box-Jenkins, descreve o comportamento da série. Essa etapa evolve a análise do comportamento das funções de autocorrelações (ACF) e das funções de autocorrelações parciais (PACF).
3. **Estimação de parâmetros** – A estimação de parâmetros implica em encontrar

os valores dos coeficientes do modelo que proporcionar o melhor ajuste para os dados.

4. **Verificação** – A fase de verificação do modelo envolve testar hipóteses para avaliar a adequação do modelo na descrição do comportamento da série. Se o modelo for considerado inadequado, é necessário voltar ao passo 2 e tentar identificar o melhor modelo.
5. **Previsão** - Uma vez que o modelo foi selecionado, estimado e verificado, geralmente, o próximo passo é fazer previsões para os valores futuros da série.

A Fig. 6 representa esquematicamente como se dá a aplicação da metodologia de Box-Jenkins. O detalhamento de todas as etapas acima se dará com a aplicação do procedimento em um caso real no Capítulo 4.

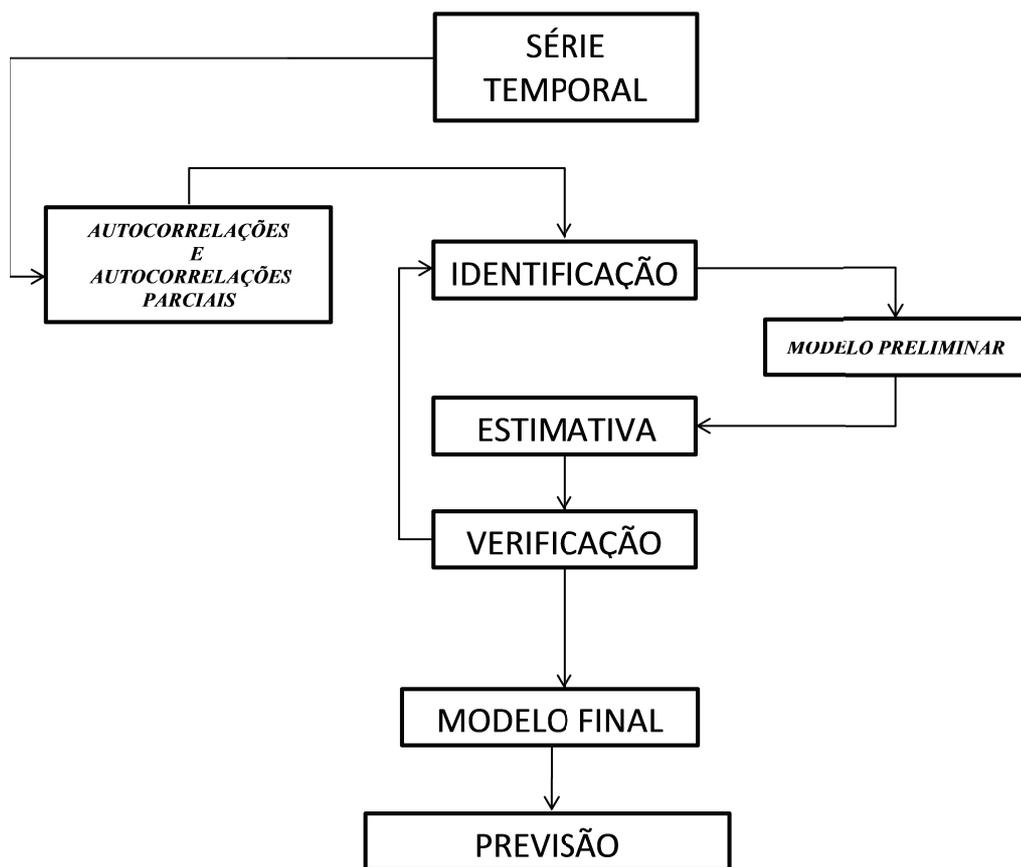


Figura 6: Estágios do ciclo iterativo para os modelos Box-Jenkins. (Adaptado: CASTRO,1995)

2.7.1 Modelos de Previsão Lineares e Estacionários

2.7.1.1 Processos Autoregressivos (AR)

Como já dito anteriormente, os modelos Autoregressivos se baseiam na idéia de que valores atuais de uma série X_t podem ser explicados em função dos valores passados da mesma série $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, onde p denota o número máximo de valores passados necessários para se prever os valores atuais. Em outras palavras, um processo Autoregressivo nada mais é do que uma equação de diferenças linear com coeficientes constantes, adicionada de um ruído branco.

Yule (1927) utilizou o processo AR para descrever o fenômeno do número de manchas solares e o comportamento de um simples pêndulo.

Uma forma de se entender um processo Autoregressivo é através de uma expansão em série de Taylor. O teorema de Taylor⁷ diz que é possível obter uma aproximação de uma função $f(z)$ k vezes diferenciável em torno de um dado ponto a , através de um polinômio de k th ordem. Pensando assim, a primeira derivada seria a diferença entre X_t e X_{t-1} . A segunda derivada dependerá dos três últimos pontos, a terceira derivada dos quatro últimos pontos, e assim por diante. A expansão em série de Taylor resultante seria uma estimativa da variável X_t em termos de uma função linear dos pontos anteriores. Tentando provar que para prever cada ponto futuro de uma série através de uma combinação linear de pontos anteriores da mesma série é o mesmo que expandir a função em série de Taylor, propõe-se uma demonstração com a função $\sin(x)$.

Adotando $a = 0$ e sabendo que:

$$\begin{aligned} f(x) &= \sin x \rightarrow f(0) = 0 \\ f'(x) &= \cos x \rightarrow f(0) = 1 \\ f''(x) &= -\sin x \rightarrow f(0) = 0 \\ f'''(x) &= -\cos x \rightarrow f(0) = -1 \\ f^{IV}(x) &= \sin x \rightarrow f(0) = 0 \end{aligned}$$

Assim tem-se que a expansão em série de Taylor é:

⁷O Polinômio de Taylor é definido por: $f(z) = f(a) + f'(a)(z-a) + f''(a)\frac{(z-a)^2}{2!} + \dots + f^k(a)\frac{(z-a)^k}{k!}$

$$\text{sen}(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \frac{\text{sen}\left(\frac{k\pi}{2}\right)}{k!}x^k \quad (2.39)$$

Para a mesma série, foi estimado um modelo AR(2) e as curvas resultantes para os dois modelos foram plotadas. Nota-se que as curvas apresentam comportamento semelhante.

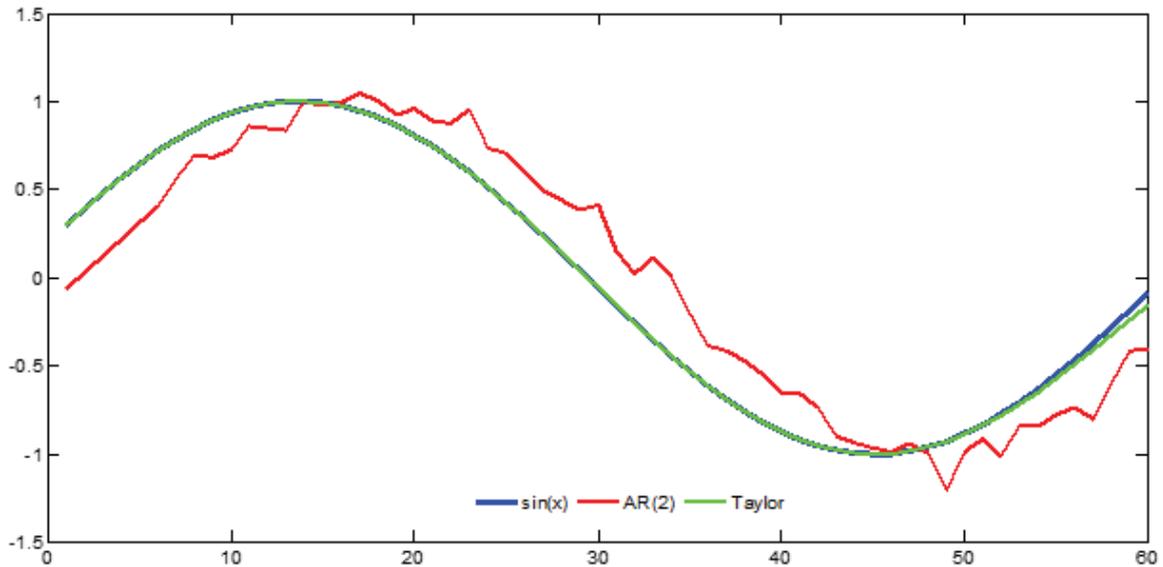


Figura 7: Comparação entre as curvas $\sin(x)$, expansão de Taylor para $\sin(x)$ e AR(2) para $\sin(x)$

O modelo mais simples para esse tipo de processo é o AR(1), descrito pela Eq. (2.38), nesse tipo de modelo o valor atual de uma série depende apenas do primeiro valor anterior. O processo AR de primeira ordem é considerado um processo de Markov. E ainda, se o erro for aleatório e identicamente distribuído, o processo é conhecido como Markov Gaussiano. (Castro, 1995)

Se um número finito de π ponderações são não nulas, isto é, $\pi_1 = \beta_1, \pi_2 = \beta_2, \dots, \pi_p = \beta_p$ e $\pi_k = 0$ para $k > p$, um processo Autoregressivo de ordem p , AR(p), pode ser descrito através da equação (WEI, 2006):

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_p X_{t-p} + \dots + \varepsilon_t \quad (t = 1, 2, 3, \dots, p) \quad (2.40)$$

onde β_0 é um parâmetro com valor constante definido por:

$$\beta_0 = \{1 - (\beta_1 + \beta_2 + \dots + \beta_p)\}\mu \quad (2.41)$$

Além disso, para um processo estacionário, $1 - (\beta_1 + \beta_2 + \dots + \beta_p) > 0$, com isso,

$\mu = 0$ se e somente se, $\beta_0 = 0$ (RUPPERT, 2011). De uma forma diferente, uma condição necessária para a estacionariedade de uma série pode ser representada por $\beta_1 + \beta_2 + \dots + \beta_p < 1$.

Outra forma de representar um modelo AR(p) pode ser dada por (MONTGOMERY, 2008):

$$\Phi(B)X_t = \beta_0 + \varepsilon_t \quad (2.42)$$

onde $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p$.

E, para que a Eq. (2.42) possa ser considerada estável, as raízes de $\Phi(B) = 0$, ou seja, $1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p = 0$ devem se encontrar fora do círculo unitário. Quando essa condição for satisfeita o processo Autoregressivo pode ser representado por um processo de médias móveis de ordem infinita (POLLOCK, 1999).

Para um processo AR(p) estacionário, a média $E(X_t)$ pode ser estimada através da seguinte equação:

$$E(X_t) = \frac{\beta_0}{(1 - \beta_1 - \beta_2 - \dots - \beta_p)} = \mu \quad (2.43)$$

A autocovariância é definida por:

$$c_j = \begin{cases} \beta_1 c_{j-1} + \beta_2 c_{j-2} + \dots + \beta_p c_{j-p} & \text{para } j = 1, 2, \dots \\ \beta_1 c_1 + \beta_2 c_2 + \dots + \beta_p c_p + \sigma^2 & \text{para } j = 0 \end{cases} \quad (2.44)$$

Dividindo a Eq. (2.38) por c_0 , obtém-se uma equação de Yule-Walker:

$$r_j = \frac{c_j}{c_0} = \beta_1 r_{j-1} + \beta_2 r_{j-2} + \dots + \beta_p r_{j-p} \quad (2.45)$$

As funções autocorrelação e autocorrelação parcial são excelentes para se identificar a ordem dos modelos AR(p), principalmente, a função autocorrelação parcial. Essas funções apresentam padrões que tornam possível a identificação do modelo em questão. Como se pode notar, a PACF para modelos Autoregressivos, é truncada no “lag” que corresponde à ordem do modelo. Já a ACF apresenta um padrão de declínio exponencial misturado a uma senóide amortecida.

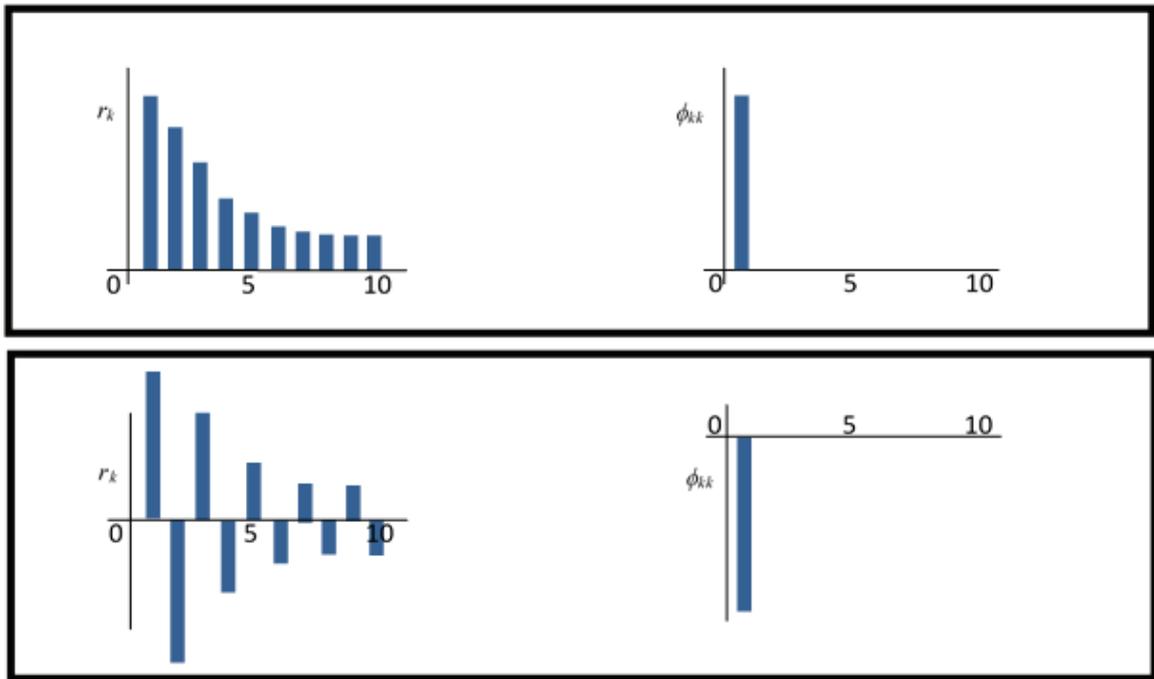


Figura 8: - Funções de autocorrelação ACF (à esquerda) e PACF (à direita) de um modelo AR(1).

2.7.1.2 Processos Média Móveis (MA)

Seja ε_t um ruído branco, e o processo representado abaixo:

$$X_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1} \quad (2.46)$$

onde μ e θ são constantes. Esse tipo de série temporal é chamada de processo de médias móveis de primeira ordem, denotado por MA(1). O termo “médias móveis” vem do fato de X_t ser construído a partir de uma soma ponderada, semelhante a uma média, de dois ou mais valores recentes de ε (HAMILTON, 1994).

De uma forma geral, a representação para um processo médias móveis de ordem q , ou seja, MA(q) é dada por:

$$X_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (2.47)$$

A média e a variância para esse tipo de modelo, podem ser denotadas respectivamente por:

$$E(X_t) = E(\mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}) = \mu \quad (2.48)$$

$$\begin{aligned}
E(X_t - \mu)^2 &= c_0 = \text{Var}(\mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}) \\
&= \sigma^2(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)
\end{aligned} \tag{2.49}$$

Similarmente, a função autocovariância é definida por:

$$\begin{aligned}
c_k &= \text{Cov}(X_t, X_{t+k}) \\
&= E(\varepsilon_t - \theta_1\varepsilon_{t-1} - \dots - \theta_q\varepsilon_{t-q})(\varepsilon_{t+k} - \theta_1\varepsilon_{t+k-1} - \dots - \theta_q\varepsilon_{t+k-q}) \\
&= \begin{cases} \sigma^2(-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q) & , k = 1, 2, \dots, q \\ 0 & , k > q \end{cases}
\end{aligned} \tag{2.50}$$

Assim, a função autocorrelação de um processo MA(q), para o “lag” k é definida por:

$$r_k = \frac{c_k}{c_0} = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q}{(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)} & , k = 1, 2, \dots, q \\ 0 & , k > q \end{cases} \tag{2.51}$$

Esta característica da ACF é muito útil na identificação do modelo MA e sua ordem. Na prática, porém, ACF amostral, não será necessariamente igual a zero depois do “lag” q , mas espera-se um valor absoluto muito próximo de zero. Em geral, a PACF de um processo Médias Móveis de ordem q , MA(q), apresenta um padrão de decaimento exponencial misturado com uma senóide amortecida.

2.7.1.3 Processos Autoregressivos de Médias Móveis (ARMA)

Dentre os modelos de processos estocásticos estacionários mais simples e mais utilizados estão os modelos Autoregressivos e Médias Móveis ou ARMA. Partindo do que já foi visto, pode-se definir um modelo AR como nada mais do que uma equação de diferenças linear com coeficientes constantes na qual a chamada “*forcing function*”⁸ é um processo de ruído branco. Outra definição pertinente é a de que um modelo MA expressa uma sequência de observações estocásticas como uma combinação linear do valor atual e de um número finito de valores defasados de um processo de

⁸Em um sistema de equações diferenciais utilizado para descrever uma série temporal, a chamada “*forcing function*” é uma função que aparece nas equações e é apenas dependente da variável tempo, e não de qualquer outras variáveis

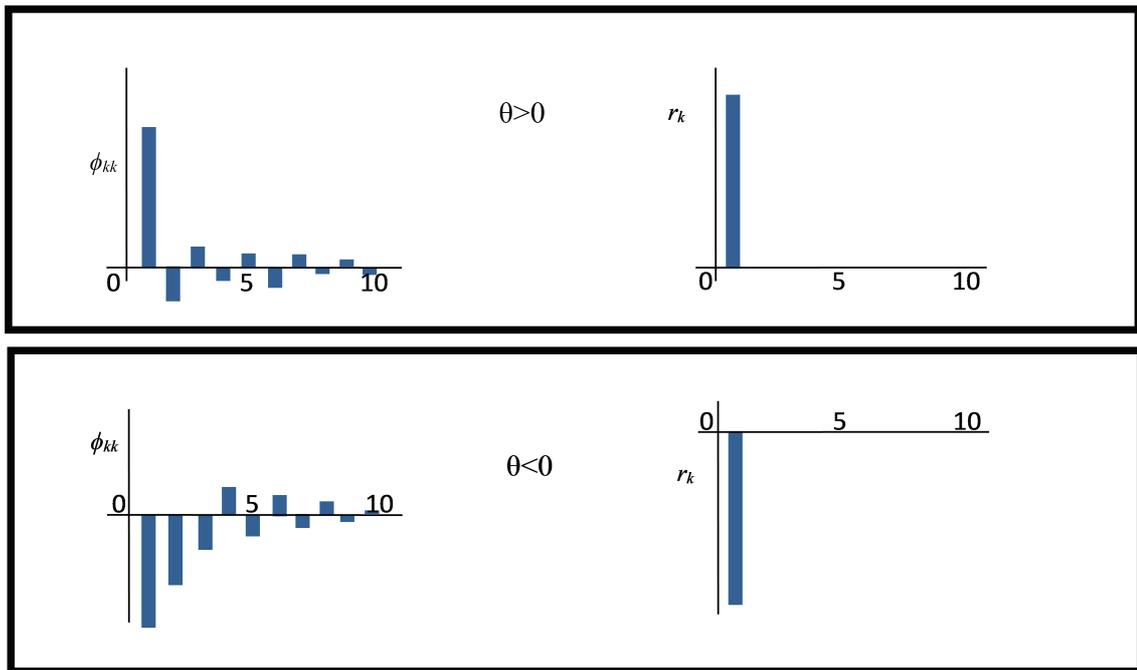


Figura 9: Funções de autocorrelação FAC (à direita) e FACP (à esquerda) de um modelo MA(1)

ruído branco. Um modelo ARMA é uma equação de diferença estocástica em que a chamada “*forcing function*” é um processo de média móvel (POLLOCK, 1999).

Um processo estacionário e inversível⁹ pode ser representado tanto através de um modelo AR como na forma de um processo de MA. Mas um problema, decorrente da utilização desses tipos de modelos, é que essas representações, muitas vezes, podem conter muitos parâmetros devido à necessidade de se criar modelos de altas ordens com o intuito de se obter uma boa aproximação (WEI, 2006).

Por esse motivo, a utilização dos modelos ARMA(p,q), onde p denota a ordem do modelo AR e q a ordem do modelo MA, é realizada em ampla escala. Através desse “*mix*” dos modelos AR e MA consegue-se ter como resultado uma boa aproximação com um menor número de parâmetros do que se fosse utilizado um modelo ou outro.

De uma forma geral, um modelo ARMA(p,q) é definido por:

$$X_t = \mu + \varepsilon_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_p X_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2.52)$$

⁹Para um processo ser inversível é necessário que as raízes de $\theta_q(B) = 0$ estejam fora do círculo unitário. Para um processo ser considerado estacionário é necessário que as raízes de $\beta_p(B) = 0$ estejam fora do círculo unitário. (WEI, 2006)

onde os β 's denotam os coeficientes de um modelo AR, θ 's denotam os coeficientes de um modelo MA, μ uma constante e ε_t o erro aleatório.

A forma mais simples para esse tipo de modelo é o ARMA(1,1), onde ambas ordens dos processos envolvidos, ou seja, AR e MA, são iguais a 1. Esse tipo de processo é definido por:

$$X_t = \mu + \varepsilon_t + \beta_1 X_{t-1} - \theta_1 \varepsilon_{t-1} \quad (2.53)$$

A média de um processo ARMA(p,q) depende apenas da parte auto-regressiva do processo, desse modo, essa pode ser encontrada através da Eq. (2.41).

Assim como a média, a estacionariedade do processo Auto-regressivo e de Médias Móveis – ARMA(p,q) - depende inteiramente dos parâmetros Autoregressivos ($\beta_1, \beta_2, \dots, \beta_p$) e é completamente independente dos parâmetros do processo de Médias Móveis (HAMILTON, 2004).

As funções ACF e PACF de um modelo ARMA(p,q) são determinadas pelos componentes AR e MA respectivamente. Isso pode ser visto, pelo fato de ambas funções apresentarem o mesmo comportamento, o que faz com que a identificação da ordem do modelo seja relativamente mais difícil. Para contornar esse problema, existem alguns recursos que podem ser utilizados, esses não serão abordados nesse trabalho, para mais informações pode-se consultar (WEI, 2006) e (MONTGOMERY, 2008).

A Tabela 1 apresenta um resumo dos comportamentos dos modelos estudados até então. Apesar das funções ACF e PACF apresentarem propriedades fortemente ligadas respectivamente aos modelos MA(q) e AR(p) é recomendável que se utilize as duas funções na identificação da ordem de modelos.

Tabela 1: Comportamento das funções ACF e PACF para os modelos AR(p), MA(q) e ARMA(p,q)

Modelo	ACF	PACF
AR(p)	Decaimento exponencial e/ou senóide amortecida	Truncada no lag p
MA(q)	Truncada no lag q	Decaimento exponencial e/ou senóide amortecida
ARMA(p,q)	Decaimento exponencial e/ou senóide amortecida	Decaimento exponencial e/ou senóide amortecida

2.7.2 Modelos de Previsão Lineares Não Estacionários

2.7.2.1 Processos Autoregressivos Integrável Médias Móveis (ARIMA)

Os modelos discutidos até então são utilizados para processos estacionários, mas na prática, a maioria das séries temporais são não estacionárias.

Os modelos ARIMA, ou como são conhecidos, modelos de Box-Jenkins, são apropriados para a modelagem das séries não estacionárias homogêneas¹⁰.

Uma série não estacionária pode ser reduzida em uma série estacionária através de uma diferenciação de grau apropriado. O modelo ARIMA(p,d,q) nada mais é, do que o modelo ARMA com uma parte integrável. Esta parte do modelo (d) determina se os valores observados serão modelados diretamente, ou se deverá ser calculado as diferenças entre observações consecutivas vezes com o intuito de tornar a série estacionária. Se $d = 0$, as observações são modeladas diretamente. Se $d = 1$, as diferenças entre observações consecutivas serão modeladas. Se $d = 2$, as diferenças entre as diferenças serão modelados, e assim por diante. Na prática, d é raramente maior que 2 (HYNDMAN, 2001).

2.7.2.2 Modelo Sazonal Autoregressivos Integrado Médias Móveis (SARIMA)

Existem também, os processos ARIMA projetados para lidar com séries temporais sazonais. Esse tipo de processo é conhecido como SARIMA.

Os modelos SARIMA contêm uma parte não sazonal com parâmetros (p,d,q) e uma sazonal com parâmetros (P,D,Q)_s, onde:

p - é o número de parâmetros Autoregressivos não sazonal.

d - é o número de diferenças não sazonal necessário para tornar a série estacionária

q - é o número de parâmetros médias móveis não sazonal.

P - é o número de parâmetros Autoregressivos sazonal

D - é o número de diferenças sazonal necessário para tornar a série estacionária

Q - é o número de parâmetros médias móveis sazonal.

s - é o tamanho do período sazonal

¹⁰Séries não estacionárias homogêneas são séries não estacionárias na média ou no nível podendo ser transformadas em estacionárias através da diferenciação sucessiva.

Desse modo, de uma maneira geral, pode-se escrever esse tipo de modelo na forma:

$$\Delta X_t = \mu + \varepsilon_t + \sum_{i=1}^p \beta_i \Delta_{t-i}^d + \sum_{t=1}^p \Lambda_i^P \Delta_{t-i}^D + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{t=1}^q \lambda_i^Q \varepsilon_{t-i} \quad (2.54)$$

onde Δ_{t-i}^d é o operador de diferenças não sazonal, Δ_{t-i}^D o operador de diferenças sazonal e λ e Λ parâmetros do modelo a ser estimados.

Todos os modelos apresentados até então são modelos lineares e não consideram a inconstância da variância das séries. Existem modelos, conhecidos como modelos heterocedásticos nos quais o comportamento da variância condicional é incluído na construção do modelo. Esses modelos não serão abordados nesse trabalho, para mais informações ver Engle (1982) e Bollerslev (1986).

3 *Técnicas Estatísticas*

São inúmeras as técnicas estatísticas que auxiliam a construção de modelos. Cada uma tem sua devida aplicação e necessidade. Este capítulo aborda algumas ferramentas estatísticas importantes que foram utilizadas durante o estudo.

3.1 *Análise das Componentes Principais*

Antes de se estimar os parâmetros de um modelo, precisa-se analisar todas as informações disponíveis a respeito do processo que se deseja modelar para então se traçar estratégias para se encontrar o modelo mais apropriado. A qualidade das análises depende integralmente da qualidade dos dados.

Foi visto no capítulo anterior que algumas características comuns em séries temporais, se não tratadas, podem atrapalhar nas análises dos dados, porém, pode-se usufruir de tal comportamento para melhor analisar o processo. Um exemplo disso é a análise das componentes principais.

É comum encontrar séries reais que estão correlacionadas com outras séries, ou seja, um processo poderia explicar o outro se bem aproveitada essa idéia. A técnica da análise das componentes principais é um método estatístico multivariado que permite interpretar a estrutura de um conjunto de dados multivariados a partir das respectivas matrizes de variâncias-covariâncias ou de correlações entre os parâmetros considerados (BASTOS et al., 2009). Consiste numa transformação linear de todas as variáveis originais em novas variáveis, de tal modo que a primeira nova variável represente a maior parte da variação possível existente no conjunto de dados, a segunda pela maior variação possível restante, e assim por diante até que toda a variação do conjunto tenha sido explicada. O objetivo principal da PCA é explicar a estrutura de variância e covariância de um vetor aleatório, através da construção de combinações lineares das variáveis originais, que são chamadas de componentes principais (PC) não correlacionadas entre si.

Seja Z um vetor composto por duas séries correlacionadas entre si. Pode-se encontrar um a série Z' que contém a maior parte das informações através da Eq. (3.1)

$$Z' = Z * V \quad (3.1)$$

Sendo V a matriz de autovetores da matriz de covariância de Z .

Assim, as informações mais importantes e relevantes dos dados complexos podem ser visualizadas de maneira mais simples, facilitando a análise dos dados e a especificação do modelo a ser utilizado para partir para a fase de estimação dos parâmetros.

3.2 Método Mínimos Quadrados Ordinários

O método dos mínimos quadrados ordinários se baseia na estimação dos parâmetros de um modelo, de modo que a soma dos quadrados dos desvios entre valores estimados e valores observados da amostra (ou seja, os resíduos) sejam mínimos.

Para melhor interpretação, considere que Y_t é uma determinada resposta que pode ser representada pela Eq. (3.2):

$$Y_i = \mu + \varepsilon_i + \beta_1 X_i \quad (3.2)$$

sendo:

ε_t é o ruído branco

μ é a interseção a se estimada

β_1 é a inclinação do modelo a ser estimada

A fim de ajustar o modelo representado pela Eq. (3.2), utilizando o princípio dos mínimos quadrados, podemos estimar os parâmetros μ e β_1 de modo que minimizem o somatório do quadrado residual.

A Eq. (3.2) pode ser escrita de tal forma que é possível determinar o valor dos desvios ε_i , ou seja, as diferenças entre os valores observados e os valores esperados, assim como na e Eq. (3.3).

$$\varepsilon_i = Y_i - \mu + \beta_1 X_i \quad (3.3)$$

Podendo esta ser escrita como:

$$\varepsilon_i = Y_i - \hat{Y}_i \quad (3.4)$$

Como esta diferença pode ser tanto positiva quanto negativa, o somatório das diferenças pode ocasionar um resultado nulo. Poderia se pensar em utilizar o módulo das diferenças, mas ao se tentar minimizá-lo, aplicando as derivadas primeiras, poderia complicar a solução dos cálculos. Por isso, uma forma eficiente de se evitar o cancelamento e a complicação dos cálculos é minimizar o quadrado da diferença. Assim, elevando-se ao quadrado os desvios e aplicando-se o somatório, tem-se que:

$$S = \sum_{i=1}^n (Y_i - \mu + \beta_1 X_i)^2 \quad (3.5)$$

Enfim de acordo com o método de Mínimos Quadrados Ordinários, os coeficientes μ_1 e β_1 são valores estimados, que minimizam o critério S para uma determinada amostra.

3.3 Método da Máxima Verossimilhança

Assim como o método dos mínimos quadrados, a máxima verossimilhança é utilizada para se estimar parâmetros de um determinado modelo. Porém, como o crescente uso de modelos estatísticos cada vez mais complexos alavancou uma maior dependência de métodos baseados em máxima verossimilhança para se estimar e testar hipóteses (ENGLE, 1982).

De uma maneira geral, dada uma determinada amostra e um modelo estatístico, o método de máxima verossimilhança estima os valores dos diferentes parâmetros do modelo estatístico de maneira a maximizar a probabilidade de se obter os dados observados, ou seja, maximizar a função de verossimilhança.

Seja y uma variável aleatória com distribuição normal, média μ e variância σ^2 . A função de densidade de probabilidade de cada observação é também normal e dada por:

$$f(y_t; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_t - \mu)^2\right) \quad (3.6)$$

e a função densidade conjunta é dada por:

$$\prod_{t=1}^T f(y_t; \mu, \sigma^2) \quad (3.7)$$

Logo, a função de verossimilhança é:

$$L = \prod_{t=1}^T f(\mu, \sigma^2; y_t) \quad (3.8)$$

e a forma mais comum de apresentação da função verossimilhança por simplificar os cálculos durante a maximização e obter o mesmo resultado que a função original, é o logaritmo natural de L que é definido por:

$$\ln L(\mu, \sigma^2; y_t) = \sum_{t=1}^T \ln f(\mu, \sigma^2; y_t) \quad (3.9)$$

$$= \sum_{t=1}^T \left[\ln \sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \mu)^2 \right] \quad (3.10)$$

$$\ln L = -\frac{T}{2} \ln 2\pi - \frac{T}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \mu)^2 \quad (3.11)$$

Estimando os parâmetros desconhecidos (μ, σ^2) que maximize o logaritmo natural da função máxima verossimilhança, maximiza-se a probabilidade de se obter a amostra específica que se deseja trabalhar.

3.4 Teste de Hipótese

Modelos que consigam representar o comportamento de um sistema bem próximo do real com um número mínimo de coeficientes é o principal objetivo dos especialistas em séries temporais. O teste de hipóteses é essencial para esse tipo de análise.

Um teste de hipótese é um método de tomada de decisões que utiliza dados de um experimento controlado ou um estudo observacional (não controlados). Em estatística,

o resultado é chamado estatisticamente significativo se é pouco provável que tenha ocorrido por acaso. O nível de significância é definido por uma probabilidade pré-determinada, essa é o limiar entre a significância ou não da hipótese testada.

Testes de hipóteses são algumas vezes chamados de análise de dados de confirmação, em contraste com a análise exploratória de dados.

Todos os testes de hipóteses são conduzidos do mesmo modo. Primeiramente, se determina a hipótese a ser testada e se formula um plano de análise. Os dados são analisados de acordo com o planejamento estipulado, para então se aceitar ou rejeitar a hipótese nula baseando-se nos resultados da análise da amostra.

Todo e qualquer teste de hipótese requer que uma hipótese nula e uma alternativa seja definida. Estas devem ser definidas de tal forma que uma seja verdadeira e a outra seja exclusivamente falsa. Geralmente, as hipóteses são representadas por H_0 e H_1 , sendo que:

H_0 = Hipótese nula - Hipótese contrária ao efeito que se deseja testar.

H_1 = Hipótese Alternativa – Hipótese que se deseja testar.

Como falado anteriormente, após a determinação das hipóteses deve-se elaborar um plano de análise. Esse plano basicamente define os critérios para a utilização da amostra para definir se deve rejeitar ou aceitar a hipótese nula. É nessa etapa que é definido o nível de significância, o qual geralmente é escolhido igual a 1%, 5% e 10%, apesar de que qualquer valor entre 0 e 100% pode ser usado.

Após determinado o nível de significância, a região crítica e a estatística teste mais apropriada deverão ser definidos. Denomina-se região crítica a área de rejeição da hipótese nula, ela compreende uma área igual ao valor de α .

A estatística teste é o escore padronizado da amostra cujo valor crítico determina o limite ou limites (no caso de um teste bilateral) entre a região de aceitação e a região de rejeição para H_0 . Tal estatística pode ser z, t, chi-quadrado, F, etc., sendo que a escolha da estatística mais apropriada depende da distribuição da amostra a ser utilizada. O valor crítico da estatística teste pode ser lido a partir das tabelas estatísticas.

Dada uma determinada estatística teste e sua distribuição amostral, pode-se encontrar as probabilidades associadas a estatística de teste. Se a probabilidade da estatística teste for menor que o nível de significância, a hipótese nula é rejeitada. Geralmente, quando essa análise é feita, envolve a comparação do chamado P-value

para o nível de significância. O P-value nada mais é que a probabilidade de observar um teste estatístico tão extremo quanto a estatística teste.

O teste de hipótese pode ser utilizado para avaliar uma média, proporção, diferença entre médias, diferença entre proporções, etc.

Geralmente, quando o teste possui hipótese nula que envolve uma média ou uma proporção, é utilizado uma das seguintes equações para o teste estatístico:

$$\text{Teste estatístico} = (\text{Estatística} - \text{Parâmetro}) / (\text{Desvio padrão da estatística})$$

$$\text{Teste estatístico} = (\text{Estatística} - \text{Parâmetro}) / (\text{Erro padrão da estatística})$$

Como é possível obter apenas dois resultados diferentes ao se fazer um teste de hipótese, existem apenas dois tipos de erros que podem ser cometidos, os erros conhecidos como do Tipo I e Tipo II.

Quando se decide aceitar a hipótese nula, pode-se estar tomando a decisão correta, ou seja, aceitar a hipótese nula sendo ela verdadeira, ou tal decisão pode estar incorreta. A decisão incorreta é a de aceitar uma hipótese nula falsa. Esta decisão é um exemplo de um erro do Tipo II sendo β a probabilidade de se cometê-lo.

Por sua vez, a decisão de se rejeitar a hipótese nula também pode estar correta ou incorreta. É possível estar equivocado e rejeitar uma determinada hipótese nula sendo que essa é verdadeira. Esse tipo de erro é o que chamamos de tipo I. A cada teste feito, há sempre alguma probabilidade de que a decisão tomada seja um erro de tipo I. Cometer esse tipo de erro é análogo ao julgar uma pessoa inocente como culpada. Para minimizar esse erro, a pessoa é assumida ser inocente ao se iniciar o julgamento. Da mesma forma, para se minimizar a probabilidade de se cometer o erro de tipo I, ao iniciar um teste de hipótese a hipótese nula sempre é assumida como verdadeira.

Sempre que se assume que a hipótese nula é verdadeira, deve-se controlar o erro do tipo I definindo um nível de significância. O nível que é estabelecido, o chamado nível alfa (α), também conhecido como *size* do teste é a maior probabilidade de se cometer um erro desse tipo. A decisão correta é rejeitar uma hipótese nula quando essa é falsa. Há sempre alguma probabilidade de se decidir que a hipótese nula é falsa quando realmente é falsa. Esta decisão é o chamado *power* do processo de tomada de decisão. Ela é chamado de *power* porque é a decisão que se almeja e é definida por $1 - \beta$.

A qualidade dos testes pode analisada através da comparação do *size* e *power*

do teste. Um teste pode ser considerado de boa qualidade, se tiver o maior *power* e *size* igual ou inferior a uma determinada quantidade estipulada. Ao aumentar o nível de significância, a região de aceitação é reduzida, conseqüentemente torna-se mais provável rejeitar a hipótese nula, sendo menos propenso aceitar a hipótese nula quando ela é falsa, ou seja, é menos provável que se cometa um erro do tipo II.

$$\alpha = P(\text{Erro do tipo I}) = \text{size} \quad \beta = P(\text{Erro do tipo II})$$

$$\text{Power} = 1 - \beta$$

Tabela 2: Erro do Tipo I e Tipo II

		Decisão	
		Aceitar H_0	Rejeitar H_0
Situação real	H_0 Verdadeira	Decisão correta $1 - \alpha$	Decisão Incorreta Erro do tipo II β
	H_0 Falsa	Decisão Incorreta Erro do Tipo I α	Decisão correta $1 - \beta$

Um teste de uma hipótese estatística, onde a região de rejeição é apenas em um dos lados da distribuição de amostragem, é chamado de teste unilateral, seja ele à esquerda como à direita. Já quando a região de rejeição é em ambos os lados o teste é conhecido como bilateral.

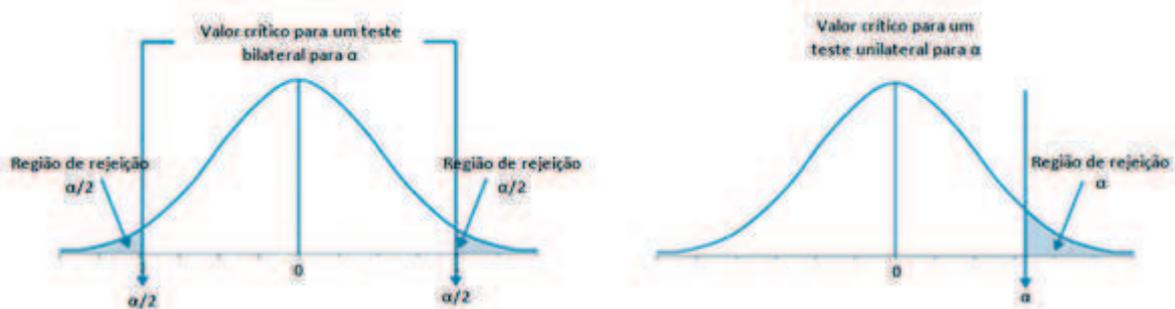


Figura 10: Representação dos testes bilateral e unilateral à direita

3.4.1 Teste de Hipóteses Para Comparação Entre Modelos

Seja A um modelo com vários preditores que representa uma determinada resposta, pode-se criticar se todos esses componentes são realmente necessários para

modelar tal resposta. Seja B um modelo menor que A, que contém um subconjunto dos preditores de A, pelo princípio de Occam's Razor, modelos parcimoniosos se qualificados para representar a resposta de interesse são preferíveis do que modelos maiores.

Para testar tal hipótese definiu-se A o representante da hipótese nula e B da hipótese alternativa.

Seja RSS_A a soma do quadrado dos resíduos do modelo maior e RSS_B a do modelo menor, se $RSS_B - RSS_A$ é muito pequeno, então B é um modelo adequado para representar a mesma resposta que a do modelo A. Isso sugere que a Eq. (3.12) poderia ser eficiente, sendo o denominador utilizado para padronização da escala.

$$\frac{RSS_B - RSS_A}{RSS_A} \quad (3.12)$$

Suponha que a dimensão de A seja q e a dimensão de B igual a p. Pelo teorema de Cochran se a hipótese nula A for verdadeira, então:

$$\frac{RSS_B - RSS_A}{RSS_A} \sim \sigma^2 \chi_{q-p}^2 \quad \frac{RSS_A}{n - q} \sim \sigma^2 \chi_{n-q}^2 \quad (3.13)$$

Sendo essas independentes. Desse modo, tem-se que:

$$F = \frac{(RSS_B - RSS_A)/(q - p)}{RSS_A/(n - q)} \sim F_{q-p, n-q} \quad (3.14)$$

Então, rejeitaríamos a hipótese nula se $F > F_{q-p, n-q}^{(\alpha)}$.

Os graus de liberdade de um modelo, geralmente são dados pela diferença entre o número de observações e a quantidade de parâmetros, assim a estatística teste pode ser escrita como:

$$F = \frac{(RSS_B - RSS_A)/(df_B - df_A)}{RSS_A/df_A} \quad (3.15)$$

onde $df_A = n - q$ e $df_B = n - p$. Esse teste é muito utilizado em regressão e análise de variância. Quando aplicado em diferentes situações, a forma da estatística teste pode ser expressa de diferentes jeitos.

3.5 Delineamento e Análise de Experimentos

De acordo com Montgomery (2001), um experimento pode ser definido como um teste ou uma série de testes, que propõe alterações nas variáveis de entrada de um sistema, de modo a identificar as alterações que ocorrem na variável de saída, sendo que geralmente o principal objetivo é desenvolver um processo robusto que seja minimamente afetado por efeitos externos.

O Delineamento de experimentos, ou DOE (*Design of experiments*) é uma metodologia que organiza os experimentos de modo que com um número mínimo de experimentos seja possível obter dados de qualidade que após serem analisados através de métodos estatísticos permitam condições apropriadas para se tirar conclusões válidas e objetivas.

Segundo Montgomery (2005), as vantagens para quem faz planejamento são: diminuição dos números de ensaios; estudo de um número considerável de fatores; detecção das interações entre os fatores; definição dos níveis ótimos; melhoria e precisão de resultados; otimização dos resultados e maximização dos lucros.

A técnica DOE teve suas primeiras aplicações por volta da década de 1920, sendo os trabalhos de Fisher e Student os pioneiros no assunto. No início da década de 1930 o método foi aprimorado por autores como Box, Wilson, Taguchi, dentre outros. Segundo Montgomery (1991) Fisher é considerado o inovador da utilização de técnicas estatísticas em projetos que envolvem experimentos por ter aplicado pela primeira vez a metodologia na área agrícola e em ciências biológicas.

Apesar de ter suas primeiras publicações nessas áreas, as técnicas do Projeto e Análise de Experimentos têm encontrado uma ampla aplicação em diversas áreas de conhecimento. Pode-se notar o aumento do interesse da aplicação do método em casos diversos na atualidade como, por exemplo, Leme et al., (2012), basearam-se no Projeto de Experimentos para identificar o impacto das variáveis ambientais sobre a ineficiência entre os diferentes cenários, e definir o fator de contexto que mais afeta a ineficiência DISCO. Balestrassi et al., (2009) utilizaram a técnica DOE para evitar a técnica de tentativa e erro no treinamento de uma RNA, em uma aplicação de séries temporais não-lineares para previsões de curto prazo para preços diários e retornos, consumo de água ou carga de eletricidade para consumidores industriais de uma companhia brasileira.

Muitas indústrias têm utilizado essa metodologia para resolução de seus problemas, conforme destacam Konda et al (1998), pois é possível um maior conhecimento

do processo, bem como posteriores melhorias em dispersões.

Para que a aplicação da técnica de planejamento e análise de experimentos tenha um bom resultado é necessário que se tenha atenção em algumas etapas importantes. Montgomery (2005) descreve com clareza os passos principais que devem ser seguidos para se ter sucesso na utilização do método.

O primeiro passo se trata da definição do problema, o experimentador deve ter o máximo de informação a respeito do problema a ser estudado. O segundo passo é a escolha dos fatores e definição dos níveis de trabalho. O terceiro passo trata-se da seleção das variáveis de resposta, essas devem ter informação suficiente para que seja possível encontrar solução para o problema em questão. O quarto passo é a escolha do projeto experimental. Após esse passo executam-se os experimentos e a análise estatística dos dados é realizada para assim então elaborar as conclusões e recomendações.

Dentre aos arranjos mais utilizados no planejamento e análise de experimentos pode-se citar:

- Planejamento fatorial completo 2^k ;
- Planejamento fatorial fracionado $2^{(k-p)}$;
- Planejamento Plackett-Burman;
- Engenharia Robusta - Taguchi;
- Metodologia da Superfície de Resposta (MSR).

Cada arranjo citado possui uma aplicação, sendo assim Tabela 3, resume de uma forma clara e objetiva as principais características relacionadas a cada uma dessas técnicas experimentais.

O presente trabalho teve por foco a otimização da modelagem de séries temporais lineares com o intuito de se evitar a identificação de modelos através da técnica de tentativa e erro. Por se tratar de modelos lineares o objetivo principal foi o de investigar a existência ou não dos parâmetros de entrada. Para isso, optou-se pela utilização dos arranjos Plackett Burman nque por conseguinte será descrito de maneira mais objetiva.

Tabela 3: Principais Técnicas de Análise de Experimentos e características fundamentais (Adaptado NILO JÚNIOR, 2003)

Projeto Experimental	Vantagens	Desvantagens	Aplicações
Fatorial Completo 2^k	Permite varredura completa da região de estudo, pois utiliza todos os fatores e respectivos níveis.	Não identifica variação intermediária, pois só trabalha em dois níveis, além de demandar um alto nº de corridas para problemas com grande nº de variáveis.	Em processos onde já se tem um prévio domínio e onde a realização das corridas não demanda alto tempo ou custo.
Fatorial Fracionado $2^{(k-1)}$	Permite uma pré-análise do processo com um nº reduzido de corridas.	Não permite varredura completa da região.	Em processos onde se deseja um pré-conhecimento e onde a literatura é limitada ou para corridas que demandam maior tempo ou custo.
Plackett Burman	Permite estimar os efeitos principais e identificar as variáveis mais relevantes de um processo com um nº reduzido de experimentos	Não permite varredura completa da região e nem identifica interações.	Em processos onde o interesse é verificar apenas os efeitos principais.
Taguchi	Permite uma análise de um processo com muitas variáveis de entrada com um nº extremamente reduzido de experimentos.	Geralmente dá uma ideia do processo, porém pode apresentar modelos matemáticos não confiáveis.	Em processos onde há pouco ou quase nenhum conhecimento prévio de comportamento, para processos com alta dispersão ou que as corridas demandem alto custo ou tempo.
MSR	Permite a verificação de variações intermediárias do processo.	Pode apresentar erros na extrapolação dos níveis +2 e -2, pois são realizadas poucas corridas nestes níveis.	Otimização de processos, principalmente bem conhecidos e com baixa dispersão.

3.5.1 Arranjos Plackett-Burman

Os arranjos do tipo Plackett-Burman podem ser classificados como arranjos fatoriais fracionados de dois níveis. Esse delineamento é indicado para projetos onde o interesse está apenas nos efeitos principais.

Do ponto de vista econômico, os arranjos Plackett-Burman são mais vantajosos do que os outros delineamentos devido a redução do número de experimentos, por exemplo, um projeto com 11 fatores pode ser analisado utilizando apenas 12 experimentos de uma forma eficiente. Porém, por ser um arranjo de resolução III, ou seja, permite o confundimento dos efeitos principais com interações duplas e das mesmas entre si, esse arranjo é utilizado como *screening designs*, em outras palavras, é utilizado em projetos de varredura onde o interesse está apenas na identificação dos efeitos principais significativos, sendo as interações insignificantes se comparadas com os mesmos. Esse é um ponto crucial para se optar pela utilização do *design* Plackett-Burman, deve-se ter a certeza de que as interações são realmente insignificantes perante a significância dos efeitos principais.

Os desenhos Plackett-Burman possuem o formato $k = (n - 1)$, isto é, o número de fatores de estudo (k) é igual ao número de experimentos (n) menos um e são formados por *runs* com número múltiplos de quatro, com sinais positivos representando os níveis superiores e negativos os níveis inferiores dos fatores do experimento.

A matriz experimental do delineamento Plackett-Burman segue o princípio do deslocamento de coluna, sendo este conhecido como variação helicoidal. A primeira linha da matriz é a linha de base, o sinal de sua última coluna desloca-se para a primeira coluna da segunda linha e o sinal de sua primeira coluna desloca-se para a segunda coluna da segunda e assim por diante, completando todas as colunas da segunda linha. A mesma regra é repetida até a penúltima linha, a última linha da matriz é completada apenas com sinais negativos, que representa todos os fatores no nível inferior.

A Tabela 4 representa a matriz experimental de um arranjo Plackett-Burman, construída para um projeto com 11 fatores e conseqüentemente 12 experimentos. Analisando a matriz experimental pode se ver perfeitamente o seguimento da regra especificada no parágrafo anterior.

No próximo capítulo será apresentada a metodologia proposta e os resultados obtidos com a utilização da técnica de delineamento de experimentos para construção de modelos de séries temporais. No trabalho em questão optou-se por utilizar

Tabela 4: Arranjo Plackett-Burman padrão para $k = 11$ $n = 12$.

+	-	+	-	-	-	+	+	+	-	+
+	+	-	+	-	-	-	+	+	+	-
-	+	+	-	+	-	-	-	+	+	+
+	-	+	+	-	+	-	-	-	+	+
+	+	-	+	+	-	+	-	-	-	+
+	+	+	-	+	+	-	+	-	-	-
-	+	+	+	-	+	+	-	+	-	-
-	-	+	+	+	-	+	+	-	+	-
-	-	-	+	+	+	-	+	+	-	+
+	-	-	-	+	+	+	-	+	+	-
-	+	-	-	-	+	+	+	-	+	+
-	-	-	-	-	-	-	-	-	-	-

o arranjo Plackett-Burman para a construção dos modelos, isso porque se trata de modelos lineares e partindo do pressuposto de que as interações são insignificantes perante aos efeitos principais. Cada linha da matriz experimental indica um modelo a ser testado, com a redução do número de experimentos a identificação do modelo mais apropriado torna-se mais rápida devido a diminuição do tempo computacional necessário para simular todos os modelos indicados na matriz.

3.6 Análise de Variância

Assim como para o método de regressão e dos mínimos quadrados, uma alternativa para se verificar a adequação de um modelo estimado utilizando a técnica de delineamento e análise de experimentos é a metodologia da análise de variância (ANOVA). Por meio dessa metodologia, é possível analisar a significância dos coeficientes e determinar quais realmente impactam na resposta de interesse.

O método em questão se baseia no princípio de teste de hipóteses. Esse testa se existe uma relação linear entre uma variável de resposta y , e um determinado subconjunto de regressores $x_1, x_2, x_3, \dots, x_k$. (Montgomery e Runger, 2003).

De uma forma geral, pode-se resumir a metodologia de análise de variância aplicada em um experimento em sete passos, são eles:

1. Obtenção do modelo matemático completo e análise da adequação do mesmo através da determinação de F.
2. Verificação da regressão mais adequada.
3. Obtenção dos valores de F para todos os coeficientes.

4. Verificação de R^2 , para determinação da porcentagem de variação explicada pelo modelo obtido.
5. Eliminação dos coeficientes com baixa probabilidade de significância e determinação do modelo matemático reduzido.
6. Verificação da adequação dos novos modelos através do teste de F.
7. Verificação do novo valor de R^2

No item de teste de hipótese apresentado anteriormente foi mostrada uma explicação de forma resumida para a obtenção dos valores da estatística F. A partir desse valor é determinada a probabilidade de se cometer um erro ao aceitar a hipótese que se deseja testar, como já dito anteriormente, o chamado P_{value} . Essa estatística também pode ser calculada para cada coeficiente do modelo e assim identificar quais realmente são significativos com a finalidade de se ajustar o modelo de forma que seja mais parcimonioso.

Outra forma de se verificar a adequação de um modelo através da análise do R^2 , o qual se refere a real porcentagem de explicação do modelo. Quanto maior o R^2 , mais eficiente está o modelo, ou seja, grande percentual de informações da dinâmica da variável de resposta pode ser explicada pelo modelo.

4 *Desenvolvimento*

Neste capítulo é apresentada a metodologia proposta para esse trabalho bem como a aplicação para doze casos genéricos e para três processos reais. Além disso, traz os resultados obtidos e vantagens e desvantagens do método.

4.1 Metodologia Proposta

A metodologia proposta por esse trabalho é a utilização da técnica de delineamento de experimentos para se tentar evitar a utilização de métodos baseados na comumente chamada tentativa e erro, o que muitas vezes ocorre em modelagens de séries temporais ao se utilizar análises das funções de autocorrelação.

A técnica de experimentação se torna muito interessante para esse trabalho pois por se tratar de experimentos simulados o único erro que está embutido durante a experimentação é o erro do próprio modelo simulado, ou seja, não existem fatores externos que influenciam nos resultados obtidos.

A metodologia foi aplicada para três tipos de modelos lineares:

- Autoregressivos (AR);
- Médias móveis (MA);
- Autoregressivos com médias móveis (ARMA).

A Tabela 5 contém os modelos genéricos criados para se testar a técnica proposta, essas foram criadas visando abranger os diferentes tipos de cenários existentes em modelagens de série temporais lineares.

Para cada modelo criado, foram gerados 1500 dados, sendo 150 utilizados apenas para validar os modelos estimados e 1350 observações para se construir os modelos.

Utilizou-se os desenhos Plackett-Burman para planejar os experimentos simulados

Tabela 5: Séries Temporais Lineares Genéricas

EQUAÇÕES	
Modelos AR	
SERIE A	$y_t = 0,55y_{t-1} + \varepsilon$
SERIE B	$y_t = 0,45y_{t-1} - 0,5y_{t-2} + \varepsilon$
SERIE C	$y_t = 0,4y_{t-1} + 0,15y_{t-2} + 0,25y_{t-3} + \varepsilon$
SERIE D	$y_t = 0,5y_{t-1} + 0,25y_{t-3} + \varepsilon$
Modelos MA	
SERIE E	$y_t = 0,3\varepsilon_{t-1} - 0,3\varepsilon_{t-2} + 0,5\varepsilon_{t-3} + \varepsilon$
SERIE F	$y_t = 0,5\varepsilon_{t-2} + \varepsilon$
SERIE G	$y_t = 0,65\varepsilon_{t-1} + \varepsilon$
SERIE H	$y_t = 0,5\varepsilon_{t-1} - 0,35\varepsilon_{t-2} + \varepsilon$
Modelos ARMA	
SERIE I	$y_t = 0,1y_{t-1} + 0,85y_{t-2} + 0,6\varepsilon_{t-1} + \varepsilon$
SERIE J	$y_t = 0,45y_{t-1} + 0,5y_{t-2} + 0,5\varepsilon_{t-2} + \varepsilon$
SERIE K	$y_t = 0,1y_{t-1} + 0,25y_{t-2} + 0,45y_{t-3} + 0,35\varepsilon_{t-1} + 0,5\varepsilon_{t-2} + 0,65\varepsilon_{t-3} + \varepsilon$
SERIE L	$y_t = 0,25y_{t-1} + 0,2\varepsilon_{t-1} + \varepsilon$

a fim de identificar quais fatores de uma série temporal linear interferem significativamente na eficiência de previsão de um modelo Box-Jenkins, ou seja, identificar a ordem do modelo a ser construído. Como já justificado anteriormente, optou-se pelos arranjos Plackett-Burman por se tratar de modelos lineares, existindo assim a premissa de que as interações são insignificantes perante aos efeitos principais. Além disso, para esse caso, esse tipo de arranjo é mais vantajoso do ponto de vista econômico, pois exige uma menor quantidade de experimentos e consegue-se obter resultados relevantes assim como nos outros tipos de delineamento.

Para os modelos AR e MA, foi construído um arranjo Plackett-Burman com 12 experimentos para se testar a significância de 6 fatores e para os modelos ARMA decidiu-se testar a significância de 11 fatores através de um arranjo com 24 experimentos. A Tabela 6, Tabela 7 e Tabela 8 contêm as matrizes experimentais dos modelos testados, onde o nível inferior, representado pelo sinal negativo (-) determina a ausência do fator, e o nível superior (+) a existência do fator. Para melhor entendimento pode-se ver Tabela 9 na um exemplo de um modelo AR representado por uma linha da matriz experimental.

Analisando a tabela em questão pode-se notar que existe a possibilidade de se encontrar modelos que dependem de observações com espaçamento de tempo diferente de um, por exemplo, um modelo AR(2) em que o fator “primeiro auto-regressivo” não seja significativo. Softwares comerciais como o MATLAB, utilizado nesse trabalho, não apresentam a possibilidade de se estimar os parâmetros de um modelo como do exemplo citado, por esse motivo, gerou-se um algoritmo, cujo o pseudo-código está re-

Tabela 6: Matriz experimental para modelos AR

	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
1	+	-	+	-	-	-
2	+	+	-	+	-	-
3	-	+	+	-	+	-
4	+	-	+	+	-	+
5	+	+	-	+	+	-
6	+	+	+	-	+	+
7	-	+	+	+	-	+
8	-	-	+	+	+	-
9	-	-	-	+	+	+
10	+	-	-	-	+	+
11	-	+	-	-	-	+
12	-	-	-	-	-	-

Tabela 7: Matriz experimental para modelos MA

	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
1	+	-	+	-	-	-
2	+	+	-	+	-	-
3	-	+	+	-	+	-
4	+	-	+	+	-	+
5	+	+	-	+	+	-
6	+	+	+	-	+	+
7	-	+	+	+	-	+
8	-	-	+	+	+	-
9	-	-	-	+	+	+
10	+	-	-	-	+	+
11	-	+	-	-	-	+
12	-	-	-	-	-	-

presentado na Fig. 11, baseado no método da máxima verossimilhança apresentado no capítulo anterior, para se estimar os parâmetros de cada modelo simulado.

Durante a simulação da rotina de maximização da função de Máxima Verossimilhança, alguns modelos necessitaram de um número muito grande de interações para convergir ou até mesmo não convergiram para uma resposta. Pensando nisso, decidiu-se inserir uma restrição do número de interações permitida, sendo assim as respostas da matriz experimental que não convergiram foram desconsideradas das análises. Esse é um dos motivos de se ter escolhido um arranjo de 24 experimentos para estudar um projeto com 11 fatores, quando que um arranjo de 12 experimentos poderia ser satisfatório. Além disso, os experimentos nos quais apontavam para a inexistência de todos os fatores também foram ignorados, visto que não se tem interesse na resposta desse modelo.

Como todo método de previsão, o método proposto também visa à obtenção de

Tabela 8: Matriz experimental para modelos ARMA

	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5
1	+	+	+	+	-	-	-	-	+	-	+
2	-	-	+	+	-	+	-	+	+	+	+
3	-	-	+	+	-	-	+	+	-	+	-
4	-	+	+	-	+	-	+	+	+	+	+
5	+	+	-	-	-	-	+	-	+	-	-
6	+	+	-	-	+	+	-	+	-	+	+
7	-	+	-	-	+	+	-	-	+	+	-
8	-	+	-	+	+	+	+	+	-	-	-
9	+	-	-	-	-	+	-	+	-	-	+
10	-	+	-	+	-	-	+	+	-	-	+
11	+	+	+	+	+	-	-	-	-	+	-
12	-	-	-	-	-	-	-	-	-	-	-
13	-	+	+	+	+	+	-	-	-	-	+
14	-	-	+	-	+	-	-	+	+	-	-
15	+	-	+	-	+	+	+	+	+	-	-
16	+	+	-	+	-	+	+	+	+	+	-
17	+	-	+	-	-	+	+	-	-	+	+
18	+	-	+	+	+	+	+	-	-	-	-
19	-	+	+	-	-	+	+	-	+	-	+
20	-	-	-	-	+	-	+	-	-	+	+
21	-	-	-	+	-	+	-	-	+	+	-
22	+	+	+	-	-	-	-	+	-	+	-
23	+	-	-	+	+	-	-	+	+	-	+
24	+	-	-	+	+	-	+	-	+	+	+

Tabela 9: Exemplo de um modelo AR representado pela linha da matriz experimental

$$\begin{array}{cccccc} \phi_0 & \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 \\ \hline + & - & + & - & - & - \end{array}$$

$$\text{Modelo simulado} = y_t = \phi_0 + \phi_2 y_{t-2} + \varepsilon$$

modelos de previsão que minimizem os valores de incertezas embutidas no modelo. Com isso, optou-se por utilizar a soma do erro quadrático, ou seja, a variância incondicional como variável de resposta, focando a análise dos experimentos na identificação de modelos parcimoniosos que resultem em um menor valor de incerteza.

Para comparar a metodologia proposta com métodos tradicionais Box-Jenkins, foram plotadas as funções de autocorrelação e autocorrelação parcial para as séries analisadas, com a finalidade de tentar encontrar algum padrão que possa auxiliar na identificação do modelo a ser construído.

Para as primeiras quatro séries (SERIE A, SERIE B, SERIE C, SERIE D) foram plotadas apenas as funções de autocorrelação parcial, por se tratar apenas de modelos Auto-regressivos. Como pode-se ver na Fig. 12 os gráficos apresentados mostram

```

01: Procedimento
02: /* Testar os modelos da matriz experimental */
03:     P = Máxima ordem do modelo AR
04:     Q = Máxima ordem do modelo MA
05:     P + Q + 1 = número de fatores da matriz experimental
06:     /*Definição do arranjo Plackett – Burman*/
07: Se (Modelo Puro) faça
08:     k = 12
09: Fim - Se
10: Se (Modelo Misto) faça
11:     k = 24
12: Fim - Se
13: /*Cálculo dos coeficientes dos modelos da matriz experimental*/
14: Para i = 1 até k faça
15:     /* i é a posição da linha da matriz experimental */
16:     x = vetor dos coeficientes do modelo i
17:     var = variância do modelo i
18: Fim - Para

```

Figura 11: Pseudo-código para cálculo das respostas para análise do DOE

que para a SERIE A apenas o lag 1 parece ser significativo o que indica um modelo AR(1). Para a SERIE B, a função mostra que o lag 1 e 2 são significativos, resultando um modelo AR(2). A SERIE C parece apresentar os 3 primeiros lags significantes, o que resulta em um modelo de terceira ordem, AR(3). O mesmo se pode notar na função de autocorrelação parcial da SERIE D, identificando outro modelo AR(3).

Aplicando a metodologia proposta no trabalho foram realizados os experimentos com os 1350 dados de cada série auto-regressiva. A Tabela 10 apresenta as respostas obtidas nas simulações dos experimentos realizados de acordo com a matriz experimental.

Analisando estatisticamente as respostas obtidas com os experimentos obtém-se os resultados apresentados na Figura 13, para a série SERIE A, SERIE B, SERIE C e SERIE D.

Pode-se verificar a influência dos parâmetros dos modelos através do gráfico dos efeitos principais. Quanto mais inclinada encontra-se a linha referente aos termos, mais significativo é o efeito. A reta pode mostrar que um efeito é significativo positivamente, ou seja, ao se incrementar a quantidade do termo no modelo o valor da resposta também aumentará, ou que um efeito é significativo negativamente, isto é, o incremento de um efeito causará um decremento na resposta.

Esse tipo de gráfico é utilizado mais para efeito visual, por não se utilizar de tes-

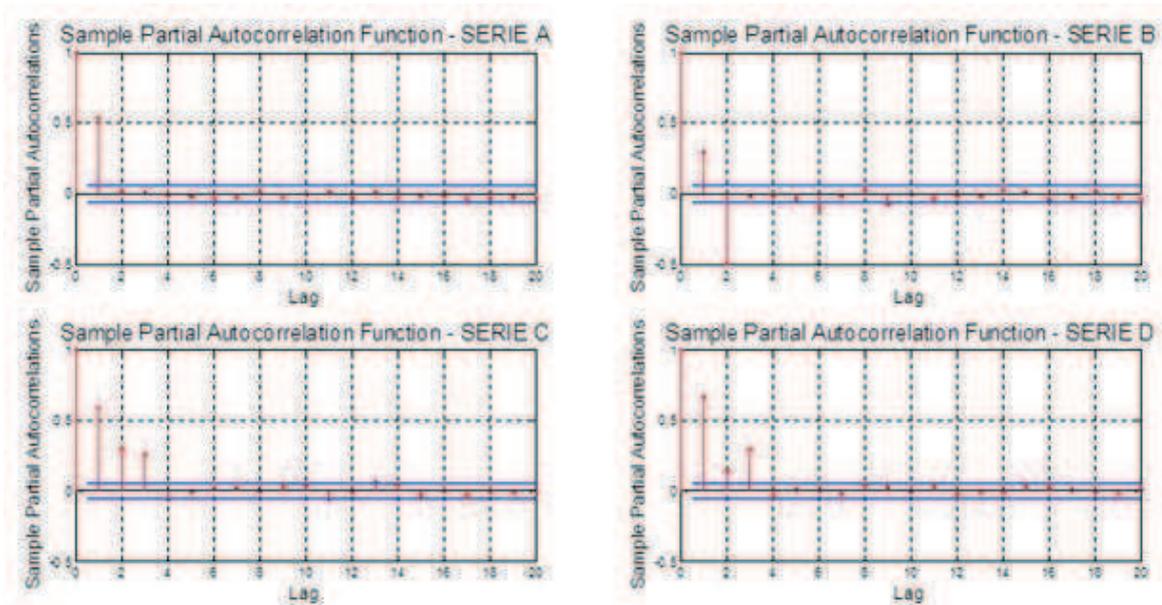


Figura 12: Autocorrelação Parcial dos modelos AR

Tabela 10: Respostas obtidas para os modelos AR

ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	SERIE A	SERIE B	SERIE C	SERIE D
+	-	+	-	-	-	-10,3698	-10,4874	-10,4993	-10,4256
+	+	-	+	-	-	*	*	*	*
-	+	+	-	+	-	-10,6174	-10,4358	-10,5828	-10,6836
+	-	+	+	-	+	*	-10,4014	-10,4863	-10,5858
+	+	-	+	+	-	-10,3714	-10,4887	-10,4983	-10,4153
+	+	+	-	+	+	-10,3711	-10,4329	-10,3709	-10,3020
-	+	+	+	-	+	-10,6174	-10,4745	-10,5835	-10,6845
-	-	+	+	+	-	-10,6177	-10,6644	-10,5499	-10,6071
-	-	-	+	+	+	*	*	-10,6171	-10,6841
+	-	-	-	+	+	*	-10,6632	-10,5480	-10,6067
-	+	-	-	-	+	-10,3066	-10,4131	-10,4096	-10,3746
-	-	-	-	-	-	-10,2889	-10,3110	-10,2389	-10,2343

* Experimentos ignorados

tes estatísticos para sua precisão. Quando se necessita de uma confirmação mais precisa, pode-se analisar estatisticamente a influência dos fatores na resposta de interesse, nesse caso, na variância incondicional, através do gráfico de Pareto como mostrado no gráfico esquerdo na Fig. 13 apresentada. A linha sólida que atravessa o gráfico indica o limite para que um efeito seja considerado significativo.

Para um nível de significância de 5%, todos os fatores em que o efeito ultrapasse a linha de restrição são considerados significativos e devem compor o modelo, do mesmo modo, todos que não ultrapassem a linha devem ser retirados do modelo.

Assim, analisando as figuras apresentadas, é possível concluir que segundo a metodologia aplicada, os termos significativos para os modelos Autoregressivos são:

Tabela 11: Efeitos significativos para os modelos AR

Séries Autoregressivas	Termos significativos
SERIE A	ϕ_1
SERIE B	$\phi_1 \phi_2$
SERIE C	$\phi_1 \phi_2 \phi_3$
SERIE D	$\phi_1 \phi_3$

Se comparado com as informações obtidas pela função de autocorrelação parcial pode-se notar que para as três primeiras séries (A, B, C) o método conduziu às mesmas respostas. Porém, para a série D, nota-se que a metodologia proposta identificou um modelo mais parcimonioso que a função de autocorrelação parcial. A função identificou que os *lags* 1, 2 e 3 eram significativos e deveriam compor o modelo; já o proposto, identificou que apenas os *lags* 1 e 3, os que realmente compõem o modelo, seriam significativos.

Assim, após a identificação dos modelos a serem construídos, passa-se para a fase de estimação dos coeficientes. Essa também foi feita através do algoritmo construído para a simulação dos modelos contidos na matriz experimental. Estimou-se então os coeficientes dos modelos através do método da máxima verossimilhança e suas estatísticas de significância. Os resultados obtidos encontram-se na Tabela 12.

Note que, se comparado com os valores reais de seus coeficientes, os valores estimados se encontram bem próximos, o que mostra a eficiência do método de estimação.

Para validar a modelagem, utilizaram-se os 150 dados nunca vistos pelo modelo e fez-se 150 previsões ao passo de um. Os resultados de previsão obtidos estão ilustrados na Fig. 14. Percebe-se que os valores previstos estão dentro do intervalo de confiança de 95%, além disso, o modelo acompanhou a média dos valores reais, o que mostra mais uma vez que o método realmente pode ser eficiente.

O método em questão também foi testado para os modelos lineares médias móveis. Para comparar com o método proposto por Box-Jenkins analisamos as funções de autocorrelação, também com a finalidade de tentar identificar os modelos estudados.

Note que as funções de autocorrelação das séries estudadas, apresentadas na Fig. 15, indicam um modelo MA(3) para a série E, um modelo com apenas o segundo termo do modelo MA significativo para a série F, um modelos de primeira ordem para a série G e um modelo de segunda ordem para a série H.

Tabela 12: Coeficientes estimados para os modelos AR

	SERIE A	SERIE B	SERIE C	SERIE D
ϕ_0	-	-	-	-
$\hat{\phi}_0$	-	-	-	-
Erro Padrão	-	-	-	-
ϕ_1	0,5500	0,4500	0,4000	0,5000
$\hat{\phi}_1$	0,5354	0,4421	0,3425	0,5215
Erro Padrão	0,0228	0,0237	0,0271	0,0241
ϕ_2	-	-0,5000	0,1500	-
$\hat{\phi}_2$	-	-0,4957	0,1863	-
Erro Padrão	-	0,0241	0,0272	-
ϕ_3	-	-	0,2500	0,2500
$\hat{\phi}_3$	-	-	0,2595	0,2826
Erro Padrão	-	-	0,0278	0,02622
ϕ_4	-	-	-	-
$\hat{\phi}_4$	-	-	-	-
Erro Padrão	-	-	-	-
ϕ_5	-	-	-	-
$\hat{\phi}_5$	-	-	-	-
Erro Padrão	-	-	-	-
k	2,5000e-5	2,5000e-5	2,5000e-5	2,5000e-5
\hat{k}	2,4504e-5	2,3395e-5	2,4513e-5	2,2925e-5
Erro Padrão	9,3001e-7	9,2310e-7	1,0064e-6	1,1182e-6

Assim como para os modelos autoregressivos, foram simulados os modelos propostos pela matriz experimental e as respostas obtidas estão apresentadas na Tabela 13.

Após as análises estatísticas plotaram-se os gráficos de Pareto e dos efeitos principais e foi identificado que para a SERIE E todos os fatores do modelo de terceira ordem eram significativos. Para a SERIE F apenas o segundo termo do modelo MA poderia ser considerado significativo, já para a SERIE G apenas o primeiro termo do modelo MA seria relevante. Finalmente para a SERIE H foi identificado um modelo de segunda ordem com todos os fatores significativos.

Veja que o método proposto encontrou o mesmo resultado que se encontraria analisando as funções de autocorrelação.

Assim como para os modelos Autoregressivos, foi utilizado o mesmo algoritmo para se estimar os parâmetros do modelo, os valores encontrados bem como os erros padrões encontram-se na Tabela 14.

Após a estimação dos parâmetros, os modelos estimados foram validados através de previsões com 10% dados separados. Os resultados apresentados na Figura 16

Tabela 13: Respostas obtidas para os modelos MA

Θ_0	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	SERIE A	SERIE B	SERIE C	SERIE D
+	-	+	-	-	-	-10,4990	-10,6164	-10,2509	-10,3770
+	+	-	+	-	-	*	*	*	*
-	+	+	-	+	-	-10,4753	-10,4128	-10,6287	-10,5908
+	-	+	+	-	+	-10,2938	-10,4118	-10,6279	-10,5763
+	+	-	+	+	-	-10,4811	-10,6154	-10,2508	-10,3820
+	+	+	-	+	+	-10,3291	-10,6149	-10,2494	-10,3682
-	+	+	+	-	+	-10,4777	-10,4129	-10,6289	-10,5909
-	-	+	+	+	-	-10,3466	-10,6177	-10,6291	-10,6761
-	-	-	+	+	+	-10,5675	-10,6151	-10,6297	-10,6721
+	-	-	-	+	+	-10,3434	-10,6164	-10,6285	-10,6718
-	+	-	-	-	+	-10,4247	-10,4113	-10,2506	-10,2863
-	-	-	-	-	-	-10,2213	-10,4126	-10,2504	-10,2905

mostram que os dados previstos pelos modelos médias móveis estimados também estão dentro do limite previsto para um intervalo de confiança de 95%.

O estudo em questão também foi proposto para modelos mistos, ou seja, modelos autoregressivos com médias móveis, porém encontrou-se um pouco de dificuldade durante o desenvolvimento. A metodologia foi aplicada para quatro séries genéricas, construídas com a finalidade de englobar os diversos comportamentos de séries temporais.

Analisando o gráfico de Pareto que compõe a Fig. 17, pode-se notar que alguns fatores estão no limiar da significância, devendo-se fazer uma análise mais profunda para tomar a decisão de aceitá-lo ou rejeitá-lo na composição do modelo. Esse é o caso da modelagem das séries J e K. Ao analisar o gráfico de Pareto para as séries mencionadas nota-se que para a série J os fatores ϕ_0 e Θ_1 e para série K o fator ϕ_4 parecem ser significativos considerando um intervalo de confiança de 95%, porém como já se conhece a série sabe-se que esses não fazem parte do modelo. Uma das justificativas para esse acontecimento poderia ser o fato de se estar utilizando o intervalo de confiança de 95% ou até mesmo por se tratar de dados finitos, sendo que a quantidade utilizada pode não ter sido suficiente para transmitir toda informação necessária para construção do modelo.

Tabela 14: Coeficientes estimados para os modelos MA

	SERIE E	SERIE F	SERIE G	SERIE H
ϕ_0	-	-	-	-
$\hat{\phi}_0$	-	-	-	-
Erro Padrão	-	-	-	-
Θ_1	0,3000	-	0,6500	0,5000
$\hat{\Theta}_1$	0,2850	-	0,6512	0,5755
Erro Padrão	0,02407	-	0,0211	0,0265
Θ_2	0,3000	0,5000	-	-0,3500
$\hat{\Theta}_2$	0,2932	0,5033	-	-0,3512
Erro Padrão	0,0237	0,0253	-	0,0266
Θ_3	0,500	-	-	-
$\hat{\Theta}_3$	0,5311	-	-	-
Erro Padrão	0,0227	-	-	-
Θ_4	-	-	-	-
$\hat{\Theta}_4$	-	-	-	-
Erro Padrão	-	-	-	-
Θ_5	-	-	-	-
$\hat{\Theta}_5$	-	-	-	-
Erro Padrão	-	-	-	-
k	2,5000e-5	2,5000e-5	2,5000e-5	2,5000e-5
\hat{k}	2,5738e-5	2,4574e-5	2,4272e-5	2,3196e-5
Erro Padrão	9,7755e-7	9,5870e-7	9,5771e-7	8,7561e-7

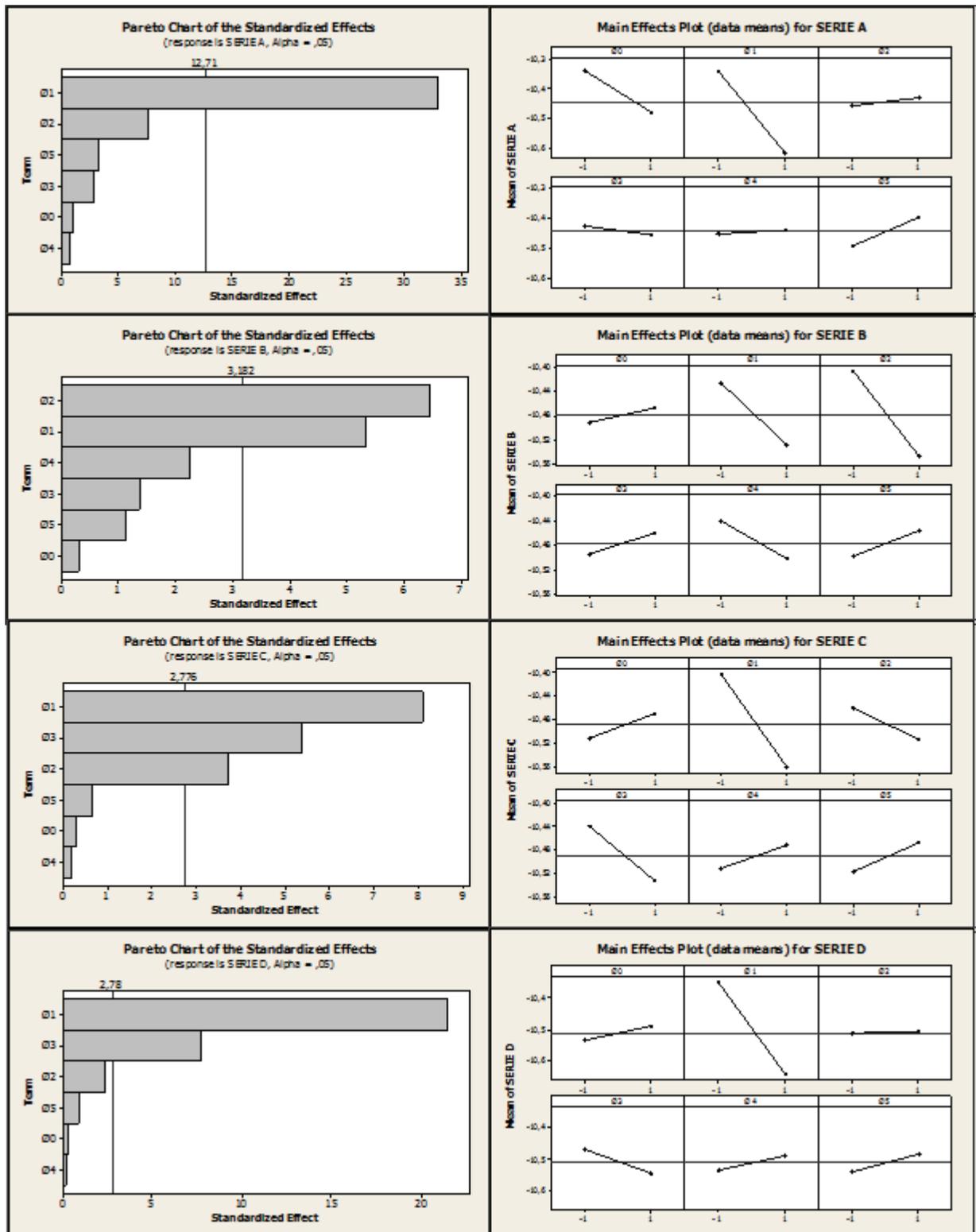


Figura 13: Diagrama de pareto e efeitos principais para os modelos AR

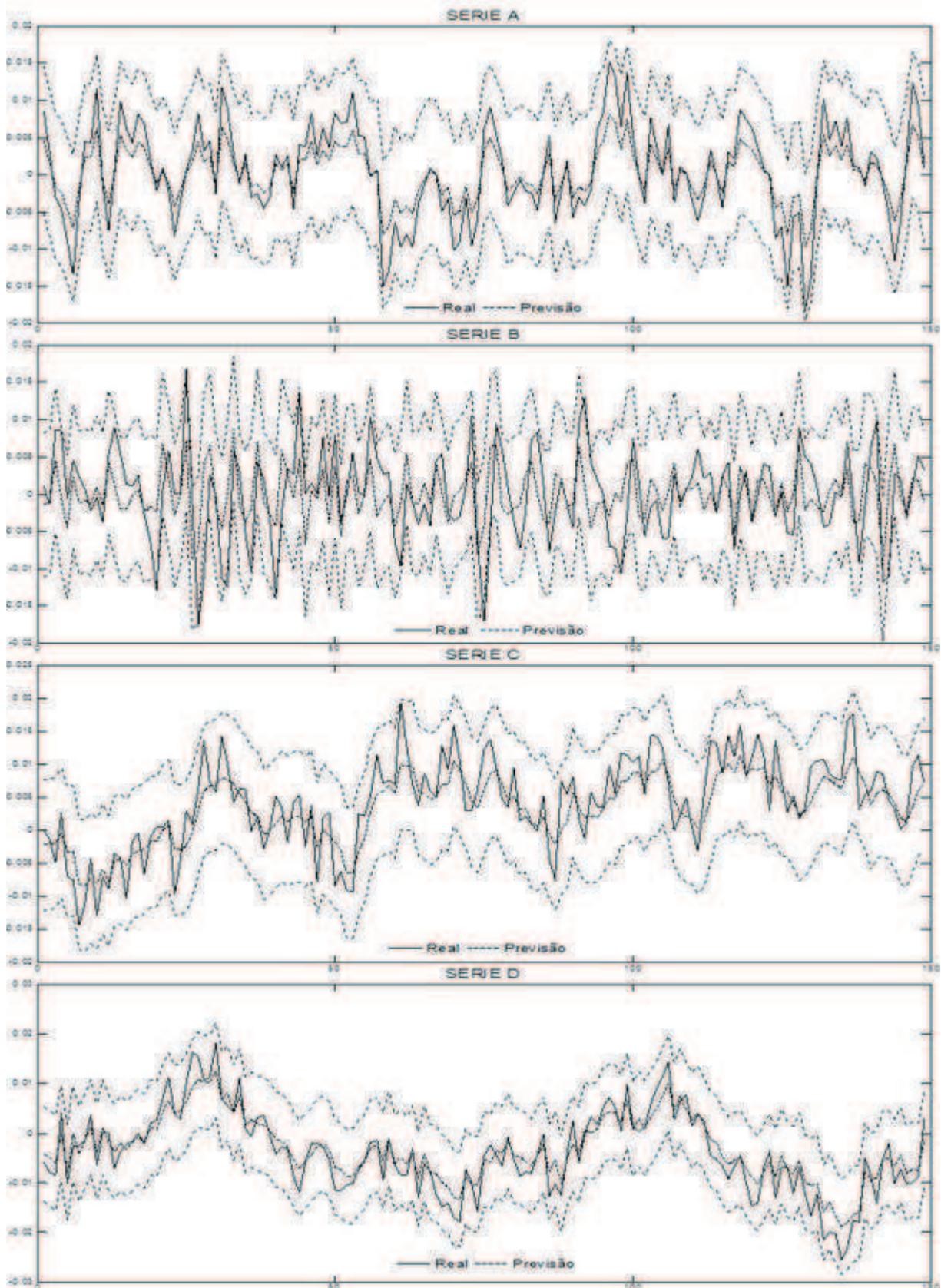


Figura 14: Validação dos modelos AR

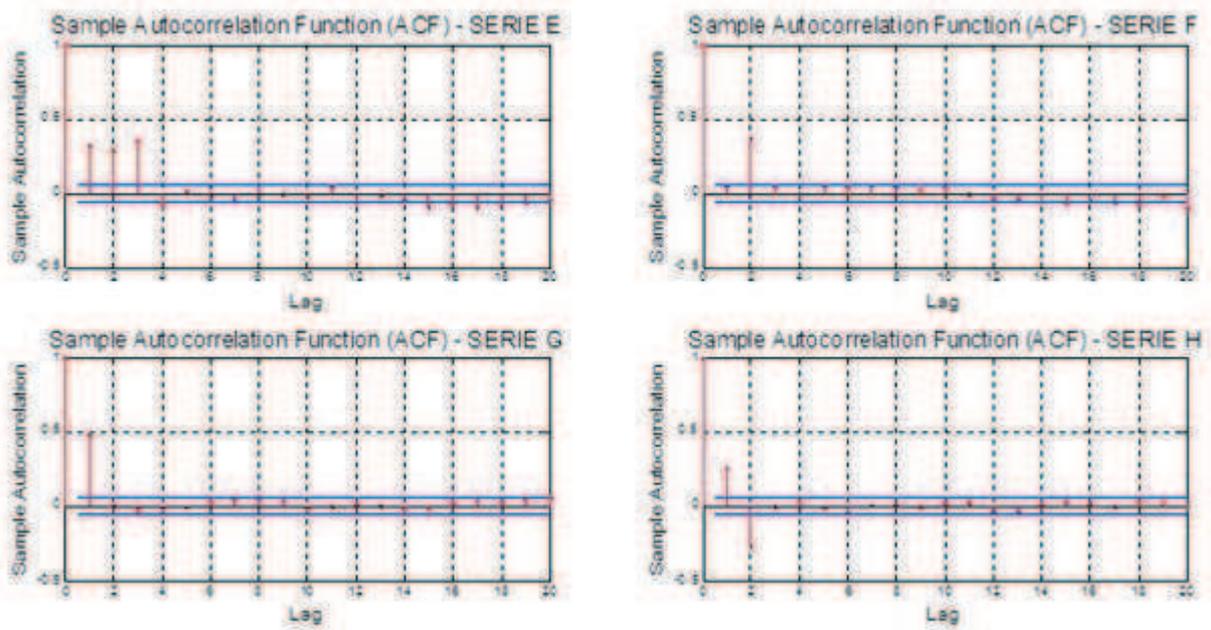


Figura 15: Autocorrelação dos modelos MA

Tabela 15: Respostas obtidas para os modelos ARMA

ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	SERIEI	SERIEJ	SERIEK	SERIEL
+	+	+	+	-	-	-	-	+	-	+	-10,443	-10,599	-10,352	-10,613
-	-	+	+	-	+	-	+	+	+	+	-10,332	-10,384	-10,395	-10,214
-	-	+	+	-	-	+	+	-	+	-	-10,503	-10,598	-10,336	-10,626
-	+	+	-	+	-	+	+	+	+	+	-10,517	-10,604	-10,571	-10,629
+	+	-	-	-	-	+	-	+	-	-	-10,140	-10,597	-10,212	-10,525
+	+	-	-	+	+	-	+	-	+	+	-10,484	-10,601	-10,040	-10,625
-	+	-	-	+	+	-	-	+	+	-	-9,959	-10,594	-10,250	-10,511
-	+	-	+	+	+	+	+	-	-	-	-10,467	-10,602	-10,502	-10,618
+	-	-	-	-	+	-	+	-	-	+	-9,853	-10,377	-9,603	-9,8283
-	+	-	+	-	-	+	+	-	-	+	-10,402	-10,598	-10,420	-10,583
+	+	+	+	+	-	-	-	-	+	-	-10,510	-10,599	-10,406	-10,618
-	-	-	-	-	-	-	-	-	-	-	*	*	*	*
-	+	+	+	+	+	-	-	-	-	+	-10,503	-10,602	-10,472	-10,625
-	-	+	-	+	-	-	+	+	-	-	-10,291	-10,376	-10,184	-10,211
+	-	+	-	+	+	+	+	+	-	-	-10,509	-10,600	-10,449	-10,626
+	+	-	+	-	+	+	+	+	+	-	-10,513	-10,602	-10,565	-10,590
+	-	+	-	-	+	+	-	-	+	+	-10,455	-10,599	-10,006	-10,622
+	-	+	+	+	+	+	-	-	-	-	-10,488	-10,599	-10,389	-10,624
-	+	+	-	-	+	+	-	+	-	+	-10,458	-10,601	-10,492	-10,626
-	-	-	-	+	-	+	-	-	+	+	-9,678	-10,575	-9,529	-10,048
-	-	-	+	-	+	-	-	+	+	-	-9,544	-10,357	-9,915	-9,762
+	+	+	-	-	-	-	+	-	+	-	-10,509	-10,599	-10,138	-10,586
+	-	-	+	+	-	-	+	+	-	+	-10,253	-10,378	-10,370	-10,214
+	-	-	+	+	-	+	-	+	+	+	-9,723	-10,578	-10,106	-10,107

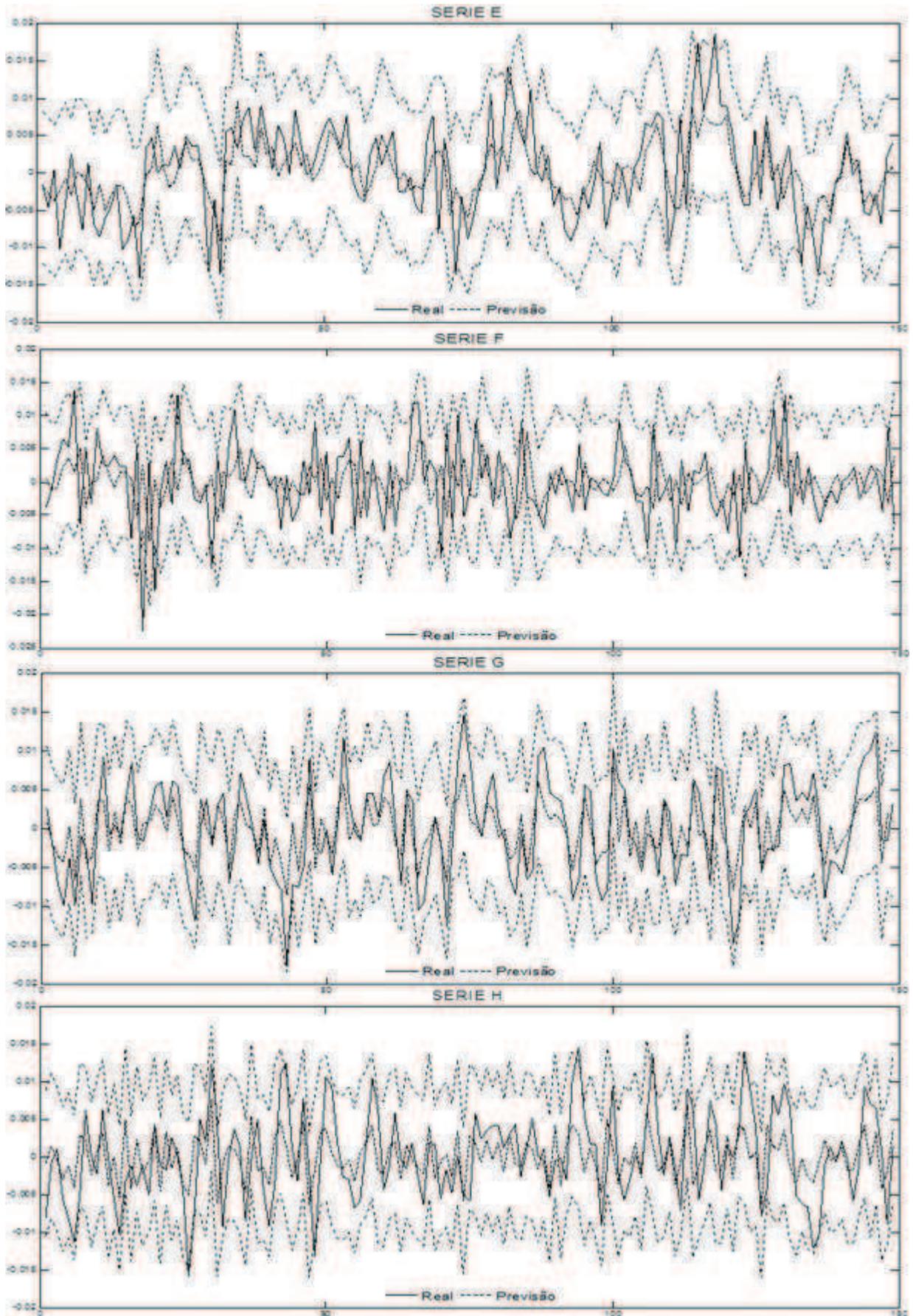


Figura 16: Validação dos modelos MA

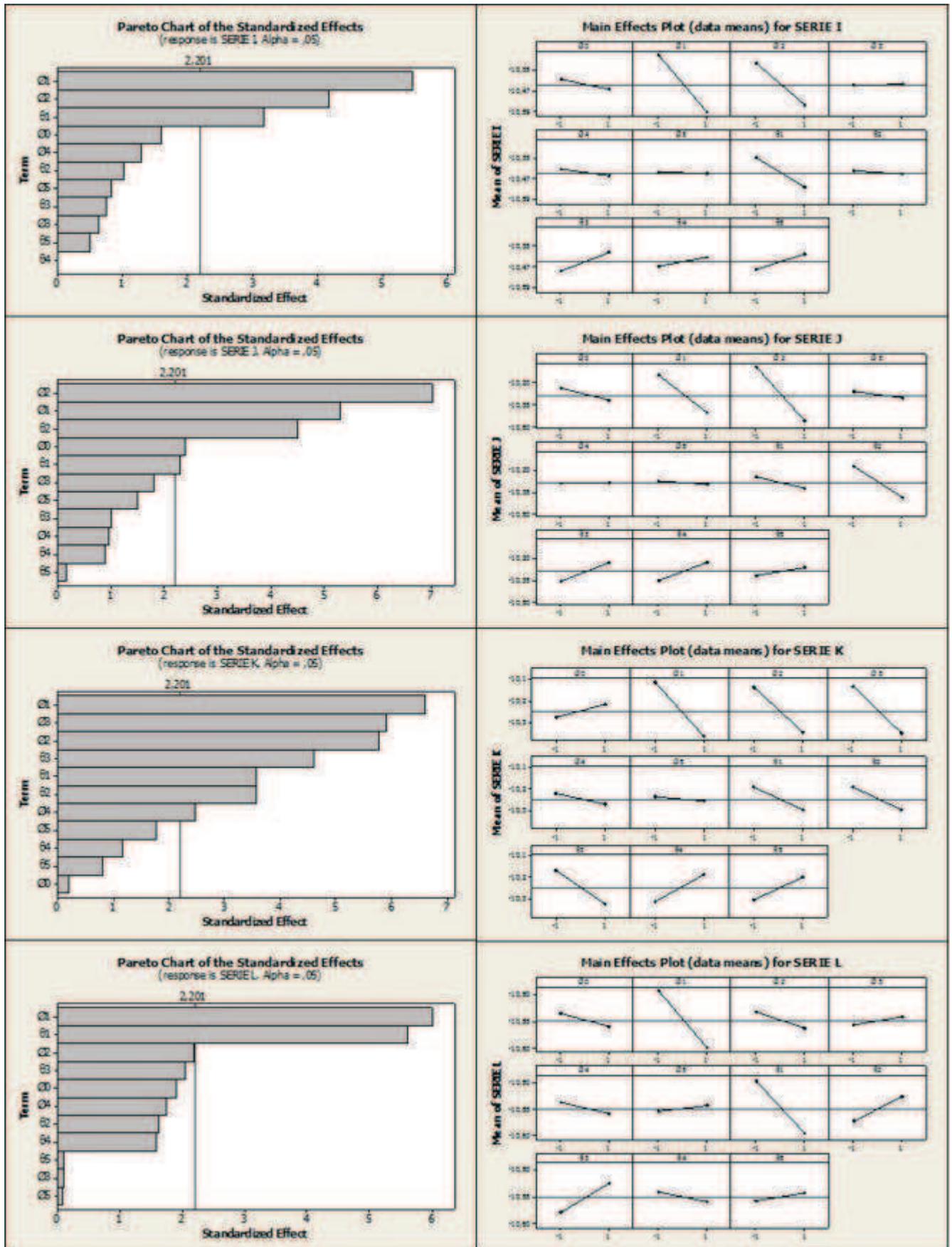


Figura 17: Diagrama de pareto e efeitos principais para os modelos ARMA

Tabela 16: Coeficientes estimados para os modelos ARMA

	SERIE I	SERIE J	SERIE K	SERIE L
ϕ_0	-	-	-	-
$\hat{\phi}_0$	-	-	-	-
Erro Padrão	-	-	-	-
ϕ_1	0,1000	0,4500	0,1000	0,2500
$\hat{\phi}_1$	0,1039	0,4397	0,1161	0,3245
Erro Padrão	0,0263	0,0208	0,0287	0,0446
ϕ_2	0,8500	0,5000	0,2500	-
$\hat{\phi}_2$	0,8346	0,5178	0,2435	-
Erro Padrão	0,0246	0,0214	0,0300	-
ϕ_3	-	-	0,4500	-
$\hat{\phi}_3$	-	-	0,4386	-
Erro Padrão	-	-	0,0277	-
ϕ_4	-	-	-	-
$\hat{\phi}_4$	-	-	-	-
Erro Padrão	-	-	-	-
ϕ_5	-	-	-	-
$\hat{\phi}_5$	-	-	-	-
Erro Padrão	-	-	-	-
Θ_1	0,600	-	0,3500	0,200
$\hat{\Theta}_1$	0,6105	-	0,3316	0,1752
Erro Padrão	0,0404	-	0,0249	0,0468
Θ_2	-	0,5000	0,5000	-
$\hat{\Theta}_2$	-	0,4199	0,5105	-
Erro Padrão	-	0,0265	0,0197	-
Θ_3	-	-	0,6500	-
$\hat{\Theta}_3$	-	-	0,6517	-
Erro Padrão	-	-	0,0234	-
Θ_4	-	-	-	-
$\hat{\Theta}_4$	-	-	-	-
Erro Padrão	-	-	-	-
Θ_5	-	-	-	-
$\hat{\Theta}_5$	-	-	-	-
Erro Padrão	-	-	-	-
k	2,5000 e-5	2,5000 e-5	2,5000 e-5	2,5000 e-5
\hat{k}	2,4413 e-5	2,7278 e-5	2,5454 e-5	2,4980 e-5
Erro Padrão	9,3580 e-7	1,0731 e-6	9,5966 e-7	9,8533 e-7

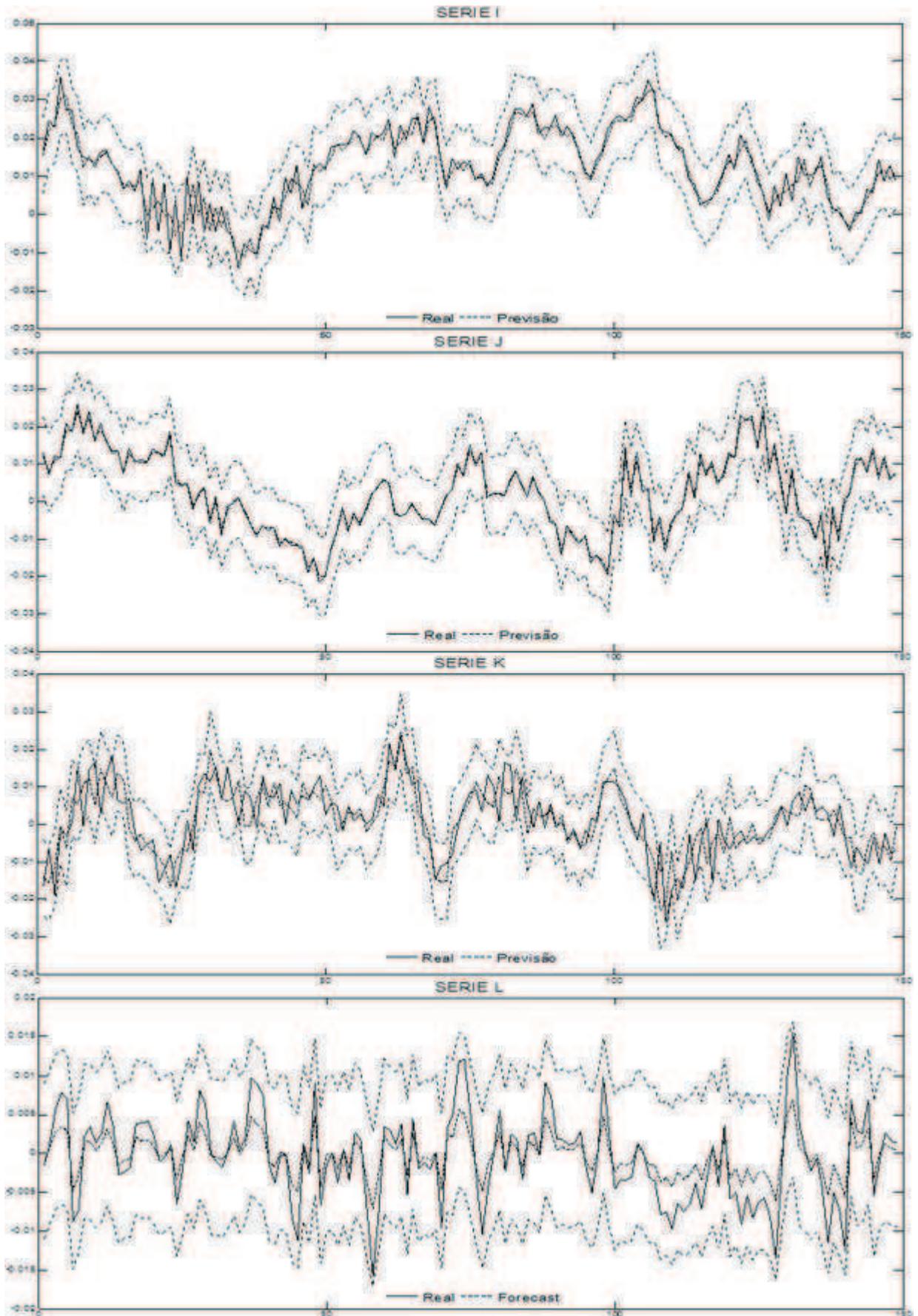


Figura 18: Validação dos modelos ARMA

4.2 Resultados Obtidos

Com o intuito de comprovar a eficiência da metodologia proposta, decidiu-se testá-la para previsões de processos reais. Para isso, escolheu-se três casos com aplicações distintas, buscando abranger comportamentos diferentes de séries temporais.

Para esses casos, foram escolhidas a mesma matriz experimental utilizada para a identificação de modelos ARMA, pois se acreditou que se o processo puder ser representado por modelos puros, ou seja, AR ou MA, a metodologia identificará tais modelos.

Para o primeiro teste da metodologia em casos reais utilizou-se o site <http://datamarket.com/> para a retirada dos dados. Escolheu-se então, modelar a média mensal de temperatura de outubro de 1955 a fevereiro de 1986 da cidade de Porto Alegre no estado do Rio grande do Sul por ser uma série que possuía uma quantidade de dados considerável e ter uma continuidade nas medições, mantendo um padrão mensal. Além disso, a opção por essa série se deu pelo seu comportamento linear e devido ao grande número de aplicações de modelos de séries temporais na área de meteorologia. A falta de dados ininterruptos forçou o estudo da média de temperatura apenas para o período citado.

Analisando a autocorrelação dos dados estudados (Fig. 19), pode-se notar que um mesmo padrão se repete a cada 24 observações. Isso significa que a série estudada apresenta um comportamento sazonal, necessitando então de uma transformação específica para a retirada dessa sazonalidade e aproximação do comportamento para o mais próximo de um estacionário.

Assim, antes da aplicação da metodologia, utilizou-se o artifício da diferenciação, nesse caso em 24, para contornar o problema de estacionariedade da série. A Fig. 20 apresenta a função de autocorrelação da série após utilizar-se a técnica da diferença, nota-se que seu comportamento se tornou próximo de um estacionário.

Após a transformação separou-se 10% dos dados para a validação do modelo e então, executou-se o algoritmo com o restante dos dados. Analisando a distribuição dos dados e visto que a quantidade de informação é limitada, optou-se por construir o modelo utilizando a função de máxima verossimilhança para a distribuição *t-Student* e assim como para as séries genéricas, foi utilizado o logaritmo natural da variância como resposta de estudo para a identificação do modelo através do DOE.

O segundo processo escolhido para se testar a metodologia foi o da descida de

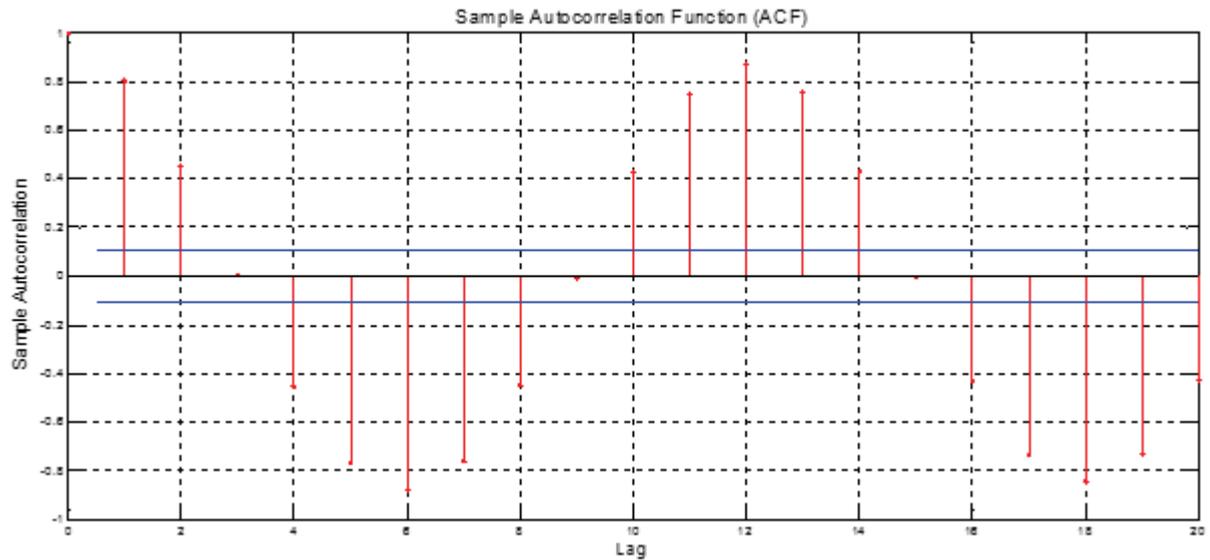


Figura 19: Autocorrelação da série M

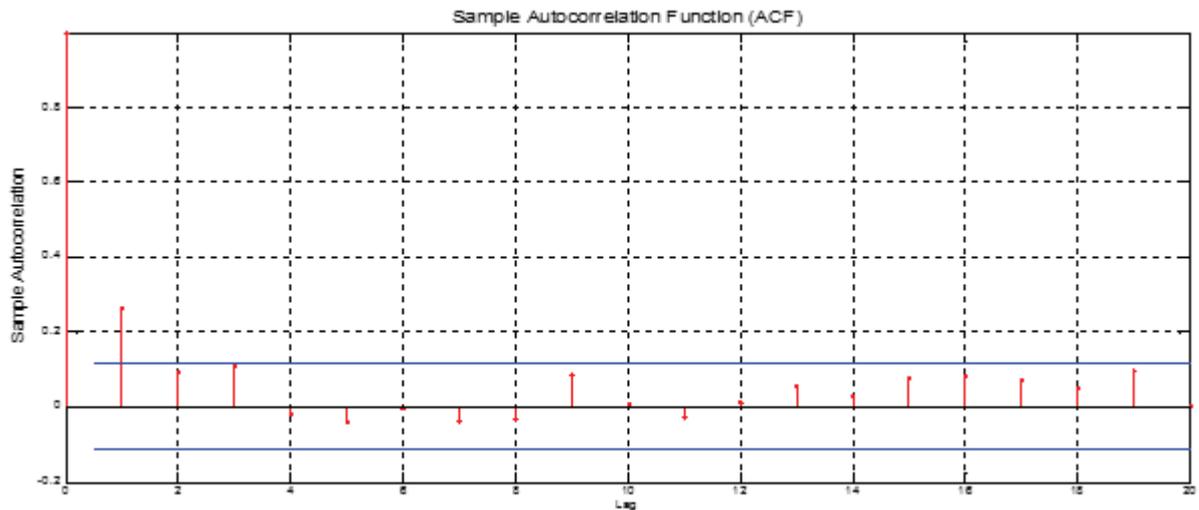


Figura 20: Autocorrelação da série M após a diferenciação

um registrador de pressão e temperatura em um poço de petróleo e gás natural.

Para se avaliar se é viável ou não a exploração de um poço de petróleo, deve-se colocá-lo em produção e coletar informações que ajudem na decisão de explorá-lo ou não. A temperatura e a pressão de um poço são variáveis fundamentais que auxiliam na determinação do tamanho de reservatório de petróleo e de sua estimativa de produção. Para se coletar tais informações utiliza-se um equipamento, denominado registrador que desce via cabo em uma coluna dentro do poço que se deseja testar e durante sua descida e subida são coletados dados de temperatura e pressão discretamente, com frequência de amostragem pré-determinada.

No caso da série estudada, a frequência de amostragem é de 5 segundos, sendo

que o teste teve duração de aproximadamente 8 horas.

Analisando a autocorrelação das séries de temperatura e de pressão, pode-se notar que ambas não têm um comportamento estacionário. Porém, diferente da série anterior, essas não possuem sazonalidade, necessitando apenas de uma transformação simples.

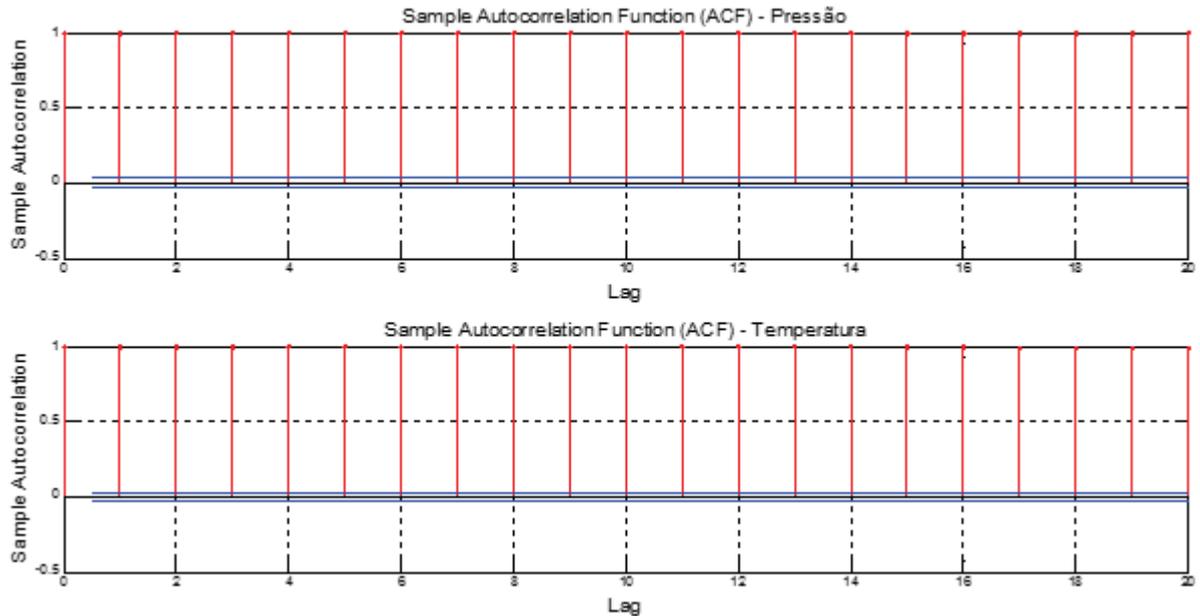


Figura 21: Autocorrelação das séries de pressão e temperatura de um poço de petróleo

Após algumas análises, concluiu-se que a segunda diferença seria suficiente para tornar as séries estacionárias. Por se já conhecer o processo estudado, o que levou a analisar correlação de uma série com a outra, descobriu-se que ambas estão correlacionadas entre si, e para usufruir desse comportamento optou-se por utilizar a técnica da análise das componentes principais introduzida no capítulo anterior.

Para aplicar a metodologia apresentada encontraram-se os autovetores e autovalores das séries e gerou-se uma série não correlacionada que continha 99% das informações do processo. Essa série já transformada, portanto estacionária, foi utilizada para construir os modelos de previsão para as duas séries.

Diferente da série M utilizou-se a função da máxima verossimilhança para uma distribuição normal para a estimação dos parâmetros do modelo. Apesar de se acreditar ser mais apropriada a utilização da função para distribuição t, a mesma não mostrou ser a mais adequada para o processo em questão.

Por fim, o último processo escolhido para se testar a metodologia de soldagem MIG/MAG por curto-circuito. Nesse tipo de processo os valores de tensão e corrente

podem transparecer a qualidade da solda. Se existe um processo estável, a probabilidade de se ter uma solda livre de defeitos aumenta, caso contrário, a instabilidade dos parâmetros do processo gera imperfeições no processo resultando em uma solda passível de fragilidade. A indicação de um valor de corrente e/ou tensão fora dos limites de previsão poderá indicar uma instabilidade no processo, permitindo com que o operador tome decisões adequadas para torna o processo estável novamente e tenha como resultado uma solda de qualidade.

Assim, como as outras séries essas não possuíam um comportamento estacionário, e fez-se necessário uma análise mais profunda das funções de autocorrelação e autocorrelação parcial para se tentar identificar padrões nos dados que auxiliem na transformação de seu comportamento não-estacionário para próximo de um estacionário.

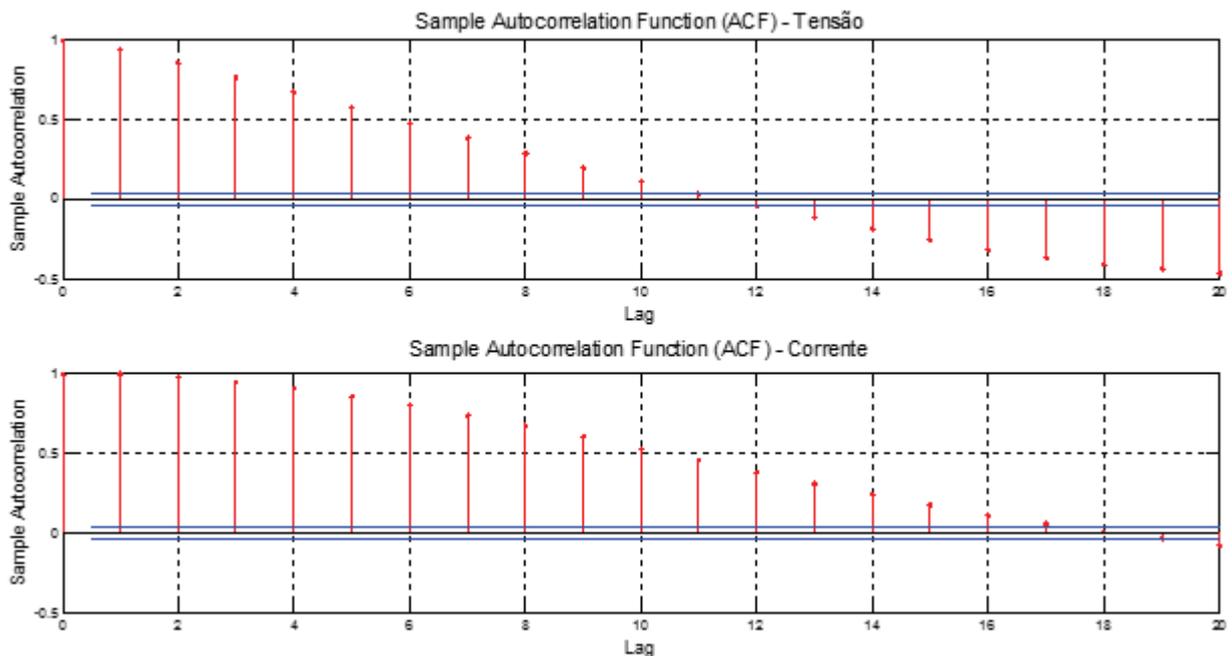


Figura 22: Autocorrelação das séries de tensão e corrente de um processo de soldagem MIG/MAG por curto-circuito

Ao se analisar tais funções, percebeu-se que o processo também possuía uma instabilidade, o que dificultava a identificação de padrões. Apesar de apresentar características que aproximava as séries de um padrão sazonal, não foi possível identificar seu período de sazonalidade devido a instabilidade do processo, porém pode-se verificar que a segunda diferença seria o suficiente para tornar as séries estacionárias.

Assim, como para as séries de pressão e temperatura, conhecendo as características do processo de soldagem MIG/MAG por curto-circuito, verificou-se que era interessante aproveitar a correlação entre os dados de tensão e de corrente para se

utilizar a técnica de componentes principais.

Para aplicar a metodologia, assim como para o processo anterior encontrou-se os autovetores e autovalores das séries e gerou-se uma série não correlacionada que também continha 99% das informações do processo. Dos 3999 dados disponíveis, utilizou-se 3600 para a construção do modelo de previsão e 399 para sua validação.

Assim como para pressão e temperatura do poço de petróleo, para esse processo utilizou-se a função da máxima verossimilhança para uma distribuição normal para a estimação dos parâmetros do modelo. Cabe ressaltar que existe a necessidade de um aprofundamento no estudo para se utilizar a distribuição t para a modelagem desses processos.

A Tabela 17 apresenta a matriz experimental com as respostas obtidas para as séries reais estudadas. Assim como para as séries genéricas, os resultados que não tiveram seu valor convergido foram desconsiderados.

Tabela 17: Respostas obtidas para os modelos das séries reais

ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	SERIE M	SERIE N	SERIE O
+	+	+	+	-	-	-	-	+	-	+	0,96474	-3,11284	2,73324
-	-	+	+	-	+	-	+	+	+	+	*	-2,80999	2,86154
-	-	+	+	-	-	+	+	-	+	-	0,95998	-3,11301	2,73206
-	+	+	-	+	-	+	+	+	+	+	0,91548	-3,11324	2,73253
+	+	-	-	-	-	+	-	+	-	-	0,96494	-3,10995	2,73299
+	+	-	-	+	+	-	+	-	+	+	*	-3,09306	2,73281
-	+	-	-	+	+	-	-	+	+	-	0,96495	-3,06292	2,74252
-	+	-	+	+	+	+	+	-	-	-	0,96428	-3,11299	2,73308
+	-	-	-	-	+	-	+	-	-	+	1,03940	-2,79475	2,87373
-	+	-	+	-	-	+	+	-	-	+	0,96474	-3,11285	2,73303
+	+	+	+	+	-	-	-	-	+	-	0,96264	-3,11300	2,73317
-	-	-	-	-	-	-	-	-	-	-	*	*	*
-	+	+	+	+	+	-	-	-	-	+	0,96420	-3,11284	2,73297
-	-	+	-	+	-	-	+	+	-	-	1,03391	-2,80988	2,87357
+	-	+	-	+	+	+	+	+	-	-	0,92463	-3,11290	2,67858
+	+	-	+	-	+	+	+	+	+	-	0,92431	-3,11298	2,73174
+	-	+	-	-	+	+	-	-	+	+	0,97706	-3,09312	2,73286
+	-	+	+	+	+	+	-	-	-	-	0,96343	-3,11282	2,73317
-	+	+	-	-	+	+	-	+	-	+	0,91735	-3,11284	2,73287
-	-	-	-	+	-	+	-	-	+	+	0,94792	-3,09020	2,72697
-	-	-	+	-	+	-	-	+	+	-	1,03226	-2,80383	2,87559
+	+	+	-	-	-	-	+	-	+	-	0,96533	-3,11300	2,73317
+	-	-	+	+	-	-	+	+	-	+	1,02786	-2,80970	2,87131
+	-	-	+	+	-	+	-	+	+	+	0,89581	-3,10684	2,68707

Analisando o gráfico de Pareto e dos efeitos principais encontrados na Fig. 23, pode-se verificar que foi identificado um modelo ARMA(1,1) para a série M, um modelo

ARMA(2,1) para a série N e um modelo ARMA(1,1) para a série O. Os valores dos coeficientes estimados encontram-se na Tabela 18. É interessante ressaltar que o erro padrão dos coeficientes para o modelo da série M possuem valores maiores do que os dos próprios coeficientes, o que corresponde a uma estatística com valor pequeno, que demonstra que os coeficientes são insignificantes.

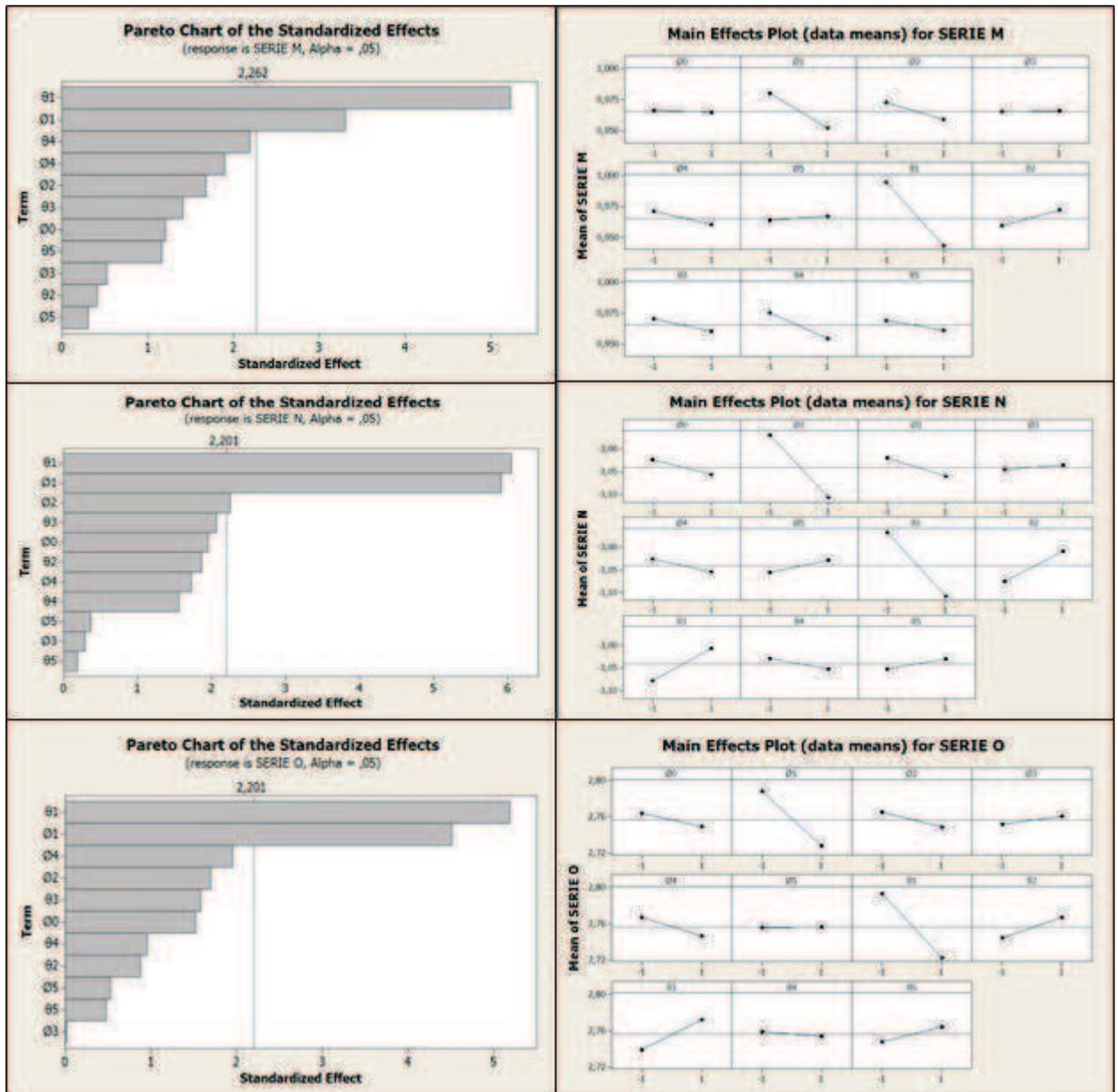


Figura 23: Diagrama de pareto e efeitos principais para os modelos das séries reais

Apesar dos coeficientes estimados para o modelo da série M apresentarem um alto erro padrão, a previsão e a análise de resíduos mostrou que o modelo construído foi eficiente, pois conseguiu prever valores dentro do intervalo de confiança de 95% e satisfaz as premissas estatísticas para as análises de resíduos.

Para as outras séries o erro padrão calculado para os coeficientes estimados teve

Tabela 18: Coeficientes estimados para os modelos das séries reais

	SERIE M	SERIE N	SERIE O
$\hat{\phi}_0$	-	-	-
Erro Padrão	-	-	-
$\hat{\phi}_1$	0,1172	-0,5661	0,1171
Erro Padrão	0,1946	0,0227	0,0335
$\hat{\phi}_2$	-	0,2453	-
Erro Padrão	-	0,0122	-
$\hat{\phi}_3$	-	-	-
Erro Padrão	-	-	-
$\hat{\phi}_4$	-	-	-
Erro Padrão	-	-	-
$\hat{\phi}_5$	-	-	-
Erro Padrão	-	-	-
$\hat{\Theta}_1$	-0,1471	0,0291	0,2715
Erro Padrão	0,1917	0,0230	0,0342
$\hat{\Theta}_2$	-	-	-
Erro Padrão	-	-	-
$\hat{\Theta}_3$	-	-	-
Erro Padrão	-	-	-
$\hat{\Theta}_4$	-	-	-
Erro Padrão	-	-	-
$\hat{\Theta}_5$	-	-	-
Erro Padrão	-	-	-
Dof	8,2289	-	-
Erro Padrão	0,2917	-	-
\hat{k}	2,6652	0,0444	15,37
Erro Padrão	12,1720	0,0001	0,1168

um baixo valor e todos foram considerados significantes de acordo com a estatística t. Porém, os resíduos que deveriam seguir uma distribuição normal, parecem seguir a distribuição t, apesar de se ter tentado aplicar a função de máxima de verossimilhança para uma distribuição t.

Tanto para a pressão quanto para a temperatura, o modelo estimado para a série N pareceu ser eficiente para as previsões feitas. Para a série O a validação do modelo para a série de tensão também apresentou resultados satisfatórios, porém, a série de corrente apresentou um pequeno deslocamento de média apesar de os valores previstos acompanharem o comportamento real da série.

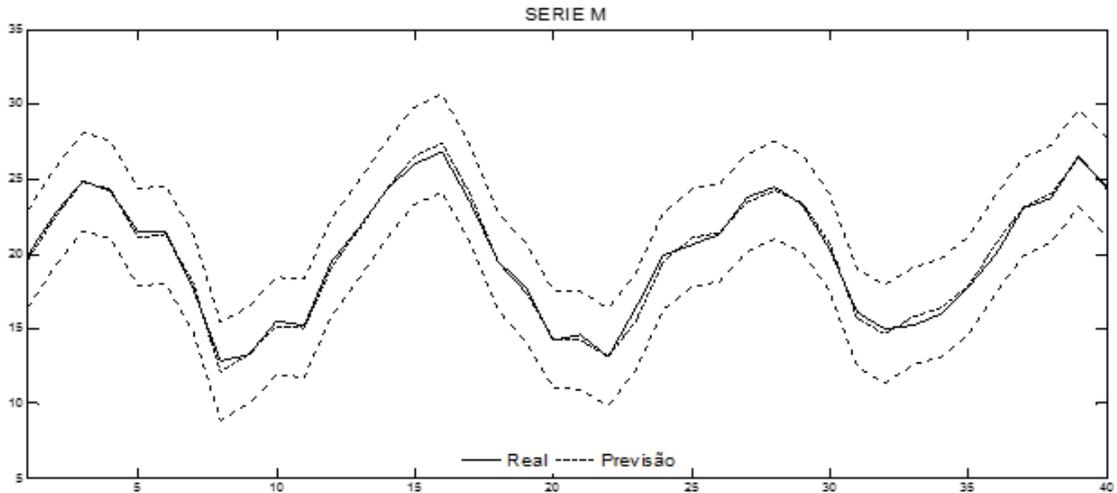


Figura 24: Validação do modelo de previsão da Série M

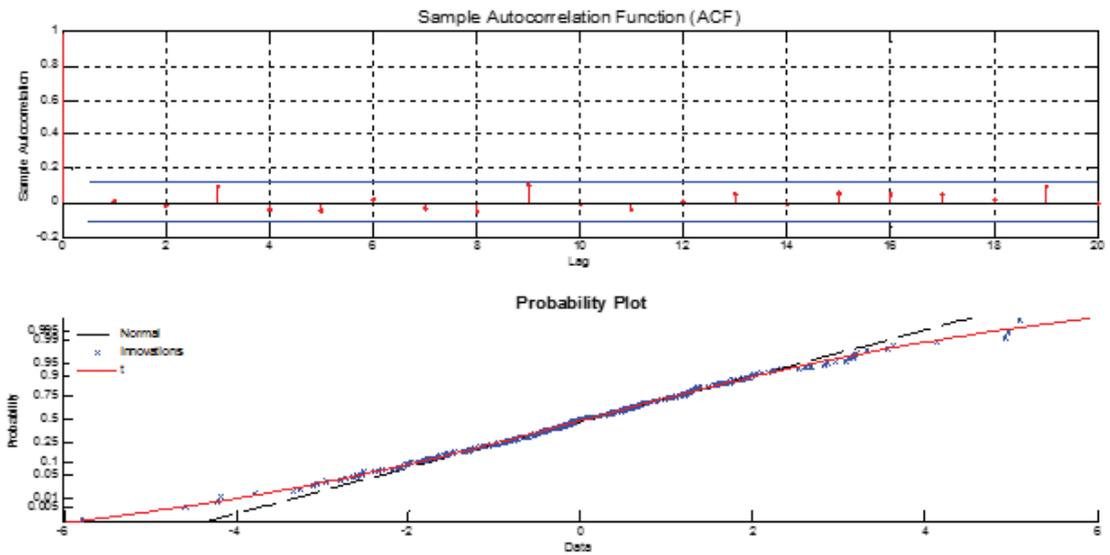


Figura 25: Análise dos resíduos do modelo da série M

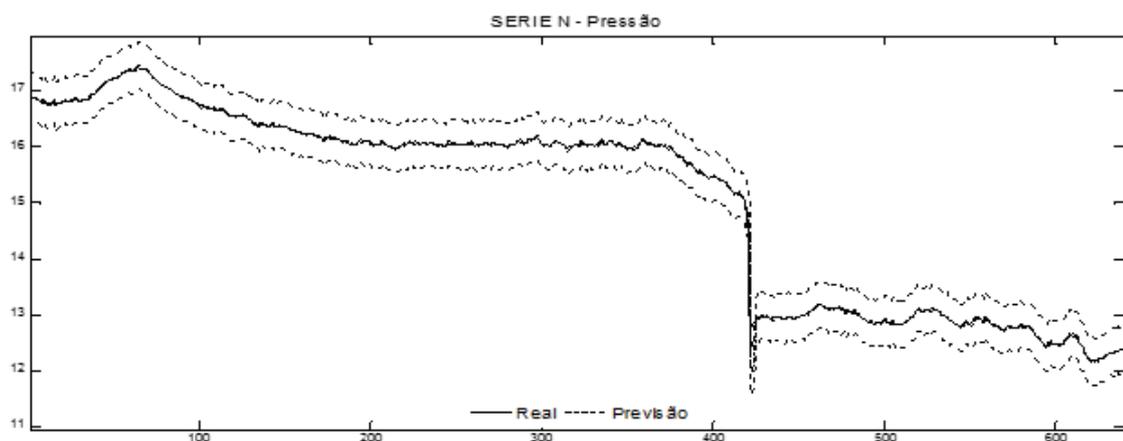


Figura 26: Validação do modelo para os dados de Pressão da série N

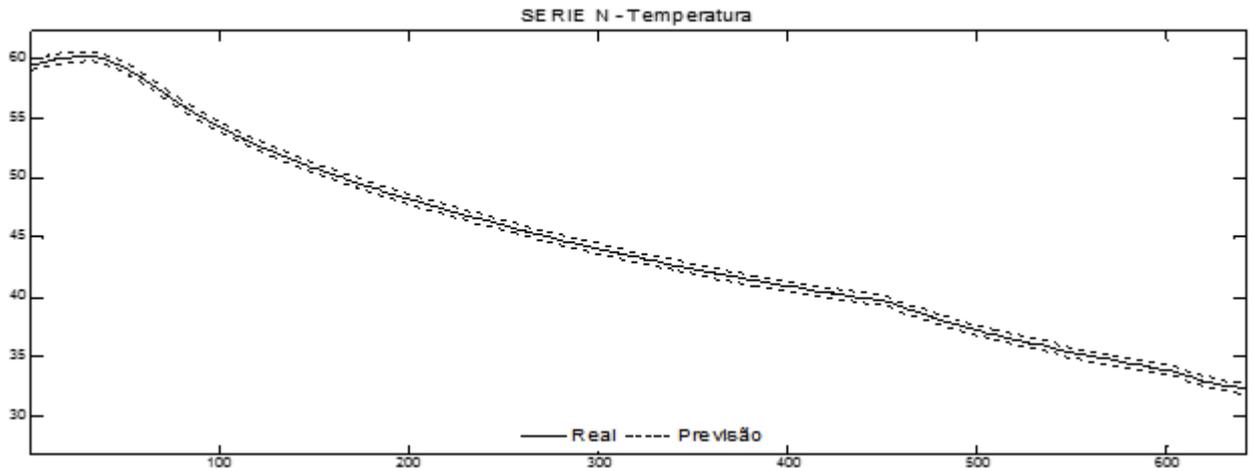


Figura 27: Validação do modelo para os dados de Temperatura da Série N

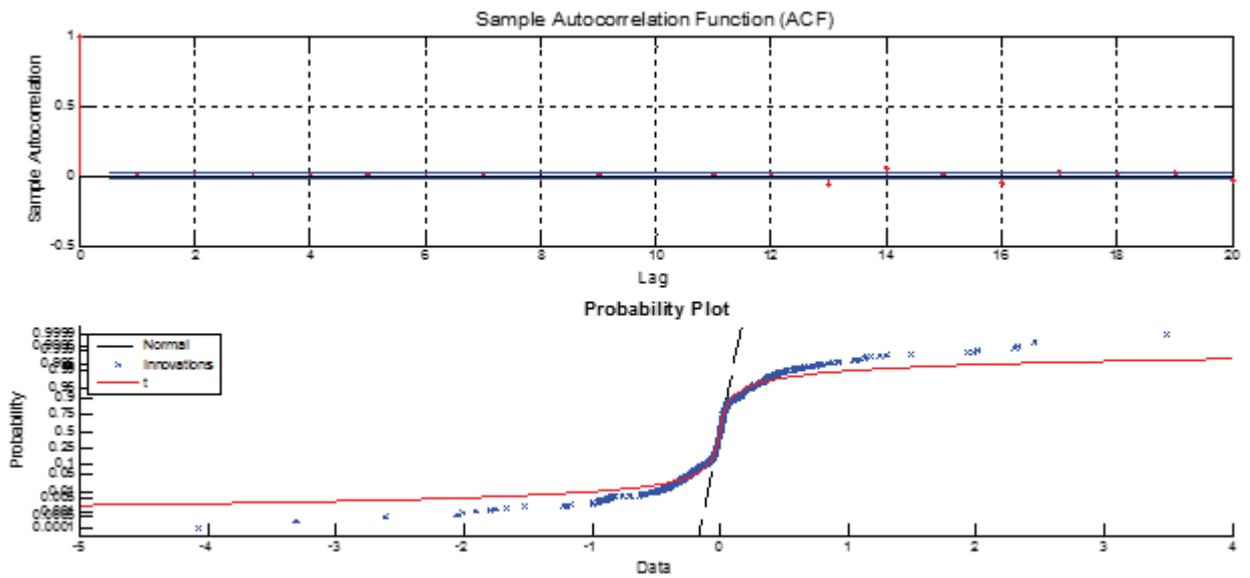


Figura 28: Análise de resíduos para o modelo da série N

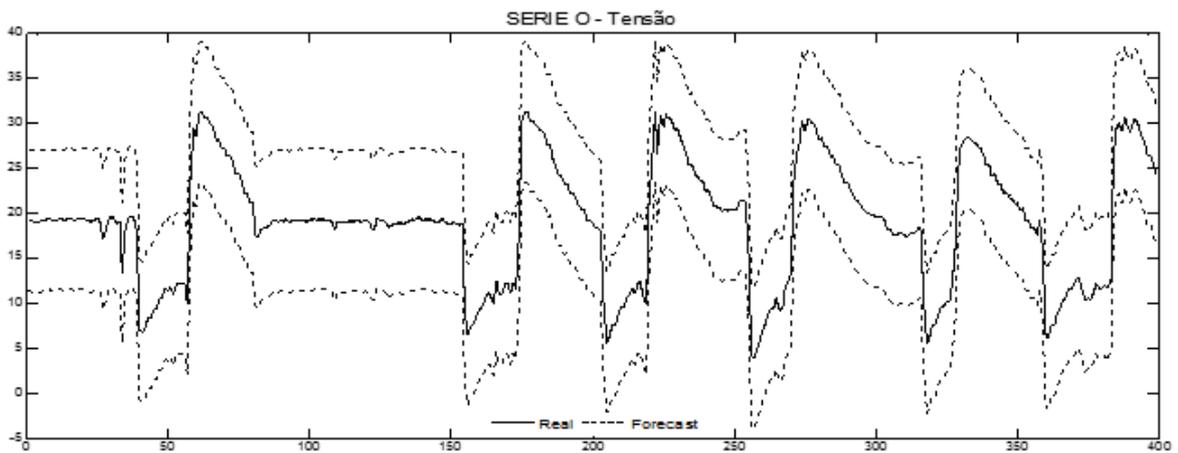


Figura 29: Validação do modelo para os dados de tensão da série O

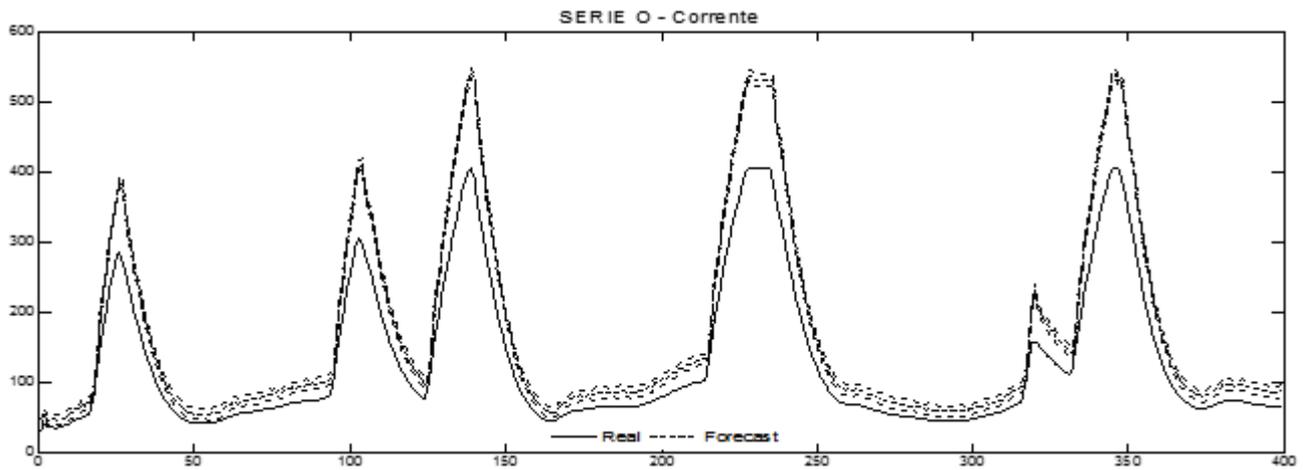


Figura 30: Validação do modelo para os dados de corrente da Série O

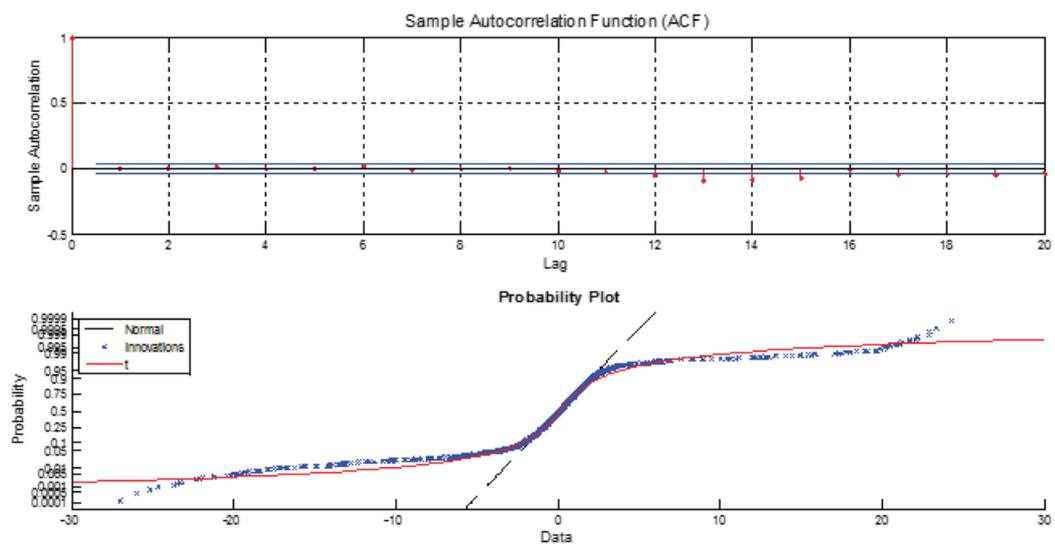


Figura 31: Análise de resíduos para Série O

5 Conclusões

Este capítulo traz comparações entre o objetivo proposto e os resultados alcançados. Além disso, aborda as principais vantagens e desvantagens encontradas na simulação realizada. Traz ainda sugestões para trabalhos futuros.

5.1 Considerações Iniciais

Neste ponto, é importante ressaltar que o objetivo inicial do trabalho, o de propor uma nova metodologia para obtenção de modelos de séries temporais, foi alcançado.

A seguir serão apresentadas conclusões referentes aos principais resultados obtidos durante o estudo, seguindo-se de sugestões para trabalhos futuros trabalhos relacionados ao tema escolhido.

5.2 Conclusões

No início deste trabalho, o objetivo principal proposto foi apresentar uma metodologia capaz de identificar modelos quantitativos parcimoniosos consistentes e representativos para a previsão da média de séries temporais lineares. Além disso, apesar de não ser o foco do trabalho foram realizadas análises de transformações de séries através das funções de autocorrelação e autocorrelação parcial para encontrar padrões que não permitiam o comportamento estacionário das séries.

Além das transformações, outra técnica aplicada no trabalho foi as análises das componentes principais. Essa se mostrou eficiente quando aplicada em séries que estão correlacionadas entre si e apresentou os mesmos resultados dos que os obtidos em análises das séries separadamente.

Foi elaborado um algoritmo na linguagem Matlab para estimar coeficientes de modelos propostos por matrizes experimentais construídas através do planejamento

Plackett-Burman. Através do algoritmo foi possível calcular o logaritmo natural da variância e utilizá-lo como resposta para se fazer análises estatísticas e definir o melhor modelo para as séries.

As previsões e análises de resíduos das séries modeladas mostrou que a metodologia pode resultar em modelos eficientes, porém, deve-se implantar melhorias em muitos aspectos, um deles é a otimização computacional do programa, pois dependendo da quantidade de dados utilizados o tempo de ciclo do algoritmo se torna muito elevado. O teste de Ljung-Box aplicado aos resíduos, para algumas séries demonstrou a existência de uma estrutura de autocorrelação, o que implica na possibilidade de que algumas características das séries podem não ter sido modeladas. Isso pode ter sido resultado da utilização da função de máxima verossimilhança inadequada, esse problema é verificado durante a análise dos resíduos quando se nota que a distribuição dos mesmos não seguem uma distribuição normal como deveriam.

Durante a aplicação da metodologia para modelagem das séries genéricas foram encontrados alguns problemas durante a identificação dos efeitos significantes devido a existência de modelos equivalentes, principalmente quando se trabalhou com modelos de ordens maiores. À medida que a ordem dos modelos crescia o aparecimento de modelos equivalentes aumentavam o que dificulta encontrar o modelo exato que foi construído.

No início do trabalho, decidiu-se utilizar os arranjos fatoriais, porém, foi identificada a significância de algumas interações entre fatores. Como se tratava de modelos lineares e o objetivo inicial do trabalho era o de identificar modelos com os mesmos padrões dos modelos Box-Jenkins decidiu-se descartar os efeitos das interações e utilizar os arranjos Plackett-Burman, visto que economizaria tempo computacional com a redução do número de experimentos.

Por fim, pode-se concluir que a metodologia em alguns casos teve resultados mais eficientes do que a técnica convencional, porém necessita de um estudo mais aprofundado em alguns pontos. A seguir serão apresentadas as sugestões para trabalhos futuros.

5.3 Recomendações Para Trabalhos Futuros

1. Adequar a metodologia proposta para construção de modelos de séries temporais não lineares.

2. Adequar a metodologia para modelos com ordens superiores aos estudados.
3. Estudar a utilização de outros tipos de arranjos da técnica de delineamento e análise de experimentos.
4. Estudar as interações entre os coeficientes.
5. Encontrar um método que otimize a definição dos coeficientes através da função da máxima verossimilhança.
6. Aprofundar o estudo com relação a função de máxima verossimilhança que deverá ser aplicada de acordo com a distribuição dos dados estudados.

Referências

- AVIV, Y. **Gaining benefits from joint forecasting and replenishment processes: The case of auto-correlated demand**. *Manufacturing & Service Operations Management*, INFORMS, v. 4, n. 1, p. 55–74, 2002.
- BALESTRASSI, P. P.; POPOVA, E.; PAIVA, A.; LIMA, J. M. **Design of experiments on neural network's training for nonlinear time series forecasting**. *Neurocomputing*, Elsevier, v. 72, n. 4, p. 1160–1178, 2009.
- BERNARDI, J.; LACERDA, L.; DÓREA, J.; LANDIM, P.; GOMES, J.; ALMEIDA, R.; MANZATTO, A.; BASTOTHERS. **Aplicação da análise das componentes principais na ordenação dos parâmetros físicos-químicos no alto rio madeira e afluentes**, Amazônia ocidental. *Geochimica Brasiliensis*, v. 23, n. 1, p. 79–90, 2012.
- BOX, G. E.; PIERCE, D. A. **Distribution of residual autocorrelations in autoregressive integrated moving average time series models**. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 65, n. 332, p. 1509–1526, 1970.
- BROCKWELL, P. J.; DAVIS, R. A. **Time Series: Theory and Methods**. 2. ed. [S.I.]: Springer, 1990.
- BROCKWELL, P. J.; DAVIS, R. A. **Introduction to Time Series and Forecasting**. 2. ed. [S.I.]: Taylor & Francis US, 2002.
- BYRON, R. P.; ASHENFELTER, O. **Predicting the quality of an unborn grange**. *Economic Record*, Wiley Online Library, v. 71, n. 1, p. 40–53, 1995.
- CARDOSO, M. M. **Simulação de Modelos Garch Para Séries Temporais Univariadas de Demanda de Energia Elétrica Para Consumidores Livres em Regime de Curto Prazo**. Dissertação (Dissertação (Mestrado em Engenharia de Produção)) — Universidade Federal de Itajubá, 2005.
- CASTRO, F. C. C. **Estudo de Series Temporais Aplicado a Perfis de Poços de Petróleo**. Dissertação (Mestrado) — Universidade Estadual de Campinas, 1995.
- CHATFIELD, C. **Time-series Forecasting**. [S.I.]: Chapman and Hall/CRC, 2000.
- CHEN, C.; LIU, L.-M. **Forecasting time series with outliers**. *Journal of Forecasting*, Wiley Online Library, v. 12, n. 1, p. 13–35, 1993.

- ENGLE, R. F. ***Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation***. *Econometrica: Journal of the Econometric Society*, JSTOR, v. 50, p. 987–1007, 1982.
- HAMILTON, J. D. ***Time Series Analysis***. [S.I.]: Princeton University Press, 1994.
- HAYASHI, F. ***Econometrics***. [S.I.]: Prince, 2000.
- HYNDMAN, R. J. ***Business Forecasting Methods***. 2009. Acesso: Fevereiro de 2012.
- INVESTOPEDIA. Acesso em 22 de março de 2012. Disponível em: <<http://www.investopedia.com/terms/v/volatility.asp>>.
- KIRCHGASSNER, G.; WOLTERS, J.; HASSLER, U. ***Introduction to Modern Time Series Analysis***. [S.I.]: Springer, 2012.
- MAKRIDAKIS, S. ***Forecasting Methods and Applications***. [S.I.]: John Wiley & Sons Inc, 1998.
- MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; MCGEE, V. E. ***Fundamentals of Quantitative Forecasting***. [S.I.]: John Wiley & Sons, 1983.
- MARTINS, P. G.; LAUGENI, F. P. ***Administração da Produção***. [S.I.]: Saraiva, 1999.
- MILESKI, A. ***Análise de Métodos de Previsão de Demanda Baseados Em Séries Temporais em Uma Empresa do Setor de Perfumes e Cosméticos***. Tese (Dissertação (Mestrado)) — Universidade Católica do Paraná, 2007.
- MILONE, G.; ANGELINI, F. ***Estatística Aplicada***. [S.I.]: Atlas, 1995.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. ***Introduction to Time Series Analysis and Forecasting***. [S.I.]: Wiley, 2008.
- MOREIRA, P. H. V.; FIGUEIREDO, M. R. B.; PINHEIRO, E. da S.; BORDALO, A. de C. ***Aplicação de Séries Temporais na Previsão da Média Mensal da Taxa de Câmbio do Real Para o Dólar Comercial de Compra Usando o Modelo de Holt***. 2010.
- NELSON, H. L.; GRANGER, C. W. J. ***Experience with using the box-cox transformation when forecasting economic time series***. *Journal of Econometrics*, Elsevier, v. 10, n. 1, p. 57–69, 1979.
- POLLOCK, D. S. G.; GREEN, R. C.; NGUYEN, T. ***Handbook of Time Series Analysis, Signal Processing and Dynamics***. [S.I.]: Academic Press, 1999.
- RUPPERT, D. ***Statistics and data analysis for financial engineering***. [S.I.]: Springer, 2011.

SÁFADI, T.; NOGUEIRA, D. A. **Previsão de Preços na Bovinocultura de Corte**. 2000.

SILVEIRA, D. **O Impacto da Previsão de Demanda nas Operações de Campo dos Vendedores em Empresa de Venda Direta**. 2002. Acesso: Fevereiro.

TOMÉ, E. A.; LATORRE, M. R. D. d. O. **Tendência da Mortalidade Infantil no Município de Guarulhos: Análise do Período de 1971 a 1998**. [S.l.]: Rev Bras Epidemiol, 2001.

TSAY, R. S. **Analysis of Financial Time Series**. 2005.

TUBINO, D. F. **Manual de Planejamento e Controle da Produção**. [S.l.]: Atlas, 2000.

WEI, W. W. S. **Time Series Analysis: Univariate e Multivariate Methods**. [S.l.]: Pearson Addison wesley, 2006.

WENER, L.; RIBEIRO, J. L. D. **Previsão de demanda: Uma aplicação dos modelos box-jenkins na Área de assistência técnica de computadores pessoais**. Gestão & Produção, v. 10, n. 1, p. 47–67, 2003.

ZAGRAJEK, J. N.; WERON, R. **Modeling electricity loads in California: Arma models with hyperbolic**. Signal Processing, v. 82, p. 1903–1915, 2002.