UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
DEPARTAMENTO DE ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA E MODELAGEM QUANTITATIVA

PREVISÃO DO PREÇO DOS PRINCIPAIS GRÃOS DO RIO GRANDE DO SUL

MONOGRAFIA DE ESPECIALIZAÇÃO

Analice Marchezan

Santa Maria, RS, Brasil
2007

PREVISÃO DO PREÇO DOS PRINCIPAIS GRÃOS DO RIO GRANDE DO SUL

por

Analice Marchezan

Monografia apresentada ao curso de Especialização em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa.

Orientador: Prof. Dr. Adriano Mendonça Souza

Santa Maria, RS, Brasil

2007

Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Departamento de Estatística
Programa de Pós-Graduação em Estatística e Métodos Quantitativos

A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

PREVISÃO DO PREÇO DOS PRINCIPAIS GRÃOS DO RIO GRANDE DO SUL

Elaborada por
Analice Marchezan

Como requisito parcial para a obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa

COMISSÃO EXAMINADORA:

Adriano Mendonça Souza, Dr. (UFSM)
(Presidente/Orientador)

Roselaine Ruviaro Zanini, Drª. (UFSM)

Luciane Flores Jacobi, Drª. (UFSM)

Santa Maria, 17 de Setembro de 2007.

AGRADECIMENTO

Ao Orientador, Adriano Mendonça Souza, pela compreensão, estímulo e compartilhamento de conhecimentos;

À Emater/RS – ASCAR por fornecer os dados para esta pesquisa;

E a todas as pessoas que, de uma forma ou outra, partilharam deste momento com dedicação e paciência.

RESUMO

Monografia de Especialização
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria

PREVISÃO DO PREÇO DOS PRINCIPAIS GRÃOS DO RIO GRANDE DO SUL

AUTORA: ANALICE MARCHEZAN
ORIENTADOR: ADRIANO MENDONÇA SOUZA
Data e Local da Defesa: Santa Maria, 17 de Setembro de 2007.

Este trabalho tem por objetivo realizar previsões para o ano de 2007, referente ao preço das principais culturas das Lavouras Temporárias desenvolvidas no Estado do Rio Grande do Sul. Os dados-base para a definição dos principais grãos a serem analisados foram a média anual, de 1995 a 2006, da quantidade produzida, e as previsões são refentes aos preços desses grãos, de janeiro de 1995 a dezembro de 2006, no estado gaúcho. A importância do estudo está no fato de a agricultura estar mostrando-se receptiva às tecnologias, bem como aos recursos que agregam valores ao produto, tornando a atividade cada vez mais competitiva. A viabilização da atividade, com a obtenção de lucro satisfatório, está sempre em risco, em decorrência de fatores controláveis e incontroláveis, que definem a produção agrícola. Logo, tendo o conhecimento dessas previsões, os gaúchos dispõem de informações para desenvolverem um planejamento através das expectativas de lucros e benefícios esperados, agregando valor ao seu trabalho. Para realizar as previsões do preço das culturas de soja, arroz e milho no Estado do Rio Grande do Sul, foi utilizada Metodologia Box & Jenkins. Para as culturas destacadas, a metodologia captou vários modelos significativos, mas através dos Critérios Penalizadores, AIC, SBC e Erro Quadrado Médio, foi possível escolher o melhor modelo para cada série, sendo o modelo *ARIMA (0,1,2)(0,1,0)* para a soja; o modelo S*ARIMA(1,1,0)(1,0,1)* para o arroz e o Modelo *ARIMA (2,1,1)(0,1,0)* para o milho.

Palavras-Chave: Box & Jenkins; previsão; grãos; preço; arroz; soja; milho.

ABSTRACT

Monograph of Specialization
Post-graduation Program in Statistics and Quantitative Modeling
Federal University of Santa Maria

FORECAST ABOUT THE MAIN GRAIN PRICES IN Rio Grande do Sul

Author: ANALICE MARCHEZAN
Advisor: ADRIANO MENDONÇA SOUZA
Place and local of defense: Santa Maria, September 17th, 2007.

This work deals with forecasts to the year of 2007 in relation to the main cultivations prices of temporary farming developed in Rio Grande do Sul state. The database for the definition of the main grains to be analyzed were selected through the annual average from 1995 to 2006 of the produced amount. The forecasts are referred to the grain prices from January 1995 to December 2006 in the 'gaúcho' state. This study is important since agriculture is likely to be receptive to new technologies in addition to resources that attribute aggregated values for products, turning the activity much more competitive. The viability of the activity, with satisfactory profit, is always in risk due to controllable and uncontrollable factors that define the farming production. Having knowledge about these forecasts, people who live in Rio Grande do Sul have tools in order to develop their planning through gain and benefit expectations, aggregating value to their work. To accomplish the forecasts in relation to prices of soybean, rice and sweetcorn in RS state, one did use of the Box & Jenkins methodology. To such cultivations, the methodoly showed many significant models, however, through the penalizing criterias, AIC e SBC, in addition the square average error, it was possible to choose the best model for each serie, being the model *ARIMA (0,1,2)(0,1,0)* to soybean; the model S*ARIMA (1,1,0)(1,0,1)* to rice and the model *ARIMA (2,1,1)(0,1,0)* to sweetcorn.

Key-words: Box & Jenkins; forecast; grains; prices; soybean; rice; sweetcorn.

LISTA DE FIGURAS

FIGURA 1 - Filtro Linear com entrada a_t, saída Z_t e Função Transferência $\Psi(B)$ 27

FIGURA 2 - Comportamento das funções de autocorrelação e autocorrelação parcial de um modelo AR (1) 32

FIGURA 3 - Comportamento das funções de autocorrelação parcial e autocorrelação de um modelo MA (1) 34

FIGURA 4 - Comportamento da série do preço da Soja, para o período de janeiro de 1995 a dezembro de 2006 51

FIGURA 5 - Funções de Autocorrelação da Série Original do Preço da Soja 51

FIGURA 6 - Série Original e Diferenciada do Preço da Soja 52

FIGURA 7 - Função de Autocorrelação e Autocorrelação Parcial Residual do Preço da Soja 54

FIGURA 8 - Série Original e a Previsão do Preço da Soja para o ano de 2007 55

FIGURA 9 - Comportamento do preço do Arroz, para o período de janeiro de 1995 a dezembro de 2006 56

FIGURA 10 - Funções de Autocorrelação da Série Original do Preço do Arroz 57

FIGURA 11 - Série Original e Diferenciada do Preço do Arroz 58

FIGURA 12 - Função de Autocorrelação e Autocorrelação Parcial Residual do Preço do Arroz 59

FIGURA 13 - Série Original e a Previsão do Preço do Arroz para o ano de 2007 60

FIGURA 14 - Comportamento do preço do Milho, para o período de janeiro de 1995 a dezembro de 2006 61

FIGURA 15 - Funções de Autocorrelação da Série Original do Preço do Milho 62

FIGURA 16 - Série Original e Diferenciada do Preço do Milho 63

FIGURA 17 - Função de Autocorrelação e Autocorrelação Parcial Residual do Preço do Milho 64

FIGURA 18 - Série Original e a Previsão para o ano de 2007, do Preço do Milho 65

LISTA DE TABELAS E QUADRO

TABELA 1 - Padrões Teóricos das FAC e FACP..... 37

TABELA 2 - Média, em tonelada, da quantidade produzida nas lavouras temporárias de 1995 a 2006. 49

TABELA 3 - Modelos significativos para a Série do Preço da Soja no Estado do Rio Grande do Sul. 53

TABELA 4 - Previsão da Serie Original da Soja, no RS, para o Ano de 2007..... 54

TABELA 5 - Modelos significativos para a Série do Preço do Arroz no Estado do Rio Grande do Sul..... 59

TABELA 6 - Previsão da Serie Original do Arroz, no RS, para o ano de 2007..... 60

TABELA 7 - Modelos significativos para a Série do Preço do Milho no Estado do Rio Grande do Sul..... 63

TABELA 8 - Previsão da Serie Original do Milho, no RS, para o Ano de 2007..... 64

QUADRO 1 - Quadro representativo dos Estágios do Ciclo Iterativo de Box & Jenkins..... 40

SUMÁRIO

1 INTRODUÇÃO 9
1.1 Tema da Pesquisa 11
1.2 Justificativa e Importância 11
1.3 Objetivos 12
1.3.1 Objetivo Geral 12
1.3.2 Objetivos Específicos 12
1.4 Delimitação da Pesquisa 12
2 REVISÃO DE LITERATURA 13
2.1 Análise de Séries Temporais 13
2.1.1 Série Temporal 17
2.1.2 Modelos para representar uma Série Temporal 18
2.1.3 Componentes de uma Série Temporal 20
2.2 Metodologia Box &Jenkins 24
2.2.1 Modelos Estacionários 27
2.2.2 Modelos Não-Estacionários 38
2.2.3 Modelos Sazonais 46
3 METODOLOGIA 48
4 RESULTADOS E DISCUSSÕES 50
4.1 Soja 50
4.2 Arroz 57
4.3 Milho 62
5 CONCLUSÃO 67
6 BIBLIOGRAFIA 68
6.1 Referências Bibliográficas 68
6.2 Bibliografias Consultadas 69

1 INTRODUÇÃO

Antes mesmo da Revolução Industrial e do processo de mecanização da atividade agrícola, os agricultores já se mostravam capazes de reconhecer a variabilidade espacial de certas características das áreas cultivadas. Até então, o uso de trabalho braçal ou de tração animal permitia aos agricultores tratar áreas com menor ou maior fertilidade, com infestação de pragas, doenças e plantas daninhas de forma diferenciada.

Atualmente, identificam-se áreas manejadas nas regiões coloniais gaúchas, com propriedades de topografia acidentada e com pequenas áreas de cultivo, sendo possível observar uma agricultura diversificada, em que a alocação das culturas é feita de acordo com as características de cada região. O produtor nessas áreas, conhece o potencial e pratica o manejo específico em cada situação. Esse tipo de manejo, que prescinde de equipamentos de alta tecnologia, pode ser considerado, também, uma forma de agregar maior precisão à atividade agrícola, embora empírica e muitas vezes inconsistente. A variabilidade dos fatores de produção está associada a múltiplas causas, desde a variabilidade climática até a representada pelo ambiente em torno de uma única semente que é depositada no solo.

O Rio Grande do Sul é um Estado modelo do Brasil, pela diversidade econômica que é capaz de aglomerar. É nele que se concentra um enorme número de empresas e indústrias, sejam elas de pequeno, médio ou grande porte. Isso porque o Estado consegue disponibilizar mão-de-obra qualificada para as mais diversas atividades, bem como adaptar-se facilmente aos avanços tecnológicos.

Esses avanços são percebidos em todos os setores; por isso, é que o povo gaúcho busca especializar-se, como forma de garantir um bom emprego. Eles também atingem a agricultura, fazendo com que os alimentos cheguem ao seu destino com melhor qualidade. Nesse sentido, o alto investimento acaba sendo compensado pela redução de perdas, agilidade no plantio e colheita, bem como pelo reconhecimento da produção com qualidade.

A diversidade, a qualidade e a quantidade do desenvolvimento agrícola gaúcho ocorrem por haver, no Estado, um clima diversificado e a não resistência aos avanços tecnológicos, porque, nos tempos atuais, é preciso ter incentivos agrícolas, boas máquinas, irrigação e combate a pragas, a fim de se obter uma boa safra. Mas, para se ter acesso a todos esses implementos, os gaúchos necessitam organizar-se, prever quais serão os seus benefícios, pesando os lucros e prejuízos.

Dentre as ferramentas que visam a auxiliar indiretamente os agricultores estão as análises estatísticas, as quais mostram uma síntese da realidade da agricultura. Nesse sentido, apontam aspectos importantes na tomada de decisões dos agricultores, das empresas e do próprio Estado. Tais decisões podem ser embasadas em análises descritivas e previsões futuras.

Deve-se ter bem claro que a análise descritiva gera uma síntese dos dados que se deseja analisar, enquanto que uma previsão é uma manifestação relativa a sucessos desconhecidos em um futuro determinado. Para Morettin (1989), a previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisão, visando a atingir determinados objetivos.

De acordo com Wheelwrigth (1985 *apud* Mueller, 1996), a maioria dos métodos de previsão de séries temporais se baseia na suposição de que observações passadas contêm todas as informações sobre o padrão de comportamento da série temporal, sendo esse padrão recorrente no tempo. O propósito dos métodos de previsão consiste em distinguir o padrão de qualquer ruído que possa estar contido nas observações e, então, usá-lo para prever os valores futuros da série temporal. Assim, pela identificação desse componente, a previsão para períodos de tempo subseqüentes ao observado pode ser desenvolvida e representada por um modelo matemático.

O destaque é atribuído ao modelo de George Box e Gwilyn Jenkins, devido, principalmente, a sua fundamentação teórica, sendo a princípio capaz de manipular séries temporais de qualquer natureza. O método de Box & Jenkins consiste na busca de um modelo ARIMA (Auto *Regressive Integrate Moving Average)* e na literatura em português por Auto-regressivos Integrados de Médias Móveis, que represente o processo estocástico gerador da série temporal, a partir de um modelo ARIMA, aplicável na descrição de séries temporais estacionárias, estendendo esse conceito para séries temporais não-estacionárias.

Portanto, os modelos ARIMA são modelos matemáticos que almejam captar o comportamento da correlação seriada, ou autocorrelação entre os valores da série temporal, e, com base nesse comportamento, realizar previsões futuras.

Diante disso, esta pesquisa realizará a previsão do preço dos principais grãos produzidos no Rio Grande do Sul, a fim de auxiliar os agricultores gaúchos na tomada de decisões e na criação de políticas de incentivo. Essas previsões podem servir de suporte para a organização do próprio governo que, ao detectar necessidades no Estado, poderá desenvolver os incentivos necessários para sanar problemas enfrentados.

1.1 Tema da Pesquisa

O tema da pesquisa estará abordando Séries Temporais, mais precisamente o método de Box & Jenkins, a fim de efetuar previsões referente ao preço dos principais grãos produzidos no Estado do Rio Grande do Sul para o ano de 2007. Esses grãos estão relacionados às Lavouras Temporárias, que são aquelas culturas que precisam ser replantadas após cada colheita, por possuírem um curto ciclo vegetativo.

1.2 Justificativa e Importância

A economia gaúcha é formada por um conjunto de atividades, dentre as quais a agricultura também tem a sua contribuição. Das várias culturas bem desenvolvidas no Estado, optou-se pelas principais produções das Lavouras Temporárias. A importância do estudo está no fato de que a agricultura está mostrando-se receptiva às tecnologias, bem como aos recursos que agregam valores ao produto, tornando a atividade cada vez mais competitiva. Isso também exige do produtor maior nível de especialização, capacidade de gerenciamento e profissionalismo. Os produtores, além de administradores, cada vez mais têm de assumir a função de produtores pesquisadores de suas áreas, atuando diretamente na coleta de informações, interagindo com novas técnicas e tomando decisões eficazes de manejo. A viabilização da atividade, com a obtenção de lucro satisfatório, está sempre em risco, em decorrência de fatores controláveis e incontroláveis que definem a produção agrícola. Portanto, a aplicação dos recursos disponíveis de forma eficiente é indispensável como garantia de sucesso. Para que isso possa acontecer, é crucial a obtenção de informações sobre os fatores de produção que interagem na lavoura, e sobre como eles podem ser maximizados ou previstos. Logo, prevendo o preço das principais culturas desenvolvidas no Estado, os gaúchos podem desenvolver um planejamento através das expectativas criadas em torno do seu trabalho.

1.3 Objetivos

1.3.1 Objetivo Geral

Realizar previsões, a curto prazo, do preço dos principais grãos do Estado do Rio Grande do Sul, por meio da metodologia de Box & Jenkins, a fim de fornecer o conhecimento prévio do comportamento dos preços da cultura investigada.

1.3.2 Objetivos Específicos

- Fazer uma análise descritiva dos dados;
- Identificar as principais culturas do Estado do Rio Grande do Sul;
- Entender o comportamento gerador da série do preço das principais culturas do Estado;
- Fazer previsão dos preços das principais culturas do Estado gaúcho para o ano de 2007.

1.4 Delimitação da Pesquisa

A agricultura brasileira, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), possui duas divisões:

Lavouras Permanentes: são culturas que não precisam ser replantadas após cada colheita, pelo seu longo ciclo vegetativo. É possível elencar, como exemplo, as lavouras de laranja, limão, café, entre outros.

Lavouras Temporárias: são culturas que possuem um curto período vegetativo, necessitando ser replantadas após cada safra. Tem-se, como exemplo, as lavouras de soja, milho, arroz e feijão.

Devido à enorme abrangência de culturas que a agricultura engloba, optou-se como delimitação desta pesquisa, o comportamento do preço das principais culturas das Lavouras Temporárias. Um fator delimitante desta pesquisa o uso de modelos de Séries Temporais univariados, não se estendendo a outros tipos de modelos, como por exemplo, os modelos fracionários e os de análise de regressão.

2 REVISÃO DE LITERATURA

Neste capitulo, estarão sendo apresentadas definições de Séries Temporais e a metodologia de Box & Jenkins, fundamentais para previsão dos preços dos principais grãos produzidos no Estado do Rio Grande do Sul.

2.1 Análise de Séries Temporais

Uma série temporal é a coleção das observações feitas seqüencialmente ao longo do tempo. As características mais importantes dessas observações vizinhas é sua dependência, a qual se tem por interesse analisar e modelar, considerando a ordem dos dados um fator crucial. Vale observar, também, que o tempo pode ser substituído por outra variável, como espaço e profundidade. A técnica de análise de Séries Temporais vem sendo utilizada com freqüência em várias áreas de pesquisa e, como a maior parte de procedimentos estatísticos foi desenvolvida para analisar observações independentes, o seu estudo requer o uso de técnicas específicas, pois os dados surgem dos mais diversos campos de conhecimento, como Ciências Sociais e Políticas, Sociologia, História, Meio Ambiente, Economia (preços diários de ações, taxas de desemprego, produção industrial, etc.), Medicina (eletrocardiograma, eletroencefalograma, etc.), Epidemiologia (números mensais de meningite, etc.), Metereologia (precipitação pluviométrica, temperatura diária, velocidade do vento), entre outros.

Do exposto, podem-se citar as seguintes características de uma Série Temporal:

- Observações correlacionadas são mais difíceis de analisar, além de essa análise requerer técnicas específicas;
- A ordem temporal das observações deve ser considerada;
- Fatores complicadores como tendência e variação sazonal ou cíclica podem ser difíceis de estimar e remover;
- A seleção de modelos pode ser bastante complicada, e as ferramentas podem ser de difícil interpretação;
- É mais difícil lidar com observações perdidas e dados discrepantes, devido à natureza seqüencial.

Com base nessas particularidades, os principais objetivos da análise de séries temporais consiste em:

- Compreender o mecanismo gerador da série:
  - Descrever efetivamente o comportamento da série;
  - Encontrar periodicidades na série;
  - Tentar obter razões para o comportamento da série (possivelmente através de variáveis auxiliares);
  - Controlar a trajetória da série.

- Predizer o comportamento futuro da série, possibilitando:
  - Fazer planos a longo, médio e curto prazo;
  - Tomar decisões apropriadas.

Sabe-se que esses objetivos estão relacionados entre si, sendo possível a ocorrência quando o modelo encontrado é adequado. O pesquisador deve ter consciência de que o futuro é incerto. Nesse sentido, as previsões podem não ser perfeitas. Portanto, a análise de Séries Temporais também consiste em reduzir ao máximo os erros de previsão.

Para Souza (1989), a classe de fenômenos cujo processo observacional e conseqüente quantificação numérica geram uma seqüência de dados distribuídos no tempo é denominada *série temporal*.

Com relação ao número de observações necessárias para um estudo em séries temporais, Box *et al*. (1994, *apud* Barbiero, 2003, p. 37) sugerem ser suficiente para a aplicação dos métodos 50 a 100 observações, e Kendall (1976 *apud* Barbiero, 2003, p. 37) descreve que:

> (...) o número de observações em uma série temporal não é por si uma medida completa de informação. O que pode ser observado é que o acréscimo ou decréscimo de **n** auxilia no reconhecimento de algum comportamento nos dados.

Granger (1977) coloca como objetivo inicial da análise de séries temporais, a realização de inferências sobre as propriedades ou características básicas do mecanismo gerador do processo estocástico das observações da série. Assim, através da abstração de regularidades contidas nos fenômenos observáveis de uma série temporal, existe a possibilidade de se construir um modelo matemático como uma representação simplificada da realidade. Barbancho (1970).

Anderson (1971 *apud* MUELLER, 1996) considera que a natureza de uma Série Temporal e a estrutura de seu mecanismo gerador está relacionada com o intervalo de ocorrência das observações no tempo. Caso o levantamento das observações da Série possa ser feito a qualquer momento do tempo, a Série Temporal é dita *contínua*, sendo denotada por Z_t, segundo Granger (1977). Além disso, na maioria das séries, as observações são tomadas em intervalos de tempo discretos e eqüidistantes, conforme Nelson (1973).

Uma série temporal pode ser classificada como determinística ou estocástica. Diz-se que é determinística quando os futuros valores da série podem ser estabelecidos precisamente por uma relação funcional matemática do tipo $y = f(t)$; e estocástica quando seus valores futuros só puderem ser expostos em termos probabilísticos, uma vez que a Série está descrita por meio de uma relação funcional que envolve não só o tempo, mas também uma variável aleatória do tipo $y = f(t, \alpha)$, em que α é o tempo aleatório residual, cuja inclusão torna-se necessária quando não se consegue explicar completamente algum movimento irregular da Série através, unicamente, da relação matemática.

Para Barbancho (1970), a investigação do poder preditivo do modelo especificado como o mais adequado para explicar o mecanismo gerador das observações de uma Série Temporal é um processo empírico de verificação. Isso acontece porque são feitas comparações entre as previsões e as observações, a fim de confirmar a habilidade do modelo matemático em descrever a estrutura definida pelos dados da série temporal analisada. O autor ainda considera importante salientar que o caráter previsional de um determinado modelo não pode ser sustentado quando forem identificadas mudanças estruturais nas observações obtidas a partir da origem, em relação àquelas utilizadas para a elaboração do modelo.

Após a formulação do modelo matemático obtido pela seleção entre as alternativas de classes de modelos, identificadas como apropriadas para essa representação e subseqüente estimação de seus parâmetros, é possível utilizá-lo para testar alguma hipótese ou teoria a respeito do mecanismo gerador do processo estocástico e realizar a previsão de valores futuros da Série Temporal, conforme Granger (1977) e Nelson (1973).

Mueller (1996) descreve uma previsão como uma manifestação relativa a sucessos desconhecidos em um futuro determinado. Morettin e Toloi (1987) complementam que a previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisão, visando a atingir determinados objetivos.

Wheelwrigth (1985 *apud* MUELLER, 1996) define um método de previsão como sendo o conjunto de procedimentos usados no desenvolvimento de uma determinada previsão.

Os métodos de previsão de Séries Temporais baseiam suas previsões na extrapolação de características de observações passadas e no inter-relacionamento entre essas observações, fornecendo previsões acuradas se o futuro apresentar comportamento similar ao passado. A idéia manifestada pelo autor é que a maioria dos métodos de previsão de Séries Temporais se baseia na suposição de que observações passadas contêm todas as informações sobre o padrão de comportamento destas, sendo esse padrão recorrente no tempo. O propósito dos métodos de previsão consiste em distinguir o padrão de qualquer ruído que possa estar contido nas observações, e, então, usá-lo para prever os valores futuros da Série Temporal. Assim, pela identificação desse componente, a previsão para períodos de tempo subseqüentes ao observado pode ser desenvolvida.

Para o autor, há uma variedade enorme de métodos de previsão de Séries Temporais, cada qual com suas capacidades e limitações. Qualquer que seja a classificação desses métodos, é possível utilizar um número muito grande de métodos diferentes para descrever o comportamento de uma série particular. A seleção do método de previsão adequado depende de vários fatores, tais como o comportamento do fenômeno observável ou o conhecimento *a priori* que se tenha sobre a sua natureza e do objetivo da análise.

Para Souza (1989), os métodos de decomposição assumem que uma Série Temporal é constituída por um conjunto de componentes não-observáveis. Dessa forma, pela identificação das componentes individuais presentes no padrão básico da série histórica de dados (tendência, ciclo, sazonalidade e aleatoriedade), a extrapolação para o futuro pode ser realizada.

Dependendo do valor assumido pelo horizonte de previsão, Souza (1989) e Wheelwrigth (1985 *apud* MUELLER, 1996) classificam as previsões dos valores futuros de uma Série Temporal como de curto, médio ou longo prazo. Desse modo, diante da possibilidade de existência de diferentes horizontes de previsão, existem específicas técnicas distintas para encontrar os valores futuros de uma Série Temporal:

- *Previsão múltiplos passos*: essa abordagem, adotada para longos horizontes de previsão, procura identificar as tendências gerais e os pontos de inflexão mais relevantes na Série Temporal. Na *previsão múltiplos passos*, o conjunto de valores correntes é empregado na realização da previsão para determinado instante. Essa previsão é, então, introduzida entre as observações passadas, compondo, dessa forma, um novo conjunto de dados, sobre o qual será obtida a previsão do tempo subseqüente.
- *Previsão simples passo*: nessa técnica, não há incorporação de previsões aos dados utilizados para encontrar a próxima previsão, sendo esta independente dos valores

anteriormente previstos. A previsão é feita apenas para o período de tempo imediatamente posterior ao atual, a partir das observações da Série Temporal.

Souza (1989) garante que a otimalidade das previsões de uma Série Temporal somente é alcançada adotando como horizonte de previsão o instante de tempo imediatamente subseqüente à origem *t*.

Os modelos utilizados para a análise desta pesquisa serão os Modelos Box & Jenkins, por serem modelos matemáticos que visam a captar o comportamento da correlação seriada, ou autocorrelação entre os valores da Série Temporal, e, com base nesse comportamento, realizar previsões futuras. É uma das técnicas mais difundidas e descritas por esses autores na década de 70.

Ao analisar uma Série Temporal, a representação gráfica dos dados, seqüencialmente, ao longo do tempo, é fundamental, e pode revelar padrões de comportamento importantes, tais como tendência de crescimento ou decrescimento, ciclos, alterações estruturais, observações aberrantes, entre tantos outros facilmente identificados. Logo, o primeiro passo em uma análise temporal é a realização do gráfico da série original. Abaixo, seguem as definições fundamentais para a análise de uma Série Temporal:

2.1.1 Série Temporal

Segundo Morettin e Toloi (2004), Souza & Camargo (1996), Barros (2006), e tantos outros autores, uma Série Temporal, também podendo ser definida por uma Série Histórica, é uma seqüência de dados obtidos em intervalos regulares de tempo, durante um período específico, e que apresentam uma dependência seriada. A notação utilizada para representar uma Série Temporal Z, no momento t será Z_t, em que $t = 1,2,\ldots,n$ indica o tamanho da Série.

Considera-se que uma Série Temporal é uma amostra do processo. O conjunto de observações ordenadas no tempo pode ser:

Discreta (como o número de atendimentos diários em um Pronto Socorro, o número mensal de casos notificados de uma doença específica, etc.);

Contínua (como o registro de um eletrocardiograma de uma pessoa, o registro dos valores de temperatura e umidade ao longo do dia, etc.).

Na análise de uma Série Temporal, primeiramente, deseja-se modelar o fenômeno estudado para, a partir daí, descrever o comportamento da Série, fazer estimativas e, por último, avaliar quais os fatores que influenciaram o comportamento da Série, buscando definir

relações de causa e efeito entre duas ou mais Séries. Para tanto, há um conjunto de técnicas estatísticas disponíveis que dependem do modelo definido (ou estimado para a Série), bem como do tipo de Série analisada e o objetivo do trabalho.

O pesquisador que se utiliza de uma Série Temporal tem por objetivo estudar o processo gerador desta, fazer previsões em função dos seus valores passados e descrever o seu comportamento. Esse estudo pode ser feito no domínio do tempo, por meio das funções de autocorrelações, ou no domínio da freqüência, por meio de análise espectral.

Deve-se sempre levar em consideração que as previsões devem ser a curto prazo, visto que o nível de incerteza aumenta em proporção com o horizonte das previsões.

2.1.2 Modelos para representar uma Série Temporal

- Processo Estocástico

Os modelos utilizados para descrever Séries Temporais são Processos Estocásticos, isto é, processos controlados por leis probabilísticas. É uma maneira mais formal utilizada para dizer que uma Série Temporal é uma realização de um processo estocástico. Denomina-se um processo todas as possíveis curvas obtida no gráfico da Série Histórica. Essas curvas são chamadas de trajetórias.

Quando existe a semelhança entre partes de uma Série Temporal apenas nas suas propriedades médias, é necessária a sua descrição através de leis probabilísticas ou modelos. Assim, os valores possíveis de uma Série em um determinado tempo t são descritos por uma variável aleatória $X(t)$ e sua associação a uma distribuição de probabilidade. O comportamento da Série Temporal, para qualquer tempo t, pode ser descrito por um conjunto de variáveis aleatórias, em que t varia aos extremos do infinito $(-\infty;+\infty)$, fazendo com que as Séries sejam descritas com qualquer conjunto de tempo $t_1, t_2, ..., t_n$. Logo, um Processo Estocástico é o conjunto ordenado de variáveis aleatórias $\{X(t)\}$ associado a sua distribuição de probabilidade.

Morettin e Toloi (2004, p. 21) explica um Processo Estocástico através da definição:

Definição: Seja T um conjunto arbitrário. Um P*rocesso Estocástico* é uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória.

Nessas condições, um Processo Estocástico é uma família de variáveis aleatórias, que suporemos definidas num mesmo espaço de probabilidades (Ω, A, P). O conjunto T é

normalmente tomado como conjunto dos inteiros $Z = \{0, \pm 1, \pm 2, ...\}$ ou o conjunto dos reais $\Re$. Também, para cada $t \in T$, Z(t) será uma variável aleatória real.

- Ergodicidade

Para Souza & Camargo (1996, p. 22-23), um processo estocástico é dito ergódigo se apenas uma realização dele é o suficiente para se obter todas as estatísticas. Para esses autores, pode-se facilmente verificar que todo o processo ergódigo também é estacionário, visto que uma realização de um processo não estacionário não poderá conter todas as informações necessárias para a especificação do processo.

Os processos ergódigos embasam-se nas propriedades de que a média e a variância precisam ser constantes; e as funções de autocovariância e autocorrelação, independentes da origem dos tempos.

- Estacionariedade

Conforme Morettin e Toloi (2004), uma das suposições mais freqüentes que se faz a respeito de uma Série Temporal é que ela seja estacionária, ou seja, desenvolve-se no tempo, aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. Todavia, a maior parte das Séries que encontramos na prática apresenta alguma forma de não-estacionariedade. Assim, as Séries econômicas apresentam, em geral, *tendências.* O caso mais simples é aquele em que a série flutua ao redor de uma reta, com inclinação positiva ou negativa (tendência linear). Pode ocorrer, também, uma forma de não-estacionariedade explosiva, como o crescimento de uma colônia de bactérias.

Os autores destacam que a classe dos modelos ARIMA será capaz de descrever, de maneira satisfatória, Séries estacionárias e Séries não-estacionárias, mas que não apresentem comportamento explosivo. Esse tipo de não-estacionariedade é chamado *homogêneo*. A série pode ser estacionária, flutuando ao redor de um nível, por certo tempo, e depois mudar de nível e flutuar ao redor de um novo nível, e assim por diante, ou então mudar de inclinação, ou ambas as coisas.

Esses mesmos autores destacam que, como a maioria dos procedimentos de análise estatística de Séries Temporais supõe que estas sejam estacionárias, será necessário transformar os dados originais, até estacionarizá-lo se for necessário. A transformação mais

comum consiste em tomar diferenças sucessivas da Série original, até se obter uma Série estacionária. A primeira diferença de Z_t é definida por

$$\Delta Z_t = Z_t - Z_{t-1}, \tag{1}$$

a segunda diferença é

$$\Delta^2 Z_t = \Delta(\Delta Z_t) = \Delta(Z_t - Z_{t-1}), \tag{2}$$

ou seja,

$$\Delta^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2}. \tag{3}$$

De modo geral, a n-ésima diferença de Z_t é

$$\Delta^n Z_t = \Delta\left[\Delta^{n-1} Z_t\right] \tag{4}$$

Morettin e Toloi (2004) ainda ponderam que, em situações normais, será suficiente tomar uma ou duas diferenças para que a Série se torne estacionária.

2.1.3 Componentes de uma Série Temporal

Uma Série Temporal Z_t, $t = 1,2,...,n$ é composta por três componentes, sendo eles a Tendência (T_t), a Sazonalidade (S_t)e a variação aleatória denominada Ruído aleatório ou erro (a_t).

Os modelos podem ser construídos através do relacionamento das componentes de forma:

Aditivo:

$$Z_t = T_t + S_t + a_t \tag{5}$$

Ou multiplicativo:

$$Z_t = T_t . S_t . a_t \tag{6}$$

Para Morettin e Toloi (1986), a estrutura de uma Série Temporal é aditiva se S_t independe de T_t, e é multiplicativa se existe uma dependência entre as duas componentes. Um teste para verificar se o modelo é aditivo ou multiplicativo é baseado em um gráfico da amplitude sazonal contra a tendência anual. Se a reta ajustada a esses pontos for paralela ao eixo das abscissas, há uma indicação que a amplitude sazonal não depende da tendência e o modelo é aditivo; se a reta apresentar inclinação diferente de zero, temos uma dependência de

S_t sobre T_t e o modelo é multiplicativo. Além do gráfico, pode-se encontrar a reta dos mínimos quadrados e testar a hipótese de que a inclinação é zero.

- Tendência

Uma Série Temporal possui tendência quando suas observações apresentam um comportamento crescente ou decrescente, representados pelo aumento ou declínio gradual. Em outras palavras, a tendência existe quando a Série não é estacionária, com média e variância constante.

Segundo Morettin e Toloi (2004), existem vários métodos para se estimar a tendência. Os mais utilizados consistem em:

- Ajustar uma função do tempo, como um polinômio, uma exponencial ou outra função suave de *t;*
- Suavizar (ou filtrar) os valores da Série ao redor de um ponto, para estimar a tendência naquele ponto;
- Suavizar os valores da Série através de sucessivos ajustes de retas de mínimos quadrados ponderados;
- Tomar sucessivas diferenças. Em situações normais, será suficiente tomar uma ou duas diferenças para que a Série se torne estacionária, conforme equações (1) a (4).

Após a estimativa da tendência, para analisar os outros componentes, é necessário construir uma Série livre de tendência, em que o procedimento mais utilizado é analisar as sucessivas diferenças desta.

Um teste indicado para a verificação da tendência é o Teste do Sinal (Cox-Stuart), para testar as hipóteses:

H_0 : não existe tendência

H_1 : existe tendência

Para mais informações sobre esse teste, consultar Morettin e Toloi (1986, p.28-29).

- Sazonalidade

Define-se um fenômeno sazonal como aquele que ocorre regularmente, em períodos fixos de tempo, no decorrer de um ano. A componente sazonal representa as flutuações da Série, de acordo com algum fator determinístico.

Um teste indicado para a verificação da sazonalidade é o Teste não-pamétrico de Kruskal-Wallis, em que:

H_0 : não existe sazonalidade determinística;

H_1 : existe tendência sazonalidade determinística.

Para mais informações sobre esse teste, consultar Morettin e Toloi (1986, p.46-47).

- Ciclo

O ciclo apresenta um comportamento similar à componente sazonal, com o diferencial de indicar padrões que se repetem em períodos superiores a um ano. Logo, são movimentos ondulatórios que, ao longo de vários anos, tendem à periodicidade.

- Ruído Branco ou Erro Aleatório

É tudo aquilo que não é explicado pelas outras componentes da Série, ou seja, é o que o modelo estimado não consegue captar. O erro compreende a variabilidade intrínseca aos dados e não pode ser modelado. Conforme Murteira; Müller; Turkman (1993, p. 6), a componente aleatória ou o ruído referem-se aos métodos menos sofisticados, os quais colocam tudo o que não conseguem definir ou modelar, tendo papel extremamente relevante nos modernos modelos probabilísticos.

Um erro é dito ruído branco quando possui distribuição normal, sendo a média de seus componentes zero e a variância constante, os quais são não-correlacionados, indicando, assim, que o modelo elaborado conseguiu explicar o máximo de série de dados, conforme Souza (2006).

$$a_t \sim N\left(0, \sigma_a^2\right)$$

Souza & Camargo (1996, p. 23-24) definem ruído branco como uma seqüência $\{a_t, t \in N\}$ de variáveis aleatórias independentes e identicamente distribuídas, não necessariamente, mas usualmente com distribuição Normal de média zero e variância constante σ_a^2. Tais variáveis são chamadas de choques aleatórios, e a seqüência é chamada ruído branco. Assim, o processo a_t, t = 0,1,2,... é um ruído branco se:

$a_t \sim N\left(0, \sigma_a^2\right)$, com:

- Média: $E[a_t] = 0$.

- Autocovariância: $E[a_t, a_{t+k}] = \gamma_k = 0$

  Para $k = \pm 1, \pm 2, ...$ e $\gamma_k = \sigma_a^2$ para $k = 0$.

- Autocorrelação: $\rho_k = 1$ para $k = 0$,

  e $\rho_k = 0$ para $k = \pm 1, \pm 2, ...$

- Autocorrelação parcial: $\phi_{kk} = 1$ para $k = 0$

  e $\phi_{kk} = 0$ para $k = \pm 1, \pm 2, ...$

Para Farias; Rocha; Lima (2000), a análise dos resíduos de modelos concorrentes ajustados é de extrema importância na escolha final do modelo que melhor explica a dinâmica da Série Temporal em estudo. Se os resíduos são autocorrelacionados, então a dinâmica da Série em estudo não é completamente explicada pelos coeficientes do modelo ajustado. Devem ser excluídos do processo de escolha modelo com essa característica. Uma análise da existência da autocorrelação serial de resíduos é feita com base nas Funções de Autocorrelação dos Resíduos (FAC) e Autocorrelação Parcial dos Resíduos (FACP) e seus respectivos correlogramas.

Para Vasconcelos e Alves (2000), os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem se comportar como tal se o modelo estiver adequado. Em particular, seus coeficientes de autocorrelação devem ser estatisticamente iguais a zero. Para essa verificação, é utilizado o teste de Ljung-Box, cuja estatística é dada por:

$$Q(K) = n(n+2)\sum_{j=1}^{K} \frac{\hat{r}_j^2}{(n-j)} \tag{7}$$

e terá aproximadamente uma distribuição χ^2 com K-p-q graus de liberdade (p e q ordens dos parâmetros do modelo). A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de $Q(K)$. Em geral, basta utilizar as 10 ou 15 primeiras $\hat{r}_k$ (autocorrelação dos resíduos de ordem *k*). Se $Q(K) > \chi^2$ e rejeita-se a hipótese que os *j* primeiros coeficientes de autocorrelação são nulos.

Das definições apresentadas, tem-se que a Série Temporal representa um conjunto de observações ordenadas no tempo, que exibe dependência serial e que pode ser modelado por um processo estocástico. Um processo estocástico é o conjunto das possíveis Séries Temporais, que nada mais é que a distribuição das probabilidades conjuntas das variáveis aleatórias. A variável é observada em pontos discretos de tempo, em geral, espaçada igualmente. A análise e a modelagem de Séries Temporais buscam descrever o processo que gera a seqüência da variável determinada, bem como apresentar as suas características mais

importantes. Para estimar uma Série Temporal, é necessário representar o comportamento do processo, por meio de um modelo matemático que possa ser estendido ao futuro. Uma vez que o modelo de Séries Temporais é estabelecido, a técnica apropriada de previsão pode ser desenvolvida. A Série Temporal, para ser estimada, precisa estar livre de tendências e componentes sazonais, para assim conseguir capturar a estrutura de dependência da Série observada, a partir da combinação de ruídos brancos. Logo, se a modelagem da Série Temporal estiver bem ajustada, não deve haver correlação serial entre os resíduos, e o comportamento da Série deve ser explicado usando o menor número possível de parâmetros – princípio da parcimônia.

Nota-se que para ser possível realizar as previsões, necessita-se de uma Série estacionária, sendo muito fácil encontrar Séries que não apresentem esse comportamento. Logo, necessita-se de um modelo adequado que possa captar o máximo de informações da Série em estudo. Um modelo que atende este requisito é o Box & Jenkins.

O enfoque de Box & Jenkins está em que os modelos ARIMA resultam da combinação de três componentes: Auto-regressivo, Integração e Médias Móveis. Esse modelo busca estudar a Série quando esta não se apresenta estacionária.

2.2 Metodologia Box &Jenkins

Durante a década de 60, os professores George E. P. Box e Gwilym M. Jenkins escreveram diversos trabalhos sobre a teoria de controle e análise de Séries Temporais. Em 1970, publicaram o livro *Time Series Analysis, forecasting and control,* apresentando uma metodologia para a análise de Séries Temporais e, em 1976, foi lançada a versão revisada desse livro, que normalmente é a mais mencionada. O grande mérito desse trabalho foi reunir as técnicas existentes em uma metodologia, a fim de construir modelos que descrevessem, com precisão e de forma parcimoniosa, o processo gerador da Série Temporal. Isso proporcionou previsões acuradas de valores futuros. A previsão é uma das principais razões da popularidade desta metodologia e, em muitos casos, as previsões, principalmente de curto prazo, são melhores que as obtidas com base nos modelos econométricos tradicionais.

Esta metodologia é composta pelos Modelos ARIMA, que são modelos estatísticos lineares para análise de Séries Temporais. Esses modelos são mais sofisticados por usarem a correlação entre as observações. No modelo ARIMA, a Série Temporal é gerada por um processo estocástico, cuja natureza pode ser representada através de um modelo. A notação

utilizada para designar o modelo ARIMA é $ARIMA(p,d,q)$, em que *p* é o número de termos auto-regressivos; *d,* o número de diferenciações para que a Série torne-se estacionária e *q,* o número de termos de médias móveis. Os termos *p, d* e *q* são todos inteiros maiores ou iguais a zero. São casos particulares: o modelo $ARMA(p,q)$, o modelo auto-regressivo $AR(p)$ e o modelo de médias móveis $MA(q)$, sendo os três modelos utilizados para Séries Temporais estacionárias $(d=0)$.

Uma Série dispõe de operadores que facilitam a manipulação dos modelos. Observa-se, a seguir:

Operador translação para o passado, denotado por B e definido por:

$$BZ_t = Z_{t-1} \rightarrow B^m Z_t = Z_{t-m} \tag{8}$$

Operador translação para o Futuro, denotado por F e definido por:

$$FZ_t = Z_{t+1} \rightarrow F^m Z_t = Z_{t+m} \tag{9}$$

Operador diferença:

$$\Delta Z_t = Z_t - Z_{t-1} \rightarrow (1-B)Z_t \tag{10}$$

$$\therefore \Delta = 1 - B \tag{11}$$

Operador soma, denotado por S e definido por:

$$SZ_t = \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + \cdots + (1 + B + B^2 + \cdots)Z_t \tag{12}$$

$$SZ_t = (1 - B^{-1})Z_t = \Delta^{-1} Z_t \tag{13}$$

$$\therefore S = \Delta^{-1} \tag{14}$$

Os modelos e equações apresentados neste capítulo foram pesquisados em Morettin e Toloi (2004), Souza & Camargo (1996), Barros (2006) e Vasconcelos e Alves (2000). Todo o capítulo está relacionado aos estudos desses autores.

Os modelos de Séries Temporais possibilitam descrever um processo estocástico, utilizando apenas valores passados da variável dependente e do termo de erro. Assim, dada uma Série, para cada *t* tem-se uma variável aleatória (v. a.) Z_t, que possui uma média $\mu(t)$ e

uma variância $V(t)$. Por isso, tem-se os parâmetros importantes a serem considerados em um processo estocástico: média e função de autocovariância.

Média:

$$\mu(t) = E[Z(t)] \tag{15}$$

Para a Função de Autocovariância:

$$\gamma(t_1, t_2) = E[Z(t_1)Z(t_2)] - E[Z(t_1)]E[Z(t_2)] \tag{16}$$

No caso de $t_1 = t_2 = t$, a Função de Autocovariância é igual à variância:

$$\gamma(t,t) = V(t) = E[Z^2(t)] - E^2[Z(t)] \tag{17}$$

Portanto, a Função de Autocovariância pode ser representada desta forma:

$$\gamma(t_1, t_2) = Cov[Z(t_1)Z(t_2)] \tag{18}$$

É importante o comentário de que a Função de Autocovariância desempenha um papel muito importante na análise dos modelos ARIMA. Esses modelos poderiam simplesmente ser escritos da seguinte forma:

$$Z_t = f(t) + a_t \tag{19}$$

Onde $f(t)$ é chamada de sinal e a_t de ruído.

É partindo dessa função, de acordo com as hipóteses sobre $f(t)$ e a_t, que se chega a vários modelos, sendo um deles os modelos ARIMA.

A hipótese de erros não-correlacionados é o que impõe uma variedade de limitações desse modelo, quando se deseja descrever o comportamento de séries econômicas e sociais, nas quais os erros observados são autocorrelacionados e influenciam a evolução do processo.

A importância dos modelos ARIMA encontra-se neste propósito, descrito através de três classes de processos:

Processos lineares estacionários, descritos no item (2.2.1), são representados da seguinte forma:

$$Z_t - \mu = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \sum_{k=0}^{\infty} \psi_k a_{t-k} \tag{20}$$

Em que a_t é ruído branco $\mu = E(Z_t)$, $\psi_1, \psi_2, \cdots$ é uma seqüência de parâmetros tal que:

$$\sum_{k=0}^{\infty} \psi_k^2 < \infty \tag{21}$$

Existem três casos particulares do modelo (20), que serão descritos a seguir:

–Processo auto-regressivo de ordem p: $AR(p)$;

–Processo de médias móveis de ordem q: $MA(q)$;

–Processo auto-regressivo e de médias móveis de ordem p e q: $ARMA(p,q)$

Processos lineares não-estacionários homogêneos: constituem uma generalidade dos *processos lineares estacionários,* que supõem que o mecanismo gerador da Série sejam não-estacionárias, em nível ou em inclinação. Essas Séries podem tornar-se estacionárias por um meio finito de diferenças (geralmente uma ou duas diferenças são suficientes para estacionarizar a Série).

Processo de memória longa: é um processo estacionário que possui uma longa função de autocorrelação, com decaimento muito longo (hiperbólico) e cuja análise necessitará de uma diferença fracionária $(0 < d < 0{,}5)$.

Esses modelos são bem descritos pelos modelos $ARIMA(p,d,q)$, que aceitam a inclusão de um operador sazonal. Portanto, esta metodologia caracteriza-se por Modelos Estacionários $(AR(p), MA(q), ARMA(p,q))$ e modelos Não-Estacionários $(ARIMA(p,d,q))$

2.2.1 Modelos Estacionários

Os Modelos Estacionários interpretam a Série Temporal como uma realização de um vetor aleatório do tamanho das Séries disponíveis, seguindo o princípio da parcimônia e da construção de modelos, através de um ciclo iterativo. Os modelos Box & Jenkins determinam o processo estocástico adequado para representar uma dada Série Temporal, por meio da passagem de um ruído branco por um filtro linear.

A representação do processo utilizado por Box & Jenkins pode ser observado na Figura 1.

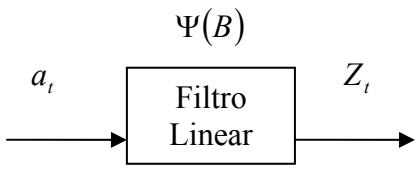


Figura 1 – Filtro Linear com entrada a_t,
saída Z_t e Função Transferência $\Psi(B)$.
Fonte: Morettin e Toloi (2004, p. 111)

Modelos derivados do filtro linear são capazes de representar estacionariedade e não-estacionariedade de uma Série Temporal. Se uma Série Temporal é estacionária, sua média, variância e autocorrelação (a defasagens diversas) permanecem as mesmas ao longo do tempo, ou seja, constantes.

Formalmente,

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \mu + \psi(B) a_t \tag{22}$$

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots \tag{23}$$

A equação (23) é denominada Função de Transferência do filtro, e μ é o parâmetro determinando o nível da série.

A equação (22) é um processo linear. Dessa forma:

$E(a_t) = 0, \forall t;$

$Var(a_t) = \sigma_a^2, \forall t;$

$E(a_t a_s) = 0, s \neq t$

Chamando $\widetilde{Z}_t = Z_t - \mu$ Série centrada na média μ, tem-se:

$$\widetilde{Z}_t = \psi(B) a_t \tag{24}$$

Se a seqüência de pesos $\{\psi_j, j \geq 1\}$ for convergente, o filtro é estável e Z_t é estacionária. Nesse caso, μ é a média do processo. Caso contrário, Z_t é não-estacionária e μ não tem significado específico, a não ser como ponto de referência da Série.

Da equação (22), tem-se que:

$$E(Z_t) = \mu + E\left(a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} \right) \tag{25}$$

E como $E(a_t) = 0, \forall t$, temos que $E(Z_t) = \mu$ se a série $\sum_{j=1}^{\infty} \psi_j$ convergir. Quanto à Função de Autocovariância γ_j de Z_t, ela é dada por:

$$\gamma_j = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+i} \tag{26}$$

Com $\psi_o = 1$. Um caso particular é para $i = 0$, onde a γ_j é igual à variância de Z_t:

$$\gamma_0 = Var(Z_t) = \sigma_a^2 \sum_{j=o}^{\infty} \psi_j^2 \tag{27}$$

A condição de existência das equações (26) e (27) é que $\sum_{j=o}^{\infty} \psi_j^2 < \infty$.

Por conseguinte, percebe-se que a média e a variância de Z_t apresentam-se constantes, e γ_j só depende dos valores de j, sendo, portanto, estacionária.

Pode-se escrever $\widetilde{Z}_t$ em uma forma alternativa, como uma soma ponderada dos valores passados e mais um ruído:

$$\widetilde{Z}_t = \pi_1 \widetilde{Z}_{t-1} + \pi_2 \widetilde{Z}_{t-2} + \cdots + a_t = \sum_{j=1}^{\infty} \pi_j \widetilde{Z}_{t-j} + a_t \tag{28}$$

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \widetilde{Z}_j = a_t \tag{29}$$

Ou, simplesmente, é possível representar um modelo estacionário da seguinte forma:

$$\pi(B) \widetilde{Z}_j = a_t \tag{30}$$

Em que $\pi(B)$ é o operador do modelo:

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots \tag{31}$$

Unindo as equações (30) e (24), resulta-se na seguinte equação:

$$\pi(B) \psi(B) a_t = a_t \tag{32}$$

Ou, também, representá-la da seguinte maneira:

$$\pi(B) = \psi^{-1}(B) \tag{33}$$

Essa equação é utilizada para obter os valores de π_j, em função ψ_j e vice-versa.

Com base nas particularidades dos processos estacionários estabelecidos acima, serão descritos, abaixo, os modelos que compõem esse processo: $(AR(p), MA(q), ARMA(p,q))$.

Modelos Auto-Regressivos de Ordem p - $AR(p)$

Esse modelo corresponde à defasagem da série Z_t. A versão mais simples de um modelo AR é o de ordem um, sendo aquele em que Z_t depende somente de Z_{t-1} e de a_t, no instante *t*. O modelo que contém $\phi(B)=1$ é chamado modelo auto-regressivo, sendo denotado por $AR(p)$, em que *p* indica a ordem do modelo, isto é, o número de defasagens.

Na equação (28), os parâmetros são representados por π_j. Se, nessa mesma equação, o parâmetro assumir o valor de $\pi_j = 0, j > p$, em que *p* é a ordem do modelo, a equação se reduzirá a um modelo auto-regressivo de ordem *p*, denotado por $AR(p)$. Então, é possível reescrever a equação auto-regressiva de acordo com a equação usual, e renomear os pesos π_j para ϕ_j. Os parâmetros desse modelo são representados por ϕ:

$$\widetilde{Z}_t = \phi_1 \widetilde{Z}_{t-1} + \phi_2 \widetilde{Z}_{t-2} + ... + \phi_p \widetilde{Z}_{t-p} + a_t \tag{34}$$

O operador auto-regressivo estacionário de ordem *p*:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p . \tag{35}$$

Podendo, então, ser reescrito da seguinte forma:

$$\phi(B)\widetilde{Z}_t = a_t . \tag{36}$$

Nas equações (37) e (38), encontram-se os modelos Auto-Regressivos de ordem 1 e 2 respectivamente:

$AR(1)$

$$\widetilde{Z}_t = \phi_1 \widetilde{Z}_{t-1} + a_t \tag{37}$$

$AR(2)$

$$\widetilde{Z}_t = \phi_1 \widetilde{Z}_{t-1} + \phi_2 \widetilde{Z}_{t-2} + a_t \tag{38}$$

A identificação do modelo adequado começa pelos valores dos parâmetros, ou seja, o modelo será considerado estacionário se $|\phi_1| < 1$, satisfazendo a condição da estacionaridade. A condição de estacionaridade do $AR(p)$ estabelece que todas as raízes da equação devem cair fora do círculo unitário.

Em contrapartida, o modelo que apresenta ordem finita sempre será inversível. Como $\pi(B) = \phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ é finito, não existem restrições sobre os parâmetros para assegurarem a invertibilidade.

Além de $|\phi_1| < 1$, é necessário recorrer às Funções de Autocorrelação. Voltando para equação (34), multiplica-se todos os membros desta por Z_{t-j}, e tomando-se a esperança, tem-se:

$$E\left(\tilde{Z}_t \tilde{Z}_{t-j}\right) = \phi_1 E\left(\tilde{Z}_{t-1} \tilde{Z}_{t-j}\right) + \phi_2 E\left(\tilde{Z}_{t-2} \tilde{Z}_{t-j}\right) + \ldots + \phi_p E\left(\tilde{Z}_{t-p} \tilde{Z}_{t-j}\right) + a_t E\left(a_t \tilde{Z}_{t-j}\right) \tag{39}$$

Como $\tilde{Z}_{t-j}$ só envolve ruídos até a_{t-j}, não-correlacionados com a_t e $E\left(a_t \tilde{Z}_{t-j}\right) = 0, j > 0$, resulta que:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \ldots + \phi_p \gamma_{j-p}, j > 0 \tag{40}$$

Para $j = 0$ em (40):

$$\frac{\gamma_0}{\gamma_0} = \frac{\phi_1 \gamma_{-1} + \phi_2 \gamma_{t-2} + \ldots + \phi_k \gamma_{-k} + \sigma_a^2}{\gamma_0} \tag{41}$$

como $\gamma_{-j} = \gamma_j$, tem-se que $\gamma_0 = Var\left(\tilde{Z}_t\right) = Var\left(Z_t\right)$:

$$1 = \phi_1 \rho_1 + \phi_2 \rho_2 + \cdots + \phi_j \rho_j + \frac{\sigma_a^2}{\gamma_0} \tag{42}$$

Ou seja:

$$Var\left(Z_t\right) = \sigma_z^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \cdots - \phi_j \rho_j} \tag{43}$$

Logo, dividindo (35) por $\gamma_0 = Var\left(\tilde{Z}_t\right) = Var\left(Z_t\right)$, obtem-se:

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \ldots + \phi_p \rho_{j-p}, j > 0 \tag{44}$$

Que também pode ser escrita como:

$$\phi(B)\rho_j = 0 \tag{45}$$

Em que o operador defasagem passa a agir agora em *j*:

$$B\rho_j = B\rho_{j-1} \cdots \tag{46}$$

As Funções de Autocorrelação e Função de Autocorrelação Parcial são muito importantes, pois a análise visual permite identificar a ordem do modelo:

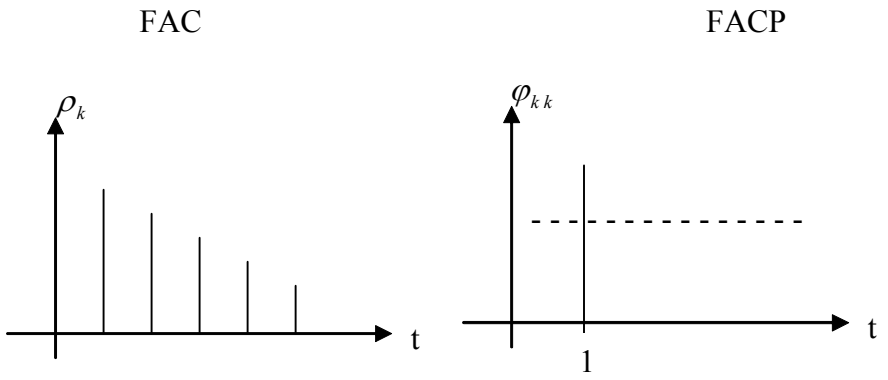


Figura 2 - Comportamento das funções de autocorrelação e autocorrelação parcial de um modelo AR (1).

A Função de Autocorrelação de um processo auto-regressivo é constituída de uma mistura de polinômios, exponenciais e senóides amortecidas, e a Função de Autocorrelação Parcial (FACP) apresenta um corte rápido no *1ag* significativo, indicando a ordem p do modelo.

Para Morettin e Toloi (2004), a função de autocorrelação decai exponencialmente quando ϕ_1 é positivo; quando ϕ_1 é negativo, a função de autocorrelação também decai exponencialmente, mas apresenta alternância de sinais positivos e negativos.

Modelo de Médias Móveis de ordem q - $MA(q)$

O modelo de Médias Móveis é a combinação do ruído branco a_t, do período atual com os ocorridos anteriormente. A notação do modelo é $MA(q)$, na qual q indica a ordem do modelo. Voltando à equação (22), suponha-se que $\psi_j = 0, j > q$, passa-se, então, a utilizar a notação:

$$Z\,t = \mu + a_t - \theta_1\, a_{t-1} - ... - \theta_q\, a_{t-q}\,; \tag{47}$$

Chamando $\widetilde{Z}_t = Z_t - \mu$, resulta que:

$$\widetilde{Z}_t = \left(1 - \theta_1 B - ... - \theta_q\, B^q\right) a_t = \theta(B) a_t \tag{48}$$

O operador de médias móveis estacionário de ordem q:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\,. \tag{49}$$

É possível observar como são descritos os modelos de ordem:

$MA(1)$

$$\widetilde{Z}_t = a_t - \theta_1 a_{t-1} \tag{50}$$

$MA(2)$

$$\widetilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \tag{51}$$

O próximo passo será analisar as condições de estacionaridade e invertibilidade, que precisam ser satisfeitas. Dado que $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$, não existem restrições sobre os parâmetros θ_j, para que o processo seja estacionário. Para isso, usaram-se argumentos completamente similares aos modelos $AR(p)$, a fim de verificar a estacionariedade, as condições de invertibilidade para um modelo $MA(q)$, observando que as raízes da equação característica $\theta(B) = 0$ estejam fora do círculo unitário. Esse modelo é igual a um modelo $AR(p)$, em que p é infinito.

A próxima análise a ser feita no modelo é a da Função de Autocorrelação. Primeiramente, a Função de Autocovariância desse modelo é:

$$\begin{aligned} \gamma_j &= E\left[\widetilde{Z}_t \widetilde{Z}_{t-j}\right] E\left[\left(a_t - \sum_{k=1}^{q} \theta_k a_{t-k}\right)\left(a_{t-j} - \sum_{l=1}^{q} \theta_k a_{t-j-l}\right)\right] \\ &= E\left(a_t a_{t-j}\right) - \sum_{k=1}^{q} \theta_k E\left(a_{t-j} a_{t-k}\right) - \sum_{l=1}^{q} \theta_1 E\left(a_t a_{t-j-l}\right) + \sum_{k=1}^{q} \sum_{l=1}^{q} \theta_k \theta_1 E\left(a_{t-k} a_{t-j-l}\right) \end{aligned} \tag{52}$$

Não esquecendo que:

$$\begin{aligned} \gamma_a(j) &= E\left(a_t a_{t-j}\right) = \begin{cases} \sigma_a^2, j = 0 \\ 0, j \neq 0 \end{cases} \\ \gamma_0 &= Var\left(Z_t\right) = \sigma_Z^2 = \left(1 + \theta_1^2 + \cdots + \theta_q^2\right) = \sigma_a^2 \end{aligned} \tag{53}$$

Em termos de $\gamma_a(j)$, a função de autocovariância de Z_t fica:

$$\begin{aligned} \gamma_j &= \left(-\theta_j \sum_{l=1}^{q-j} \theta_l \theta_{j+l}\right) \sigma_a^2 \\ \gamma_j &= \left(-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \cdots + \theta_q \theta_{j-q}\right) \sigma_a^2, j = 1, \cdots, q \\ \gamma_j &= 0, j > q \end{aligned} \tag{54}$$

Unindo as equações (53) e (54), obtem-se a Função de Autocorrelação do processo:

$$\rho_j = \begin{cases} \dfrac{-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \cdots + \theta_q \theta_{j-q}}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2} \; j = 1, \cdots, q \\ 0, j > q \end{cases} \tag{55}$$

Permite-se verificar que a função de autocorrelação de um processo $MA(q)$ é igual a zero para *lag's* maiores que *q*, ao contrário do que acontece em um processo $AR(p)$.

Dessa forma, identificando a condição de estacionariedade, o processo apresenta ordem finita, portanto, será sempre estacionário.

Quanto à condição de invertibilidade, é imposta a restrição de que o parâmetro em módulo é menor que um, ou seja, $|\theta_1| < 1$.

A identificação do modelo é feita através das autocorrelações parciais, sendo que a função de autocorrelação fornece a ordem do modelo:

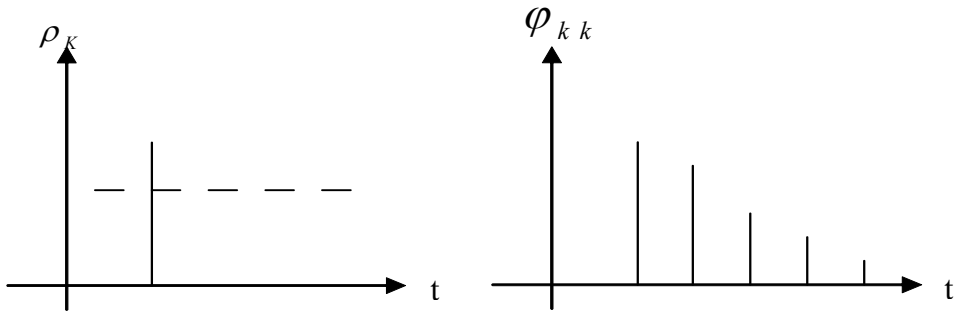


Figura 3 - Comportamento das funções de autocorrelação parcial e autocorrelação de um modelo MA (1)

Para Souza (1991), a autocorrelação apresentará um corte rápido no *1ag* significativo, indicando a ordem do modelo, enquanto que a autocorrelação parcial decairá exponencialmente, se todos os parâmetros forem positivos, conforme Figura 2. Caso contrário, formará uma senóide amortecida.

Modelo Auto-Regressivo de Médias Móveis - $ARMA(p,q)$

Esse modelo é a união dos modelos vistos anteriormente, os auto-regressivos e o de médias móveis. Ele é utilizado em Séries encontradas na prática, para as quais não se deseja números muito grandes de parâmetros.

Um modelo $ARMA(p,q)$ apresenta-se da seguinte forma:

$$\widetilde{Z}_t = \underbrace{\phi_1 \widetilde{Z}_{t-1} + \cdots + \phi_p \widetilde{Z}_{t-p} + a_t}_{AR(p)} \underbrace{- \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}}_{MA(q)} \qquad 56)$$

Se $\phi(B)$ e $\theta(B)$ são os operadores auto-regressivos e de médias móveis, respectivamente, pode-se reescrever um modelo $ARMA(p,q)$ através dos seus operadores:

$$\phi(B)\widetilde{Z}_t = \theta(B)a_t \tag{57}$$

Um modelo freqüentemente usado é o $ARMA(1,1)$, no qual a ordem de ambos os parâmetros é igual a um $(p = q = 1)$. Tem-se, também, que $\phi(B) = 1 - \phi B$ e $\theta(B) = 1 - \theta B$, ou seja, reduz-se a equação (56) em:

$$\widetilde{Z}_t = \phi\widetilde{Z}_{t-1} + a_t - \theta a_{t-1} \tag{58}$$

Facilmente, pode-se escrever (58) na forma de um processo linear ou de médias móveis infinito, apenas substituindo $\widetilde{Z}_{t-1}, \widetilde{Z}_{t-2}, \cdots$:

$$\widetilde{Z}_t = \psi(B)a_t \tag{59}$$

Em que $\psi_j = \phi^{j-1}(\phi - \theta), j \geq 1$, sendo o processo:

Estacionário: $\sum \psi_j = (\phi - \theta)\sum \phi^{j-1} < \infty$, ou seja, $|\phi| < 1$.

Também podendo ser escrito da seguinte forma:

$$\pi(B)\widetilde{Z}_t = a_t \tag{60}$$

Em que os pesos $\pi_j = \theta^{j-1}(\phi - \theta), j \geq 1$, de modo que o processo é irreversível se $\sum \pi_j < \infty$, ou seja, $|\phi| < 1$.

Em seguida, percebe-se que as condições de estacionariedade de um processo $ARMA(1,1)$ são as mesmas que as de um processo $AR(1)$, e as condições de invertibilidade são as mesmas que as de um processo $MA(1)$.

Essas conclusões se generalizam para qualquer ordem de um modelo $ARMA(p,q)$.

Dessa forma, pode-se concluir que o processo é inversível, se todas as raízes de $\phi(B) = 0$ e $\theta(B) = 0$ caírem fora do círculo unitário.

Quanto à Função de Autocorrelação, sua ordem vem da multiplicação da equação (55) por $\widetilde{Z}_{t-j}$ e tomando a esperança. Logo:

$$\gamma_j = E\left[\left(\phi_1\widetilde{Z}_{t-1} + \cdots + \phi_p\widetilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}\right)\widetilde{Z}_{t-j}\right] \tag{61}$$

$$\gamma_j = \phi_1\gamma_{t-1} + \cdots + \phi_p\gamma_{t-p} + \gamma_{za}(j) - \theta_1\gamma_{za}(j-1) - \cdots - \theta_q\gamma_{za}(j-q) \quad (62)$$

Em que $\gamma_{za}(j)$ é a covariância cruzada entre Z_t e a_t, definida por:

$$\gamma_{za}(j) = E\left(a_t\widetilde{Z}_{t-j}\right) \quad (63)$$

Como $\widetilde{Z}_{t-j}$ só depende dos choques até a_t, ocorridos até o instante $t-j$, obtem-se:

$$\gamma_{za}(j) = \begin{cases} = 0, j > 0 \\ \neq 0, j \leq 0 \end{cases} \quad (64)$$

De modo que (62) é reescrita:

$$\gamma_j = \phi_1\gamma_{j-1} + \cdots + \phi_p\gamma_{j-p}, j > q \quad (65)$$

E a Função de Autocorrelação é obtida através de (65):

$$\rho_j = \phi_1\rho_{j-1} + \cdots + \phi_p\rho_{j-p}, j > q \quad (66)$$

Observa-se que os *lag's,* $1,2,\cdots,q$ das autocorrelações são afetados diretamente pelos parâmetros das médias móveis, mas, para $j > q$, as funções de autocorrelação assumem o comportamento de um modelo $AR(p)$. Ao verificar-se que:

- $q < p$, a função de autocorrelação é uma mistura de exponenciais ou senóides amortecidas;

- $q \geq p$, os primeiros $q-p-1$ valores da função de autocorrelação não seguem este padrão.

Até agora, os processos estacionários, que são os modelos $AR(p)$, $MA(q)$, $ARMA(p,q)$, apresentam Função de Autocorrelações (FAC) especiais:

$AR(p)$**:** FAC decai de acordo com exponenciais e ou senóides amortecidas, infinita ou em extensão;

$MA(q)$: FAC finita e com um corte após o *lag q*;

$ARMA(p,q)$: tem o mesmo comportamento da FAC de um modelo $AR(p)$, após o *lag q-p*.

A FAC possui importante papel na identificação do modelo em estudo. Box, *et.al*, 1994 propõem a utilização da Função de Autocorrelação Parcial (FACP), para facilitar o

procedimento de identificação do modelo, o qual se denotará ϕ_{kj} de um modelo $AR(k)$, de tal modo que ϕ_{kk} seja o último coeficiente:

$$\rho_j = \phi_{k1}\rho_{j-1} + \cdots + \phi_{kk}\rho_{j-k}, j = 1, \cdots, k \quad (67)$$

Para os processos estudados, tem-se que:

$AR(p)$: FACP $\phi_{kk} \begin{cases} \neq 0, k \leq p \\ = 0, k > p \end{cases}$

$MA(q)$: a FACP desse modelo se comporta da mesma forma que a FAC do modelo $AR(p)$, denominado por exponenciais e senóides amortecidas.

$ARMA(p,q)$: a FACP se comporta da mesma forma que a FACP de um processo $MA(q)$.

Considerando os seguintes parâmetros, abaixo segue a Tabela 1 com o resumo dos padrões teóricos da FAC e FACP dos modelos Estacionários.

TABELA 1 - Padrões Teóricos das FAC e FACP

PROCESSO	FAC	FACP
Ruído Branco	$\rho_k = 0, \forall k$	$\rho_{kk} = 0, \forall$ k.
AR (1): $\phi_1 > 0$	Queda Exponencial Direta $\rho_k = \phi_1^k$	$\rho_{11} = \rho_1$;
AR (1): $\phi_1 < 0$	Queda com Oscilação $\rho_k = \phi_1^k$	$\rho_{11} = \rho_1$; $\rho_{kk} = 0, \forall$ k ≥ 2.
AR(p)	Queda em Direção a zero. Coeficientes podem oscilar	Apresenta Pico na Defasagem p; $\rho_{kk} = 0, \forall$ k > p.
MA(1): $0 < \theta_1$	Pico Positivo na defasagem 1 $\rho_k = 0, \forall k \geq 2$.	Queda com Oscilação $\rho_{11} > 0$
MA(1): $0 > \theta_1$	Pico Negativo na defasagem 1 $\rho_k = 0, \forall k \geq 2$.	Queda sem Oscilação $\rho_{11} < 0$.
ARMA(1,1): $\phi_1 < 0$	Queda Exponencial começando na defasagem 1 Sinal ρ_1 = Sinal ($\phi_1 + \theta_1$)	Queda Exponencial começando na defasagem 1 $\rho_{11} = \rho_1$; Sinal ρ_{kk}=Sinal ρ_{11}.
ARMA(p,q):	Queda com Oscilação começando na defasagem q	Queda com Oscilação começando na defasagem q

ρ_k = Autocorrelação;

ρ_{kk} = Autocorrelação Parcial;

k = Defasagem das FAC e FACP;

ϕ = Coeficiente do processo AR(p);

θ = Coeficiente do processo MA(q);

p = Defasagem do processo AR;

q = Defasagem do processo MA.

2.2.2 Modelos Não-Estacionários

Séries não-estacionárias são aquelas que não se desenvolvem ao longo do tempo ao redor de uma média constante. Mas, essas Séries podem tornar-se estacionárias, ao recorrer-se ao processo de diferenciação.

Seja Z_t uma Série não-estacionária:

$$W_t = Z_t - Z_{t-1} = (1 - B)Z_t = \Delta Z_t \tag{68}$$

Com um número finito de diferenças, *d,* transformou-se Z_t em estacionária. Essas Séries também são chamadas de estacionárias homogêneas ou, ainda, de portadoras de raízes unitárias.

Modelo Auto-Regressivo Integrado de Médias Móveis- $ARIMA(p,d,q)$

Dada Z_t, não-estacionária, se $W_t = \Delta^d Z_t$ for estacionária, pode-se representar W_t por um modelo ARMA (p,q), isto é:

$$\phi(B)W_t = \theta(B)a_t . \tag{69}$$

Sendo W_t uma diferença de Z_t, pode-se dizer que Z_t segue um modelo auto-regressivo, integrado de médias móveis, de ordem (p,d,q), ou seja, $ARIMA(p,d,q)$, em que *p* e *q*, são as respectivas ordens de $\phi(B)$ e $\theta(B)$.

$$\phi(B)\Delta^d Z_t = \theta(B)a_t . \tag{70}$$

No modelo (69), estão fora do círculo unitário. Portanto, escrever (70) equivale a escrever:

$$\omega(B)\, Z_t = \theta(B)a_t . \tag{71}$$

Em que $\omega(B)$ é um operador auto-regressivo não-estacionário de ordem $p+q$,com d raízes iguais a um (sobre o círculo unitário), e as restantes p fora desse círculo:

$$\omega(B)=\phi(B)\Delta^d = \phi(B)(1-B)^d . \tag{72}$$

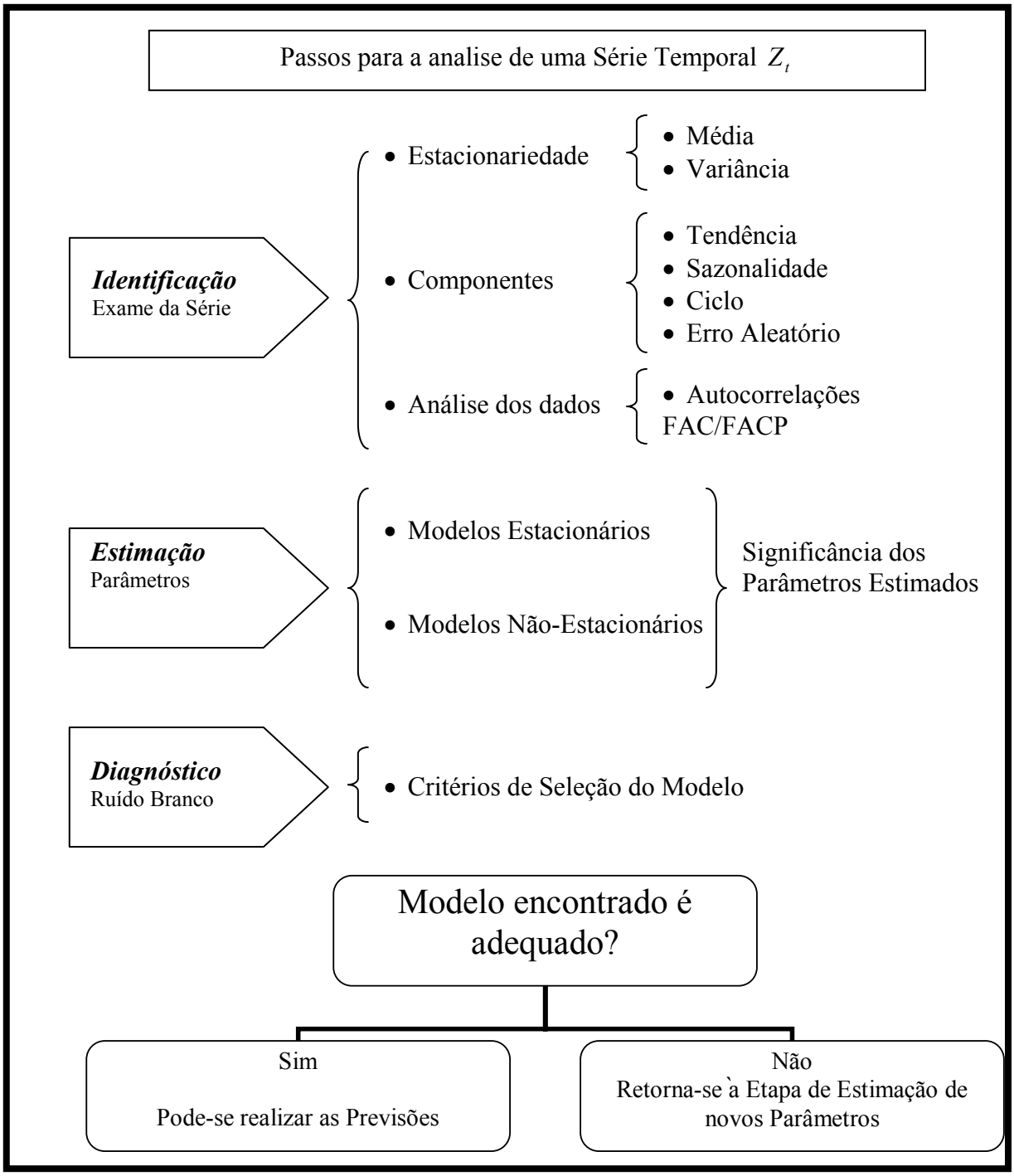
Observa-se que é indiferente escrever $\omega(B)\ Z_t$ ou $\omega(B)\ \widetilde{Z}_t$, pois $\Delta^d\ Z_t = \Delta^d\ \widetilde{Z}_t$, para $d>1$.

É importante destacar que os modelos ARIMA não apresentam especificidade gráfica, somente plotando a Série Temporal, não sendo possível identificar qual o modelo que a representa. Para que se possa determinar qual o modelo dessa família que melhor representa a Série, e que poderá ser utilizado para fazer as previsões, a sua construção será baseada no ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos próprios dados.

Caso o modelo não seja adequado, o ciclo é repetido, voltando à fase de identificação. Um procedimento que muitas vezes é utilizado é identificar não um só modelo, os quais serão então estimados e verificados. Se o propósito é previsão, escolher-se-á, entre os modelos ajustados, o melhor; por exemplo, no sentido de fornecer o menor erro quadrático médio de previsão.

Para os autores, a fase crítica do procedimento é a identificação, pois é possível que vários pesquisadores identifiquem diferentes modelos para uma única Série em estudo. Isso é possível, inclusive, pelo critério de decisão adotado para escolher o modelo adequado. Em geral, os modelos postulados são parcimoniosos, por terem um pequeno número de parâmetros e pelas previsões obtidas serem bastante precisas.

A estratégia utilizada para a construção de modelos, pela metodologia Box & Jenkins, é baseada no ciclo iterativo, descrito no Quadro 1.



QUADRO 1 – Quadro representativo dos Estágios do Ciclo Iterativo de Box & Jenkins.

Identificação do modelo

Para Barros (2006), no passo de identificação de um modelo ARIMA, busca-se encontrar procedimentos que permitem identificar os parâmetros do modelo para a Série Temporal em estudo. Um ponto importante na identificação é a parcimônia, pois de dois modelos com ajustes igualmente bons, escolhe-se aquele com menor número de parâmetros.

Isso deve-se à simplicidade da estrutura, permitindo uma melhor compreensão do processo, enquanto que os modelos superparametrizados se ajustam bem ao que acontece dentro da amostra, mas não generalizam o comportamento da Série. O primeiro passo para a identificação é realizar a ordem da diferenciação da Série, que tende a introduzir correlações negativas. Por isso, é importante ter o cuidado de não diferenciar a Série original mais do que o necessário para estacionarizá-la.

Nesse sentido, considera-se essa a fase mais crítica do processo. A escolha do melhor modelo ajustado é feita principalmente com base nas Funções de Autocorrelação (FAC) e Funções de Autocorrelação Parcial (FACP) estimadas.

A FAC ρ_j é estimada por:

$$r_j = \frac{c_j}{c_0}, j = 0,1,\cdots,N-1. \qquad 73)$$

Em que c_j é estimada pela Função de Autocovariância (FACV) γ_j:

$$c_j = \frac{1}{N}\sum_{t=1}^{N-j}\left[\left(Z_t - \overline{Z}\right)\left(Z_{t+j} - \overline{Z}\right)\right], j = 0,1,\cdots,N-1. \qquad 74)$$

Em que $\overline{Z} = \frac{1}{N}\sum_{t=1}^{N} Z_t$ é a média amostral. É importante lembrar que $r_{-j} = r_j$:

A variância aproximada de r_j:

$$Var\left(r_j\right) \cong \frac{1}{N}\sum_{\upsilon=-\infty}^{\infty}\left[\rho_\upsilon^2 + \rho_{\upsilon+j}\rho_{\upsilon-j} - 4\rho_j\rho_\upsilon\rho_{\upsilon-j} + 2\rho_\upsilon^2\rho_j^2\right]. \qquad 75)$$

Em um processo no qual as autocorrelações são nulas para $\upsilon > q$, todos os termos do lado direito de (75) anulam-se para $j > q$, exceto o primeiro, obtendo-se:

$$Var\left(r_j\right) \cong \frac{1}{N}\left[1 + 2\sum_{\upsilon=1}^{q}\rho_\upsilon^2\right], j > q. \qquad 76)$$

Como as autocorrelações são desconhecidas, substitui-se ρ_υ por r_υ, obtendo-se as estimativas para (76):

$$\sigma^2\left(r_j\right) \cong \frac{1}{N}\left[1 + 2\sum_{\upsilon=1}^{q} r_\upsilon^2\right], j > q. \qquad 77)$$

Para N suficientemente grande e sob a hipótese de $\rho_j = 0$, para $j > q$, será criado um intervalo de confiança aproximado para as autocorrelações, pois a distribuição da

estimativa aproxima-se de uma distribuição normal, com média igual a zero e variância encontrada através da equação (77). Logo:

$$r_j \pm t_\gamma . \hat{\sigma}(r_j). \tag{78}$$

Em que t_γ é o valor de t de *Student* com *N-1* graus de liberdade, tal que $P(-t_\gamma < t < t_\gamma) = \gamma$. Na prática, é utilizado $t_\gamma = 2$, sendo assim, possível considerar ρ_j como sendo significativamente diferente de zero, se:

$$|r_j| \pm 2.\hat{\sigma}(r_j), j > q. \tag{79}$$

A FACP é analisada sob a hipótese de um modelo $AR(p)$, em que:

$$Var(\hat{\phi}_{jj}) \cong \frac{1}{N}, j > p \tag{80}$$

Como as FAC são analisadas através de estimativas:

$$\hat{\sigma}(\hat{\phi}_{jj}) \cong \frac{1}{\sqrt{N}}, j > p \tag{81}$$

No caso de N apresentar-se muito grande e sob hipótese de o modelo ser $AR(p)$, $\hat{\phi}_{jj}$ terá uma distribuição aproximadamente normal, com média zero e variância encontrada através da equação (80), sendo possível considerar $\hat{\phi}_{jj}$ significativamente diferente de zero, se:

$$|\hat{\phi}_{jj}| > \frac{1}{\sqrt{N}}, j > p \tag{82}$$

O objetivo do processo de identificação é encontrar a ordem dos parâmetros do modelo $ARIMA(p,d,q)$. Essa fase do modelo consiste em três partes:

- Verificar a necessidade de transformação da Série original, com o objetivo de estabilizar a variância, que pode ser identificada através de uma análise visual do gráfico da Série;

- Tomar diferenças da Série até torná-la estacionária. O número de diferenças necessárias é alcançado quando a FAC da Série decresce rapidamente para zero;

- Encontrar o processo através das FAC e FACP estimadas, no qual essas encontram seus valores dentro dos limites de confiabilidade estipulado.

Nesse processo de identificação, busca-se determinar a ordem dos modelos, tomando-se por base o comportamento das Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP), além dos seus respectivos correlogramas (exposição gráfica das FAC e FACP contra

as defasagens temporais). Para identificar os modelos, pode-se utilizar a Tabela 1 (pág. 35) como suporte, observando o comportamento das Funções de Autocorrelação.

Estimação dos parâmetros

Nesta fase da metodologia, cada um dos modelos sugeridos na fase de identificação é ajustado, e os vários parâmetros são examinados. Os modelos ajustados são comparados usando os critérios de parcimônia e os de estacionaridade e invertibilidade.

A estimação pode ser por mínimos quadrados e por máxima verossimilhança. Para a estimação da máxima verossimilhança, é necessário admitir, inicialmente, que os ruídos brancos têm a distribuição Normal.

Experimentos com simulação pelo método de Monte Carlo, realizados por vários autores, indicam que o método de máxima verossimilhança é superior ao método de mínimos quadrados, quando o tamanho da Série é pequeno e, sobretudo, quando os valores dos parâmetros aproximam-se dos limites da condição de invertibilidade. Para mais detalhes, ver Vasconcelos e Alves (2000, p.69-82)

Diagnóstico ou adequação do modelo

Como já comentado anteriormente, existe a possibilidade de serem encontrados vários modelos, candidatos à representação em estudo. Dentre esses, somente um deve ser escolhido. Portanto, recorre-se aos seguintes critérios de decisão:

- Parcimoniosidade nos parâmetros. Deve-se escolher, preferencialmente, entre os modelos de melhor ajuste, aqueles com um número menor de parâmetros;
- Invertibilidade e estacionariedade dos coeficientes dos modelos estimados;
- Ajustamento do modelo aos dados da Série Temporal em estudo, com base nos critérios *Akaike Information Criterion,* AIC, e *Schwartz Bayesian Criterion*, SBC.

Esses critérios permitem a comparação de diferentes modelos para a Série e, por conseguinte, a escolha do modelo com melhor ajustamento. É nesse critério que se minimiza a variância dos resíduos e que se estabelece a conseqüente penalidade pela inclusão de mais parâmetros. Isso porque se busca modelos parcimoniosos e, geralmente, a inclusão de mais parâmetros leva à melhora do ajuste do modelo, mas peca pelo excesso de parametrização.

Criado por Akaike, em 1974, o critério *AIC (Akaike Information Criterion)* estabelece que o modelo ótimo para uma Série é aquele que minimiza o *AIC*, e é representado por:

$$AIC = T\ln(SQR) + 2n \tag{83}$$

Segundo Barros (2006), esse critério tende a superestimar os modelos auto-regressivos. Com o intuito de corrigir esse problema é que se desenvolveu o próximo critério, *Schwartz Bayesian Criterion*, SBC.

$$SBC = T\ln(SQR) + n\ln(T) \tag{84}$$

Em que:

SQR = soma do quadrado dos resíduos;

n = número de parâmetros estimados;

T = número de observações utilizadas.

Criado por Schwartz, em 1978, esse critério vem a ser um refinamento do critério AIC, desenvolvido justamente para resolver o problema encontrado nele.

Os critérios AIC e SBC podem ser negativos. Isso ocorre quando a soma do quadrado dos resíduos é menor que 1 e *T* ln (SQR) > *2n* ou *T* ln (SQR) > n ln(T).

Um fator a ser observado é quanto aos resíduos do modelo estimado, se são estimativas do ruído branco a_t. Assim sendo, devem comportar-se como um ruído branco, se o modelo estiver adequadamente especificado. Em particular, seus coeficientes de autocorrelação devem ser estatisticamente iguais a zero. Os resíduos se distribuem como uma Normal de média zero e variância 1/n.

Previsão

Essa é a ultima etapa do processo dos modelos ARIMA, seguindo o Ciclo Iterativo mostrado no Quadro 1. Os passos construídos até aqui se devem ao interesse de utilizar o modelo identificado e estimado para se realizar as previsões. Tais Previsões são representadas por $Z_{t+h}, h > 1$, sendo que se tem uma Série original Z_t é até o instante *t;*, e a previsão de origem *t* e *horizonte h* será denotada por $\hat{Z}_t(h)$. Também, será assumido que $W_t = (1-B)^d Z_t$ é estacionário e invertível, e os parâmetros são conhecidos.

Já que o objetivo agora é a previsão, é possível reescrever os modelos $\mathrm{ARIMA(p,d,q)}$ nas três formas básicas:

Forma de Equação de Diferença:

$$Z_{t+h} = \omega_1 Z_{t+h-1} + \cdots + \omega_{p+d} Z_{t+h-p-d} - \theta_1 a_{t+h-1} - \cdots - \theta_q a_{t+h-q} + a_{t+h} \tag{85}$$

Forma de Choques Aleatórios:

$$Z_{t+h} = \sum_{j=-\infty}^{t+h} \psi_{t+h-j} a_j = \sum_{j=0}^{t+h} \psi_j a_{t+h-j} \tag{86}$$

Em que $\psi_0 = 1$.

Forma Invertida:

$$Z_{t+h} = \sum_{j=1}^{\infty} \pi_j Z_{t+h-j} + a_{t+h} \tag{87}$$

- Previsão do Erro Quadrático Médio Mínimo - EQM

Seja $\hat{Z}_t(h)$ uma função linear, utilizada para representar as observações até o instante *t* e a equação (86) uma função de $a_t, a_{t-1}, \cdots$:

$$\hat{Z}_t(h) = \psi_h^* a_t + \psi_{h+1}^* a_{t-1} + \psi_{h+2}^* a_{t-2} + \cdots \tag{88}$$

Deseja-se determinar os valores ψ_j^* que minimizem o EQM da previsão dada por:

$$E\left[Z_{t+h} - \hat{Z}_t(h)\right] = E\left[\sum_{j+0}^{\infty} \psi_j^* a_{t+h-j} - \sum_{j+0}^{\infty} \psi_{h+j}^* a_{t-j}\right]^2 \tag{89}$$

Com base nas equações (86) e (88), chega-se ao erro de previsão dado por:

$$e_t(h) = Z_{t+h} - \hat{Z}_t(h) = \psi_0 a_{t+h} + \psi_1 a_{t+h-1} + \cdots + \psi_{h-1} a_{t+1} - \sum_{j=0}^{\infty} \left(\psi_{h+j} - \psi_{h+j}^*\right) a_t \tag{90}$$

Chegando à equação (91), devido o fato de a_t serem não-correlacionados:

$$E\left[e_t(h)\right]^2 = \left(1 + \psi_1^2 + \cdots + \psi_{h-1}^2\right)\sigma_a^2 + \sum_{j=0}^{\infty} \left(\psi_{h+j} - \psi_{h+j}^*\right)^2 \sigma_a^2 \tag{91}$$

A equação (91) é minimizada caso $\psi_{h+j} = \psi_{h+j}^*, j = 0,1,\cdots,h$. Portanto, a previsão do EQM mínimo é dada por:

$$\hat{Z}_t(h) = \psi_h a_t + \psi_{h+1} a_{t-1} + \cdots = \sum_{j=0}^{\infty} \psi_{h+j} a_{t-j} \tag{92}$$

E o erro de previsão apresentado na equação (90):

$$e_t(h) = a_{t+h} + \psi_1 a_{t+h-1} + \cdots + \psi_{h-1} a_{t+1} \tag{93}$$

Ou, também, tem-se que:

$$Z_{t+h} = e_t(h) + \hat{Z}_t(h), h \geq 1 \tag{94}$$

$$E[Z_{t+h}] = E[Z_{t+h} \mid Z_t, Z_{t-1}] \tag{95}$$

O método de mínimos quadrados tem o objetivo de estimação de um modelo ARIMA (p,d,q). Para isso, passa-se a achar o vetor auto-regressivo $\underset{\sim}{\varphi}$ e o vetor dos parâmetros médias móveis $\underset{\sim}{\theta}$, tais que minimizem a soma das diferenças quadradas entre os pontos observados na amostra e o esperado pela estimativa obtida com os parâmetros estimados. Simbolicamente, deve-se achar $\underset{\sim}{\phi}$, $\underset{\sim}{\theta}$, de modo que $S(\underset{\sim}{\phi}, \underset{\sim}{\theta}) = \Sigma a_t^2$ seja um mínimo.

2.2.3 Modelos Sazonais

Esses modelos são uma adição dos modelos ARIMA, ou seja, é um modelo ARIMA mais a Sazonalidade, chamado de SARIMA, podendo ser escrito como: $ARIMA(p,d,q)(P,D,Q))$ ou $SARIMA(p,d,q)(P,D,Q))$, no qual as letras minúsculas representam a ordem do modelo e diferenciação ARIMA; e as maiúsculas representam a ordem dos modelos e diferenciação Sazonal.

Para Farias, Rocha; Lima *(2000),* as Séries Temporais apresentam comumente padrões de comportamento periódico, ou seja, repetem-se a cada *s* períodos de tempo $(s > 1)$. Um dos casos mais comuns de dados periódicos é a série sazonal, sendo aquela em que existe o período de repetição. As Séries Temporais sazonais, em geral, exibem intervalos de tempo de um mês e período anuais $(s = 12)$; no entanto, é possível ocorrer casos em que o período é trimestral $(s = 4)$. O tratamento das Séries Temporais sazonais é normalmente feito usando-se o componente sazonal como fator de ajustamento. Todavia, alguns dos métodos de ajustamento sazonal não são adequados para fazer previsão com modelos de Séries Temporais, visto que o ajustamento sazonal provoca a perda de informações cruciais para o processo de previsão com os modelos de Séries Temporais.

Para esses autores, os modelos *ARIMA* exploram a autocorrelação entre os valores da Série em instantes sucessivos; mas, quando os dados são observados em períodos inferiores a um ano, a Série também pode apresentar autocorrelação para uma estação de sazonalidade *s*.

Os modelos que contemplam as Séries, os quais apresentam autocorrelação sazonal são conhecidos como *SARIMA*.

Os modelos *SARIMA* contêm uma parte não sazonal, com parâmetros (p,d,q); e uma sazonal, com parâmetros (P,D,Q). O modelo $ARMA(p,q)(P,Q)$ é dado por:

$$\begin{aligned}&(1-\phi_1 B-\ldots-\phi_p B^p)(1-\Phi_1 B^S-\ldots-\Phi_{PS} B^{PS})Z_t\\&=(1-\theta_1 B-\ldots-\theta_q B^q)(1-\Theta_1 B^S-\ldots-\Theta_{QS} B^{QS})a_t\end{aligned} \tag{96}$$

Conforme Farias, Rocha; Lima *(2000)*, se a Série em estudo for não-estacionária com sazonalidade, temos um $ARIMA(p,d,q)(P,D,Q)$, no qual d representa a ordem de diferenciação não sazonal; e D, a ordem de diferenciação sazonal. A diferenciação sazonal pode ser realizada da seguinte forma:

$$\Delta^D Z = Z_t - Z_{t-s} \tag{97}$$

A ocorrência de estacionaridade está relacionada a uma queda brusca no valor dos picos sazonais. A FAC também mostrará se há componente sazonal, o que pode ser observado quando esta segue um padrão de picos e depressões ano a ano. Da mesma forma, a FAC mostrará a não-estacionaridade sazonal.

A diferenciação da Série em termos não sazonal e sazonal tem por objetivo torná-la estacionária, o que possibilita a aplicação da metodologia de Box & Jenkins.

É através da metodologia apresentada, que será analisado o preço dos principais grãos produzidos no Estado do Rio Grande do Sul, os quais estão apresentados no Capítulo 3 e Capítulo 4.

3 METODOLOGIA

Neste Capítulo, serão apresentadas as etapas da análise do comportamento dos preços dos principais grãos do Estado do Rio Grande do Sul, por meio da análise de Séries Temporais.

Dados: os dados analisados referem-se às culturas das Lavouras Temporárias do Estado, as quais são assim caracterizadas pelo seu curto ciclo vegetativo e por requererem um replantio após cada colheita. Analisaram-se 24 culturas pertencentes a essas lavouras desenvolvidas no Estado, através das médias anuais da quantidade produzida, no período de janeiro de 1995 a dezembro de 2006. Esses dados foram coletados no *site* do Instituto Brasileiro de Geografia e Estatística - IBGE e, por meio da estatística descritiva, identificaram-se as três culturas mais desenvolvidas, em toneladas. Optou-se pela quantidade produzida com a finalidade de identificar as culturas que são preferência na agricultura, consequentemente as quais se tem mais interesse em prever os seus preços futuros, devido a dedicação dos agricultores

Após identificar as principais culturas produzidas no Estado, e levar em consideração a necessidade de, no mínimo, 50 observações para analisar o comportamento da Série em estudo, utilizaram-se os dados mensais, de janeiro de 1995 a dezembro de 2006, das culturas de Soja, Arroz e Milho. Os referidos dados foram fornecidos pela Emater/RS-ASCAR. Essas Séries referem-se ao preço médio mensal dessas culturas, em Reais, visto que se deseja analisar o comportamento dos preços dos principais grãos do Estado.

Análise Descritiva: *r*ecorreu-se à Análise Descritiva para identificar as principais produções desenvolvidas no Estado, dentre as 24 culturas analisadas. Escolheu-se a três culturas com maior produção, em toneladas, através das medidas descritivas: média, desvio-padrão e coeficiente de variação.

Séries Temporais: **a** análise das Séries dos preços da Soja, Arroz e Milho foi feita através da Metodologia Box & Jenkins, descrita no Capítulo 2. Para apresentar as análises, seguiu-se o Ciclo Iterativo apresentado no Quadro 1.

A seleção das lavouras temporárias se dará pela análise das médias das quantidades produzidas, no período de 1995 a 2006, no Estado do Rio Grande do Sul, é possível identificar as culturas que apresentaram o maior número de produção em Toneladas. Ganham destaque, portanto:

Tabela 2- Média, em tonelada, da quantidade produzida nas lavouras temporárias de 1995 a 2006.

Lavouras Temporárias	Média da Quantidade Produzida de 1995 a 2006	Classificação
Milho	203.354.148	1º
Soja	174.086.546	2º
Arroz	74.519.863	3º
Mandioca	61.460.898	4º
Trigo	36.753.201	5º
Batata - inglesa	13.245.288	6º
Feijão	10.725.246	7º
Fumo	7.430.523	8º
Cebola	6.899.786	9º
Tomate	4.193.963	10º
Melancia	3.695.180	11º

Fonte: IBGE – Instituto Brasileiro de Geografia e Estatística

Através da Tabela 2, pode-se perceber quais foram as culturas selecionadas para a análise, devido à enorme diferença das três primeiras culturas em relação às demais. A partir deste momento, serão apresentadas as análises dos preços das Séries dessas principais culturas: Soja, Arroz e Milho, fornecidos pela Emater/RS – ASCAR. Os dados que compõem cada uma das três Séries são referentes ao período de janeiro de 1995 a dezembro de 2006. Como a Emater forneceu os dados até o mês de maio de 2007, os dados referentes a 2007 servirão para comparação e cálculo do Erro Quadrado Médio (EQM), que será muito útil na decisão do modelo que represente a Série. Isso porque um modelo bem ajustado é aquele que apresenta um pequeno EQM.

4 RESULTADOS E DISCUSSÕES

Neste capítulo será apresentado os resultados da análise dos preços das lavouras temporárias, destacadas no Estado do Rio Grande do Sul de 1995 a 2006

4.1 Soja

A soja é uma cultura de origem Asiática que veio para a Europa e depois para o continente americano. As primeiras referências sobre a cultura em São Paulo datam de 1892 e foram realizadas pelo Dr. Franz Wilhelm Daffert, no Instituto Agronômico de Campinas. Em 1900, a Secretaria da Agricultura distribuía alguns quilos de sementes de soja. No Rio Grande do Sul, a cultura de soja chegou em 1917 e teve como centro de irradiação o município de Santa Rosa.

Entre os produtos agrícolas que presentemente alimentam o mundo, a soja vem ocupando uma posição de crescente destaque, com extraordinária expansão, considerando-se sua função de produção de grãos de uso na alimentação. Em uma década, a produção de soja é capaz de duplicar, no Brasil, por sua fácil adaptação às condições climáticas. Apesar dos problemas de condições climáticas, pragas e moléstias, que impõem à agricultura novos desafios, o crescimento de novas áreas de cultivo segue expandindo-se por todos os continentes. Hoje, a soja é a mais importante oleaginosa em cultivo extensivo ou rotacional. A soja produz mais proteínas por hectares do que qualquer outra planta de lavoura. O desenvolvimento de novas áreas de produção e a difusão do consumo de soja pode ser fundamental para o suprimento alimentar diário de grande parte da população mundial, nas diferentes latitudes.

Segundo a Emater, no segundo semestre de 2001 e no primeiro semestre de 2002, os preços da soja no mercado interno foram mais atrativos que os do milho. As cotações do dólar deram sustentação aos preços da soja no mercado interno, fator este que acabou influenciando os produtores a optarem por plantar mais soja do que milho na próxima temporada, ou seja, no ano de 2002. Outro fator relevante foi a primeira estimativa para safra 2002/03, a qual apontava que o Brasil, a Argentina e o Paraguai produziriam juntos 80,7 milhões de toneladas. Os americanos, principais produtores mundiais de soja, seriam responsáveis sozinhos por 77,8 milhões de toneladas na próxima safra.

Ainda, para a Emater, o que baliza os preços da soja no mercado internacional é o comportamento da oferta e da demanda, que faz as cotações de Chicago oscilarem para cima ou para baixo, dependendo do tamanho da safra. Outros fatores também contribuem para a formação dos preços internamente, tais como, logística, taxa de câmbio, taxa de juros, fatores climáticos. Os sojicultores devem sempre ficar atentos a todas as informações relacionadas ao mercado e ao processo produtivo, para assim tomar a decisão mais adequada possível em relação aos investimentos a realizar na sua propriedade, nunca esquecendo de analisar sistematicamente as atividades, principalmente levando-se em consideração os custos da formação da lavoura e o comportamento dos preços no mercado.

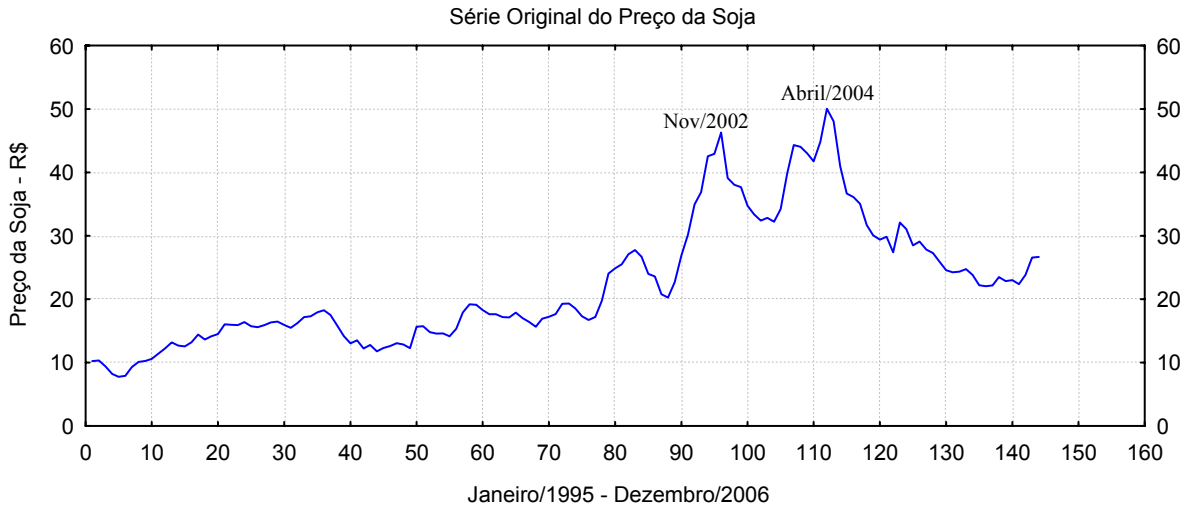


Figura 4 - Comportamento do preço da Soja, para o período de janeiro de 1995 a dezembro de 2006. Fonte dos dados é a Emater/RS – ASCAR.

Para analisar o comportamento da Série do preço da soja, de janeiro de 1995 a dezembro de 2006, no Rio Grande do Sul, o primeiro passo será desenvolver um Gráfico da Série Original, que possibilite observar o comportamento da Série ao longo do período de análise. Na Figura 4, encontra-se a Série do preço da soja.

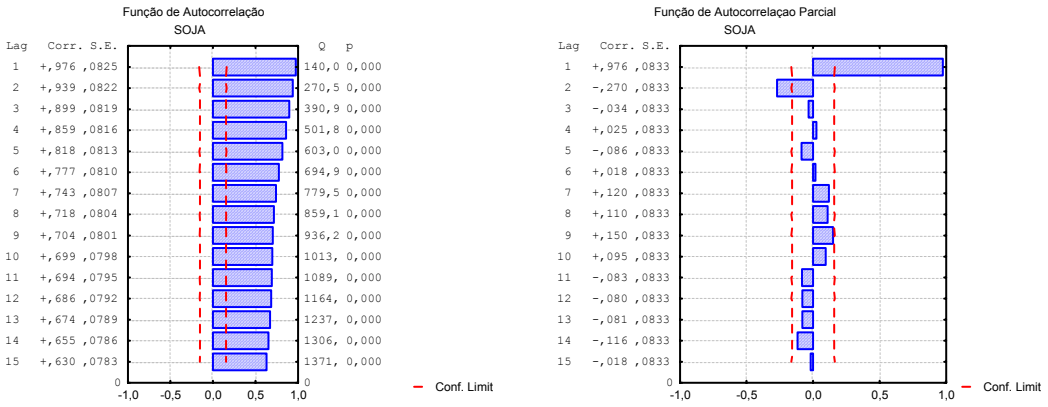


Figura 5- Funções de Autocorrelação da Série Original

É possível observar que a Série não assume um comportamento estacionário por apresentar uma tendência crescente. A Série também apresentou dois picos, que representam a alta do seu preço no mês de novembro de 2002, e voltou a apresentar esse pico em abril de 2004. Percebe-se que, a partir de abril de 2004, esses picos não tornaram a se repetir. Através das Funções de Autocorrelação da Série Original, é notável a não-estacionariedade da Série, uma vez que a FAC apresenta decaimento muito lento.

Como a série original e as funções de autocorrelação indicam uma série com tendência, sua significância será apresentada através do Teste do Sinal (Cox-Stuart), para testar as hipóteses:

H_0 : não existe tendência

H_1 : existe tendência

Como o $p = 0{,}000$, para $\alpha = 5\%$, $p < \alpha$, rejeita-se a hipótese nula, logo a tendência apresentada na série original é significativa.

Além da componente tendência a série apresenta sazonalidade, a qual pode ser verificada através do Teste não-pamétrico de Kruskal-Wallis, em que:

H_0 : não existe sazonalidade determinística;

H_1 : existe tendência sazonalidade determinística.

Como o $p = 0{,}000$, para $\alpha = 5\%$, $p < \alpha$, rejeita-se a hipótese nula, logo a sazonalidade anual apresentada na série original é significativa.

Segue, na Figura 6, a série original e diferenciada, na qual uma diferença foi suficiente para estacionarizá-la.

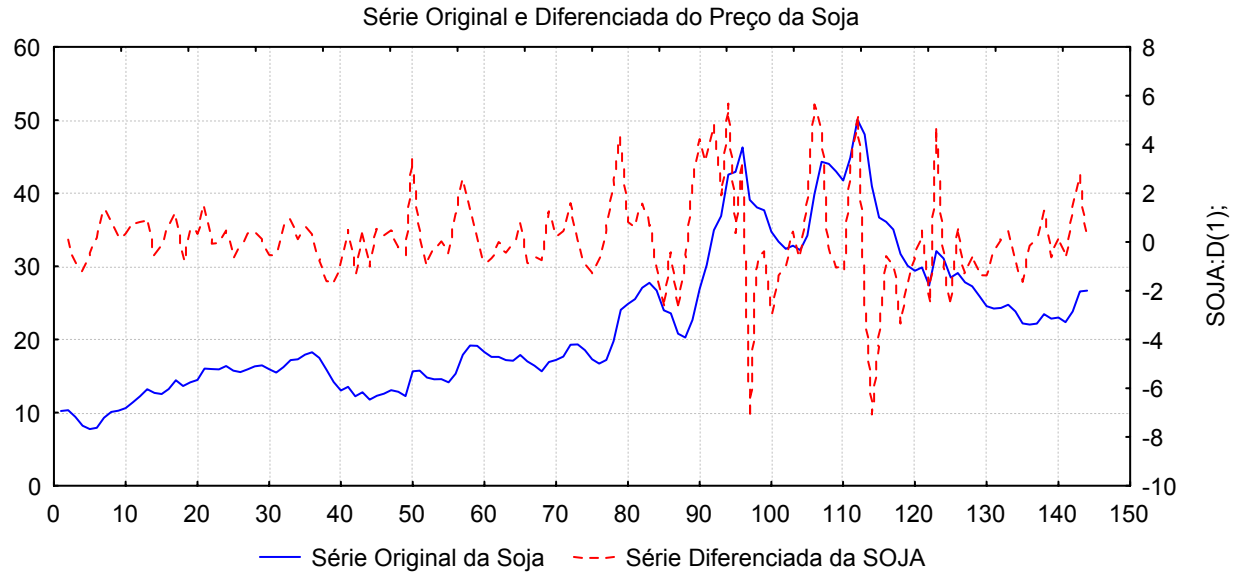


Figura 6- Série Original e Diferenciado Preço da Soja

Transformada a série em estacionária, o próximo procedimento é identificar os modelos representativos desta. Para essa série, dentre os vários modelos investigados, foram encontrados seis modelos significativos. É importante salientar que quando uma Série Temporal apresenta *outliers*, a metodologia recorre aos modelos com intervenção; mas, essa Série dispensa esse auxílio, pois quando utilizados, os modelos apresentaram-se sem significância. Isso pode estar relacionado à sazonalidade existente na Série. Como dentre os modelos significativos precisa-se optar por um, recorre-se aos critérios penalizadores AIC e SBC e ao Erro Quadrado Médio, como regra de decisão. Sabe-se que o modelo que mais se ajusta à série é o que apresentar os menores valores correspondentes a esses critérios.

O modelo que melhor representa a Série em estudo, identificado pelo menor Erro Quadrático Médio e AIC e SBC, é o $ARIMA(0,1,2)(0,1,0)_{12}$. Analisando os critérios de estacionariedade e invertibilidade, estes são satisfeitos, visto que os parâmetros de cada modelo encontrado, em módulo, são menores que um.

Reescrevendo o modelo encontrado, temos que o preço da soja é representado:

$$Z_t = a_t - 0{,}6414a_{t-1} - 0{,}358a_{t-2}$$

Tabela 3- Modelos significativos para a Série do Preço da Soja no Estado do Rio Grande do Sul.

Modelos	Estimativas dos parâmetros do modelo	P	EQM	AIC	SBC
ARIMA $(2,1,0)(0,1,0)_{12}$	$\phi_1 = 0{,}3859$ $\phi_2 = 0{,}1938$		39,91	40,67	46,59
SARIMA $(2,1,1)(1,1,0)_{12}$	$\phi_1 = 0{,}2816$ $\phi_2 = 0{,}1960$ $\theta_1 = 0{,}9999$ $\Phi_1 = 0{,}2605$		8,45	26,09	37,92
ARIMA $(2,1,2)_{12}$	$\phi_1 = 0{,}9379$ $\phi_2 = 0{,}5042$ $\theta_1 = 0{,}8540$ $\theta_2 = 0{,}7648$		4,51	367,13	378,95
SARIMA $(2,1,2)(0,0,1)_{12}$	$\phi_1 = 0{,}6495$ $\phi_2 = 0{,}8420$ $\theta_1 = 0{,}8612$ $\theta_2 = 0{,}9964$ $\Theta_1 = 0{,}2828$		11,01	355,71	370,52
SARIMA $(0,1,1)(0,1,1)_{12}$	$\theta_1 = 0{,}9999$ $\Theta_1 = 0{,}2678$		3,43	347,29	353,20
SARIMA *(0,1,2)(0,1,0)* $_{12}$	$\theta_1 = 0{,}6414$ $\theta_2 = 0{,}358$		***2,12***	***25,62***	***31,54***

O próximo passo é analisar as Funções de Autocorrelação, na Figura 7, na qual é possível verificar que o modelo ajustado apresenta alguns *lags* fora do intervalo de confiabilidade. Como a FAC e FACP aceitam até 5% dos seus valores fora do intervalo de confiabilidade, o modelo é significativo.

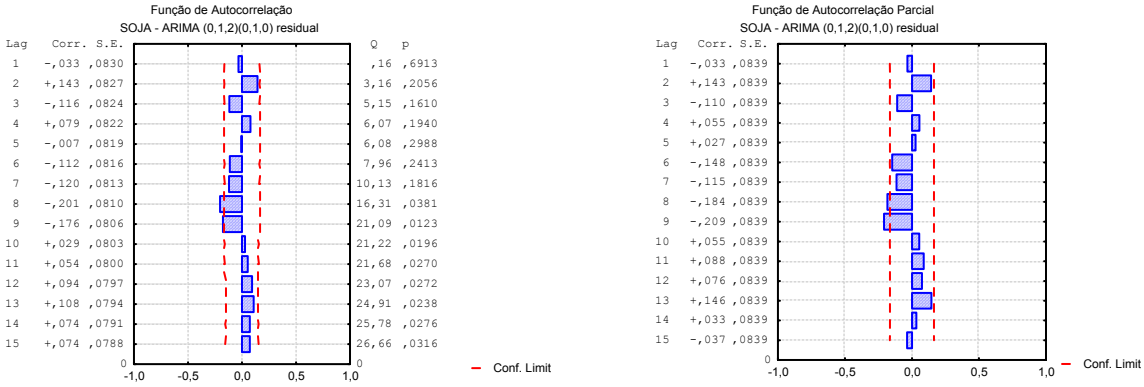


Figura 7- Função de Autocorrelação e Autocorrelação Parcial Residual

Identificado o melhor modelo, pode se realizar as previsões da Série.

Tabela 4 – Previsão da Serie Original da Soja, no RS, para o Ano de 2007.

Meses	Previsões	Valores Reais	Inferior 95%	Superior 95%	Erro Padrão
Jan/07	26,52	26,58	22,91	30,13	1,826
Fev/07	26,62	27,49	20,53	32,71	3,081
Mar/07	26,72	27,63	18,90	34,55	3,956
Abr/07	26,82	27,33	17,59	36,06	4,671
Mai/07	26,93	26,39	16,47	37,39	5,291
Jun/07	27,03		15,47	38,58	5,846
Jul/07	27,13		14,57	39,69	6,353
Ago/07	27,23		13,74	40,72	6,822
Set/07	27,33		12,97	41,69	7,262
Out/07	27,43		12,25	42,61	7,677
Nov/07	27,53		11,57	43,49	8,071
Dez/07	27,63		10,93	44,33	8,447

Transformada a Série original em estacionária, encontrou-se o modelo mais adequado através dos critérios penalizadores e identificaram-se os parâmetros menores que um e as Funções de Autocorrelação entre os limites de confiabilidade. Desse modo, pode-se realizar as previsões para o ano de 2007. As previsões, até maio, serviram para o cálculo do Erro Quadrado Médio, como um recurso adicional, aumentando a confiabilidade no modelo ajustado.

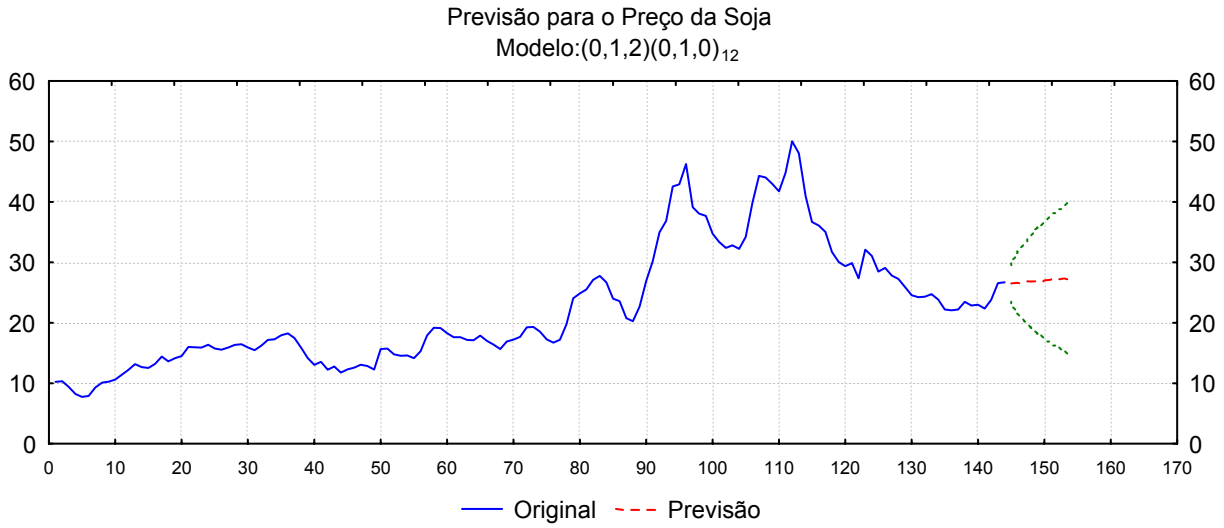


Figura 8 – Série Original e a Previsão para o ano de 2007.

Através da Tabela 4 e da Figura 8, é possível visualizar as previsões mensais para o ano de 2007. Além disso, podem-se observar os valores reais até maio de 2007, o erro padrão e os limites com 95% de confiabilidade. Comparando os valores previstos com os reais, conclui-se que esta metodologia conseguiu captar o padrão encontrado na Série em estudo e fazer a previsão para o preço da soja no ano de 2007, através do modelo selecionado: $Z_t = a_t - 0{,}64a_{t-1} - 0{,}358a_{t-2}$

4.2 Arroz

O arroz, cujo nome científico é Oryza Sativa, é o segundo alimento mais consumido no mundo e seu cultivo é tão antigo quanto à própria civilização. Diversos autores acreditam que ele seja originário da Ásia e as primeiras culturas iniciadas há sete mil anos. Porém, referências concretas aparecem por volta de 2800 a.C., na China. O arroz é rico em carboidratos e, quando combinado com feijão ou outras leguminosas, transforma-se em rica fonte de proteínas. Da Ásia, seu provável berço, foi impondo a sua presença na Europa e no norte da África, e batizado como "aruz" na Espanha, para onde foi levado pelos árabes, no século XV.

O arroz foi introduzido no Brasil pela frota de Pedro Álvares Cabral. No Rio Grande do Sul, lavouras comerciais surgem em Pelotas, no ano de 1904. Segundo o Instituto Rio Grandense do Arroz - IRGA, o Rio Grande do Sul é responsável por 50% do arroz produzido no país; outros produtores são de Mato Grosso, segundo maior produtor; Santa Catarina; Tocantins e Goiás. O Brasil é o grande produtor mundial de arroz (11 milhões de toneladas) e o seu consumo é o terceiro mais elevado da América Latina (entre 40 a 45 quilos por habitante). Já na era primitiva, os japoneses trançavam a palha de arroz na confecção de tatamis. Existem, pelo menos, nove variedades de arroz: Vermelho, Parboilizado, Basmati, Tailandês, Selvagem, Japonês, Ráris, Agulha e Integral.

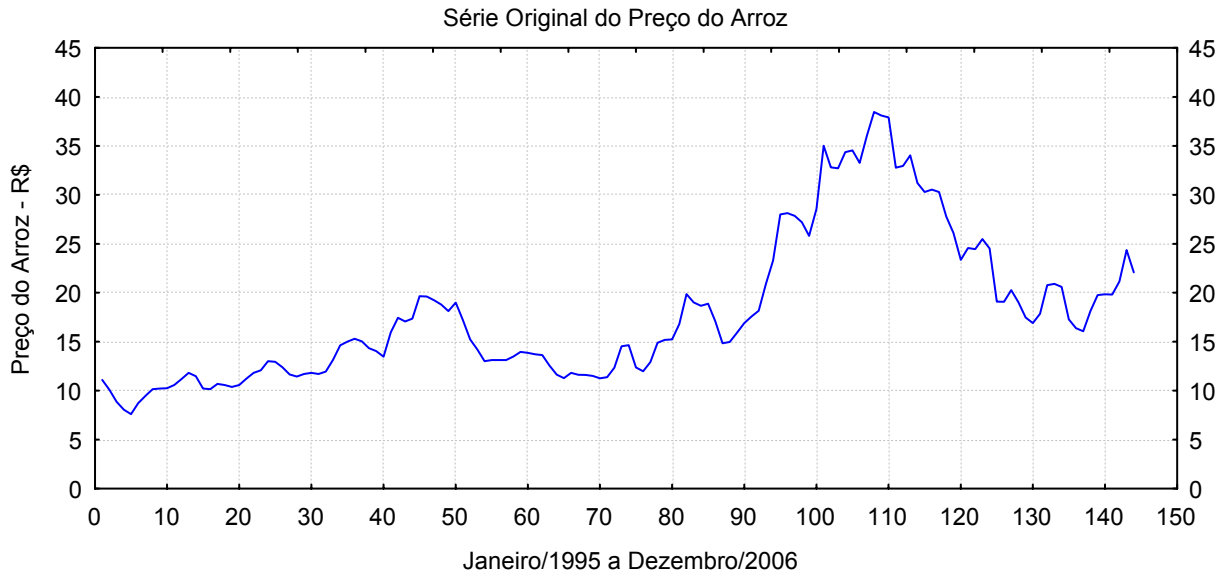


Figura 9 - Comportamento do preço do Arroz, para o período de janeiro de 1995 a dezembro de 2006.
Fonte dos dados é a Emater/RS – ASCAR.

Para analisar o comportamento do preço do arroz no Estado do Rio Grande do Sul, primeiramente, realizou-se o gráfico da Série Original, conforme a Figura 9. Nesse gráfico, é possível identificar que a Série não é estacionária, por haver em seu comportamento uma tendência crescente. Nota-se, também, através da análise visual, que esta apresenta sazonalidade.

É possível observar, na figura 10, que a Função de Autocorrelação não decresce rapidamente para zero.

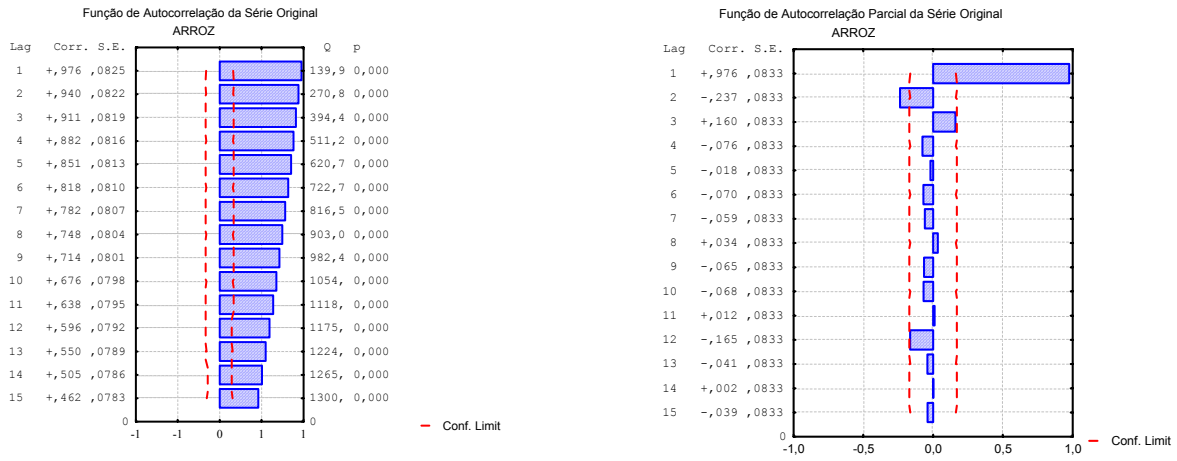


Figura 10- Funções de Autocorrelação da Série Original

Recorrendo as Funções de Autocorrelação, complementa-se que Série do Preço do Arroz no Estado do Rio Grande do Sul não é estacionária.

Para verificar se esta tendência apresentada é significativa aplicou-se o Teste do Sinal (Cox-Stuart), para testar as hipóteses:

H_0 :não existe tendência

H_1 : existe tendência

Como o $p = 0{,}000$, para $\alpha = 5\%$, $p < \alpha$, rejeita-se a hipótese nula, logo a tendência apresentada na série original é significativa.

Baseado nessa informação, como medida principal para ser possível aplicar a metodologia de Box & Jenkins, o próximo procedimento será estacionarizá-la. Além disso, pode-se perceber que a Série também apresenta sazonalidade anual ao longo do período.

A sazonalidade pode ser verificada através do Teste não-pamétrico de Kruskal-Wallis, em que:

H_0 :não existe sazonalidade determinística;

H_1: existe tendência sazonalidade determinística.

Como o $p = 0{,}000$, para $\alpha = 5\%$, $p < \alpha$, rejeita-se a hipótese nula, logo a sazonalidade anual apresentada na série original é significativa.

Transformando a Série Original, por meio de uma única diferença, é possível observar que a Série Diferenciada possui um comportamento estacionário, ou seja, oscila em torno de uma média.

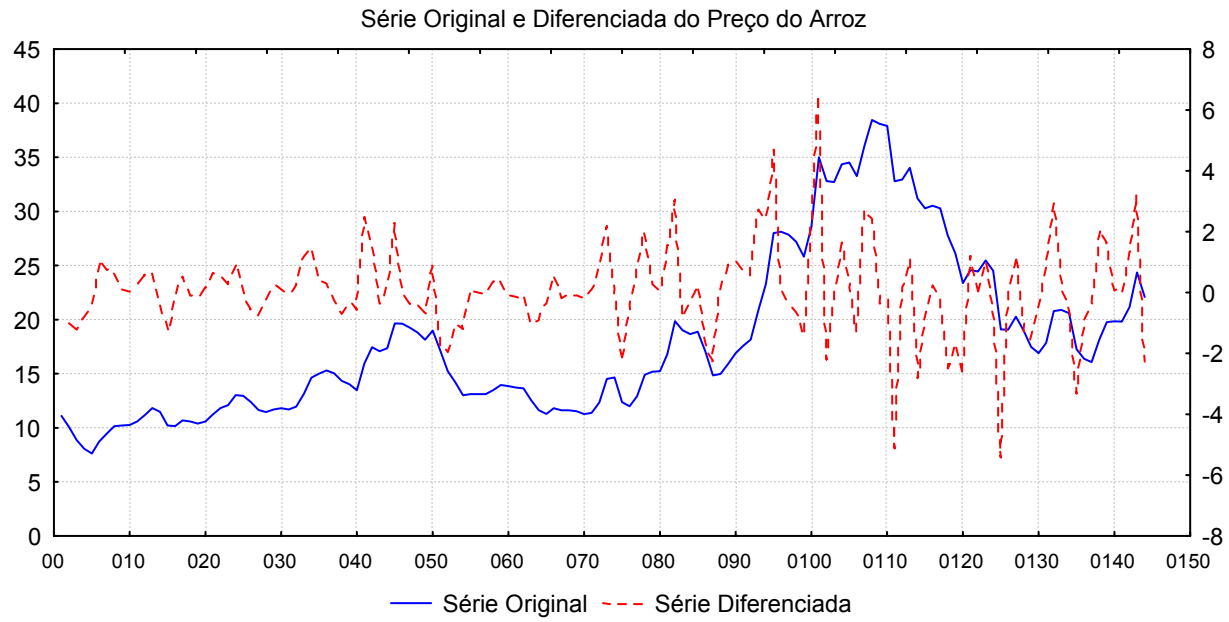


Figura 11- Série Original e Diferenciada

Após estacionarizada a Série, o passo seguinte é identificar o modelo significativo que a represente. Através da análise da Tabela 5, é possível observar que foram identificados quatro modelos significativos, dos quais se deve escolher apenas um. Como regra de decisão, tem-se os Critérios Penalizadores, o AIC e o SBC, que buscam encontrar modelos parcimoniosos, e o Erro Quadrado Médio (EQM). Quanto menor o EQM, mais condizente esse modelo está com a realidade da Série. Logo, o modelo escolhido é o SARIMA $(1,1,0)(1,0,1)_{12}$, por apresentar menor EQM e menores AIC e SBC.

Tabela 5 - Modelos significativos para a Série do Preço do Arroz no Estado do Rio Grande do Sul.

Modelos	Estimativas dos parâmetros do modelo	EQM	AIC	SBC
SARIMA (1,1,0)(1,0,1) $_{12}$	$\phi_1 = 0{,}19492$ $\Phi_1 = -0{,}8401$ $\Theta_1 = -0{,}9768$	**4,22**	**283,38**	**292,27**
ARIMA (1,1,0)(0,1,0) $_{12}$	$\phi_1 = -0{,}2767$	24,56	-336,05	-333,10
ARIMA (0,1,1)(0,1,0) $_{12}$	$\theta_1 = 0{,}90134$	17,25	348,69	351,65
ARIMA (0,1,1)(1,1,1) $_{12}$	$\theta_1 = 0{,}91879$ $\Phi_1 = -0{,}7551$ $\Theta_1 = -0{,}9088$	5,03	348,38	357,25

Esse modelo também satisfaz os critérios de estacionariedade e invertibilidade, pois durante a análise do parâmetro que representa a ordem do modelo, em módulo, observou-se que todos são menores que um. Na Série original, percebe-se que o preço do arroz teve uma alta, entre maio de 2003 e maio de 2004, mas o modelo conseguiu captar o comportamento da Série sem a necessidade da aplicação de uma intervenção.

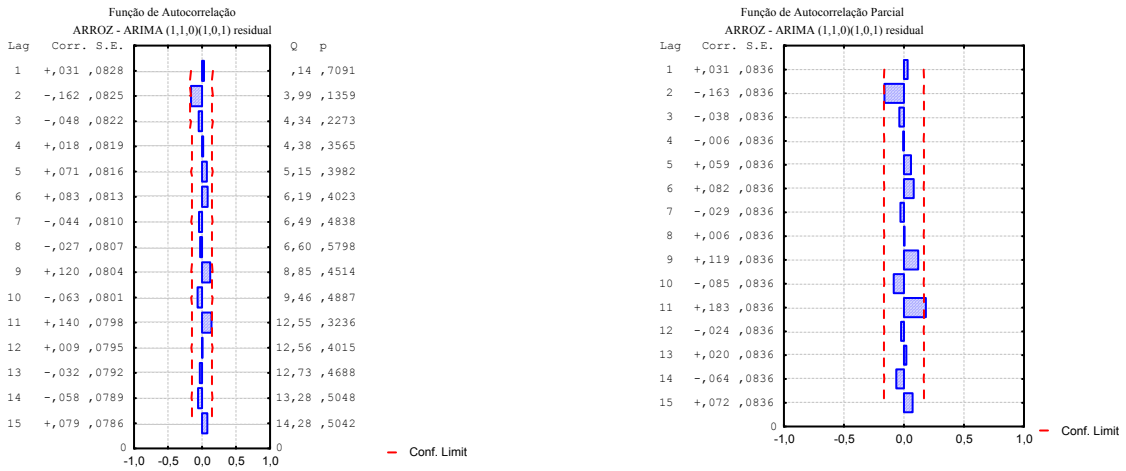


Figura 12- Função de Autocorrelação e Autocorrelação Parcial Residual

A adequação do modelo está confirmada na Figura 12, em que se pode perceber que os dados encontram-se dentro dos limites de confiabilidade.

Na Tabela 6, apresenta-se as previsões para todo o ano de 2007, bem como os valores reais, os limites de Confiabilidade em 95% e o erro.

Tabela 6 – Previsão da Serie Original do Arroz, no RS, para o ano de 2007.

Meses	Previsões	Valores Reais	Inferior 95%	Superior 95%	Erro Padrão
Jan/07	20,98	21,75	17,98	23,98	1,52
Fev/07	20,92	19,71	15,80	26,04	2,59
Mar/07	20,86	21,01	14,15	27,57	3,39
Abr/07	20,80	20,74	12,71	28,89	4,09
Mai/07	20,74	20,56	11,39	30,10	4,73
Jun/07	20,68		10,13	31,24	5,34
Jul/07	20,62		8,91	32,34	5,93
Ago/07	20,57		7,72	33,41	6,50
Set/07	20,51		6,55	34,46	7,06
Out/07	20,45		5,40	35,50	7,61
Nov/07	20,39		4,25	36,52	8,16
Dez/07	20,33		3,11	37,55	8,71

Figura 13 – Série Original e a Previsão para o ano de 2007.

Na Figura 13, encontra-se a Série Original e a Previsão para o Modelo SARIMA *(1,1,0)(1,0,1)*$_{12}$, adequado para representar a Série do Arroz no Estado do Rio Grande do Sul, e recorrer a ele para fazer as previsões para o Estado no ano de 2007.

Através da Tabela 6 e da Figura13, é possível visualizar as previsões mensais do preço do arroz para o ano de 2007. Além disso, podem-se observar os valores reais até maio de 2007, o erro padrão e os limites com 95% de confiabilidade. Comparando os valores previstos com os reais, conclui-se que esta metodologia conseguiu captar o padrão encontrado na Série em estudo e fazer a previsão para o preço da soja no ano de 2007, através do modelo selecionado: $\left(1+0{,}19492B^{1}\right)\left(1-B\right)\left(1-0{,}8401B^{12}\right)\left(1-B\right)^{12}Z_t=\left(1-0{,}97B^{12}\right)a_t$

4.3 Milho

Segundo a Embrapa, o milho é um conhecido cereal cultivado em grande parte do mundo. É extensivamente utilizado como alimento humano ou ração animal, devido às suas qualidades nutricionais. Todas as evidências científicas levam a crer que seja uma planta de origem americana, já que aí era cultivada desde o período pré-colombiano. É um dos alimentos mais nutritivos que existem, contendo quase todos os aminoácidos conhecidos, sendo exceções a lisina e o triptofano. Tem alto potencial produtivo e é bastante responsivo à tecnologia. Seu cultivo geralmente é mecanizado, beneficiando-se muito de técnicas modernas de plantio e colheita. A produção mundial de milho chegou a 600 milhões de toneladas, em 2004. No Brasil, um grande produtor e exportador, São Paulo e Paraná são os estados líderes na sua produção. Atualmente, somente cerca de 5% da produção brasileira destina-se ao consumo humano e, mesmo assim, de maneira indireta na composição de outros produtos. Isso se deve principalmente à falta de informação sobre o milho e à ausência de uma maior divulgação de suas qualidades nutricionais, bem como aos hábitos alimentares da população brasileira, que privilegia outros grãos.

Para analisar a Série do preço milho no Estado do Rio Grande do Sul, utilizou-se também a Metodologia Box & Jenkins, seguindo os mesmos procedimentos da análise das Séries da soja e do arroz. Primeiramente, analisou-se visualmente a Série Original, observando que esta apresenta comportamento tendencioso ao longo do tempo.

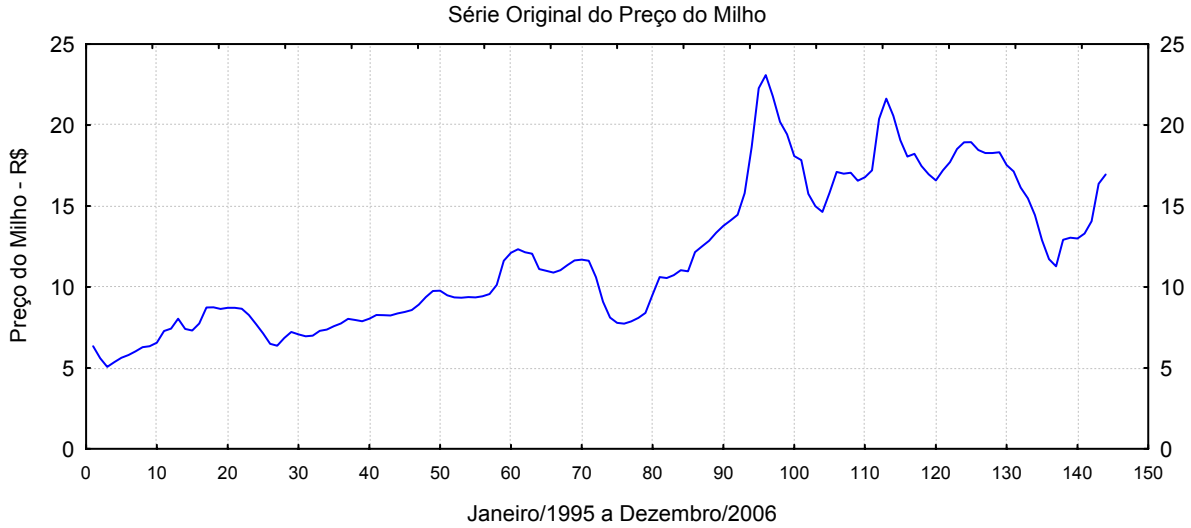


Figura 14 - Comportamento do preço do milho, para o período de janeiro de 1995 a dezembro de 2006.
Fonte dos dados é a Emater/RS – ASCAR.

Como foi feita uma análise apenas visual para comprovar que realmente a Série não se encontra estacionária, serão analisadas as Funções de Autocorrelação.

Através das FAC, comprova-se o descrito acima: a Série do Preço do Milho no Rio Grande do Sul não é estacionária, pois ela não apresenta um decaimento rápido dos *lags*. Portanto, serão realizadas diferenças sucessivas até se encontrar uma Série estacionária.

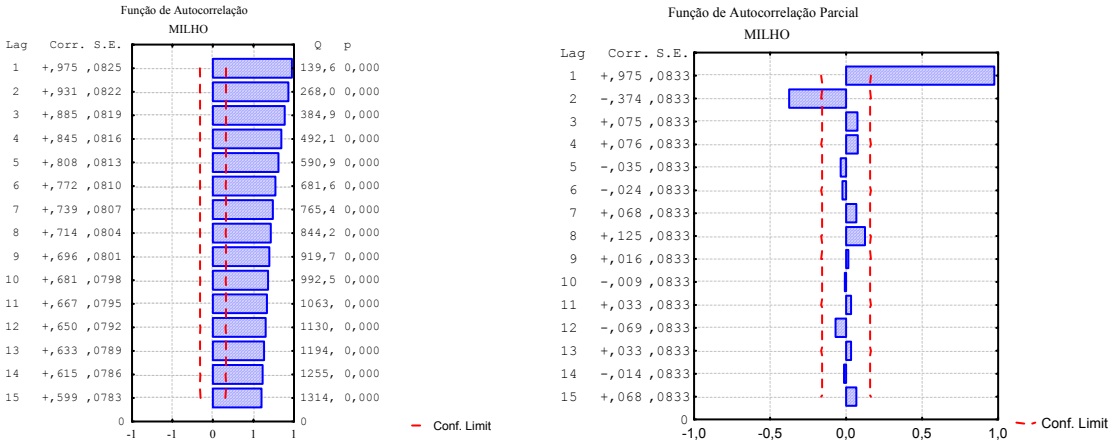


Figura 15- Funções de Autocorrelação da Série Original

Para verificar a significância da tendência apresentada aplicou-se o Teste do Sinal (Cox-Stuart), para testar as hipóteses:

H_0 :não existe tendência

H_1 : existe tendência

Como o $p = 0{,}000$, para $\alpha = 5\%$, $p < \alpha$, rejeita-se a hipótese nula, logo a tendência apresentada na série original é significativa.

Pela análise visual também é identificada uma sazonalidade anual, a qual pode ser verificada através do Teste não-pamétrico de Kruskal-Wallis, em que:

H_0 :não existe sazonalidade determinística;

H_1 : existe tendência sazonalidade determinística.

Como o $p = 0{,}000$, para $\alpha = 5\%$, $p < \alpha$, rejeita-se a hipótese nula, logo a sazonalidade apresentada na série original é significativa.

Após comprovada a tendência e a sazonalidade, o próximo procedimento será realizar diferenças sucessivas até estacionarizar a série original.

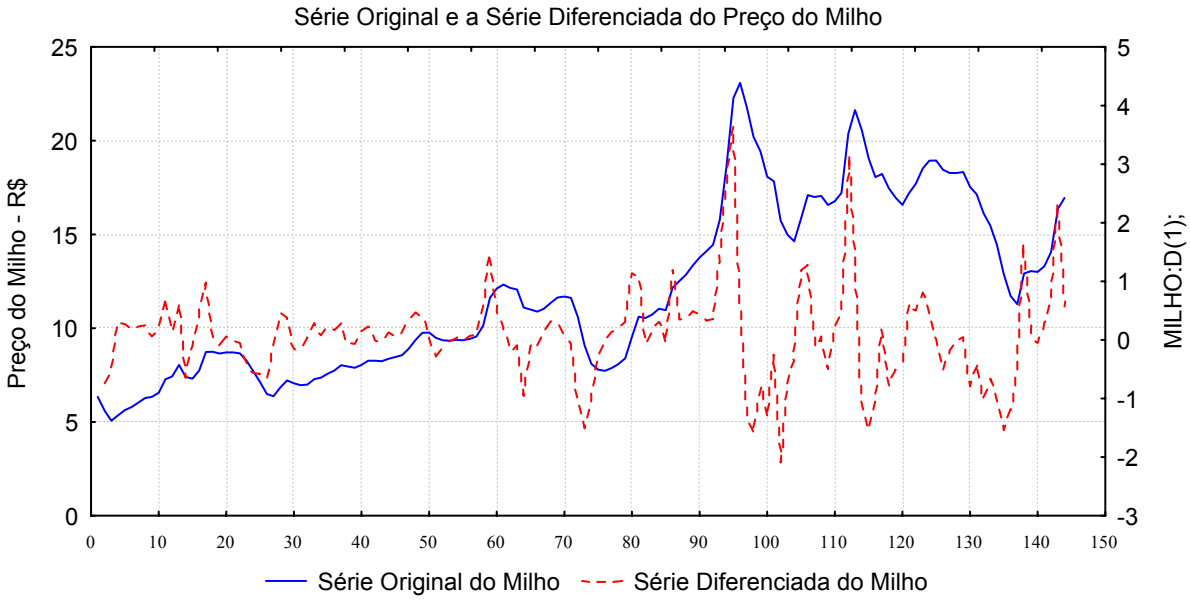


Figura 16- Série Original e Diferenciada

A Figura 16 mostra a Série Original e a Diferenciada, em que foi necessária a realização de apenas uma diferença, sendo desnecessária mais diferenciações, para não se pecar no excesso de diferenças realizadas.

Depois de encontrada a Série estacionária, pode-se partir para a identificação do modelo que represente a Série. Dentre os modelos investigados, apenas três se mostraram significativos, sendo que alguns foram descartados por realizarem previsões constantes. Esse não é o objetivo, pois nunca se desenvolverá um estudo para prever algo constante, uma vez que não ocorrerá a hipótese de, em todos os meses do ano, o preço do milho ser igual, inclusive nos centavos.

Tabela 7 - Modelos significativos para a Série do Preço do Milho no Estado do Rio Grande do Sul.

Modelos	Estimativas dos parâmetros do modelo	EQM	AIC	SBC
ARIMA $(1,1,1)(0,1,0)_{12}$	$\phi_1 = 0{,}59877$ $\theta_1 = 0{,}96874$	8,00	337,23	343,14
SARIMA $(2,1,1)(0,1,0)_{12}$	$\phi_1 = 0{,}70394$ $\phi_2 = 0{,}2065$ $\theta_1 = 0{,}95807$	**0,64**	**323,55**	**332,42**

Na Tabela 7, encontram-se os modelos ajustados, com seus respectivos Parâmetros, EQM e Critérios Penalizadores, AIC e SBC. Através dos critérios estabelecidos, o modelo

que melhor representa a Série do preço do milho, sendo o mais adequado para realizar as previsões, é o modelo *ARIMA (2,1,1)(0,1,0)*$_{12}$.

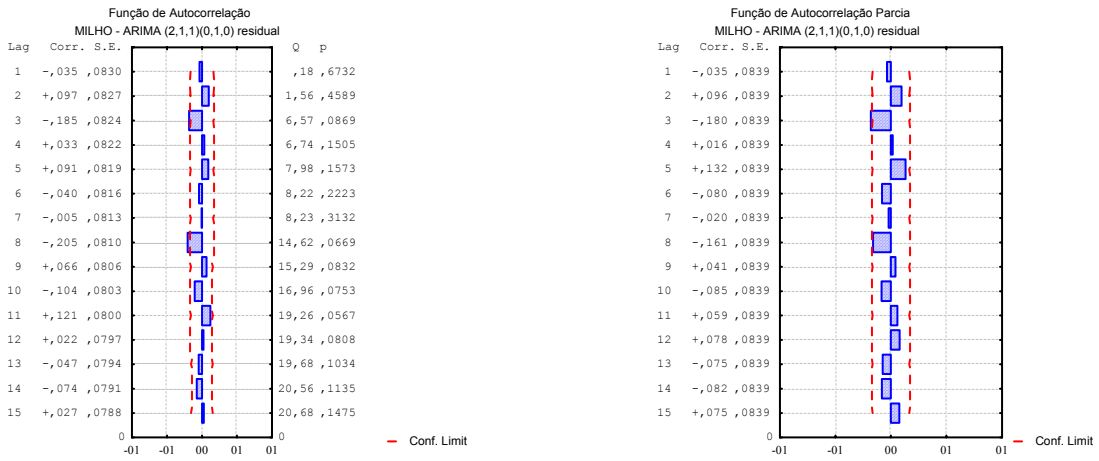


Figura 17- Função de Autocorrelação e Autocorrelação Parcial Residual

A adequação do Modelo pode ser verificada através das Funções de Autocorrelação, em que os resíduos encontram-se, no mínimo, 95% dentro dos limites de confiabilidade. Os Parâmetros dos modelos são descritos conforme a ordem estabelecida na Tabela 7, em que é possível identificar que todos os modelos significativos satisfazem a condição de estacionariedade e invertibilidade, porque os valores destes, em módulo, são menores que um.

Tabela 8 – Previsão da Serie Original do Milho, no RS, para o Ano de 2007.

Meses	Previsões	Valores Reais	Inferior 95%	Superior 95%	Erro Padrão
Jan/07	16,89	16,55	15,54	18,25	0,69
Fev/07	16,76	16,68	14,03	19,48	1,38
Mar/07	16,69	16,92	12,74	20,64	2,00
Abr/07	16,69	16,24	11,71	21,68	2,52
Mai/07	16,73	16,22	10,84	22,63	2,98
Jun/07	16,78		10,06	23,49	3,40
Jul/07	16,82		9,33	24,31	3,79
Ago/07	16,86		8,63	25,10	4,17
Set/07	16,91		7,94	25,87	4,53
Out/07	16,95		7,27	26,62	4,89
Nov/07	16,99		6,61	27,37	5,25
Dez/07	17,03		5,95	28,11	5,61

Na Tabela 8, encontram-se as previsões realizadas com o modelo ajustado *ARIMA (2,1,1)(0,1,0)*$_{12}$. Encontram-se, também, na tabela, os valores reais e os intervalos com 95% de confiabilidade, além do erro.

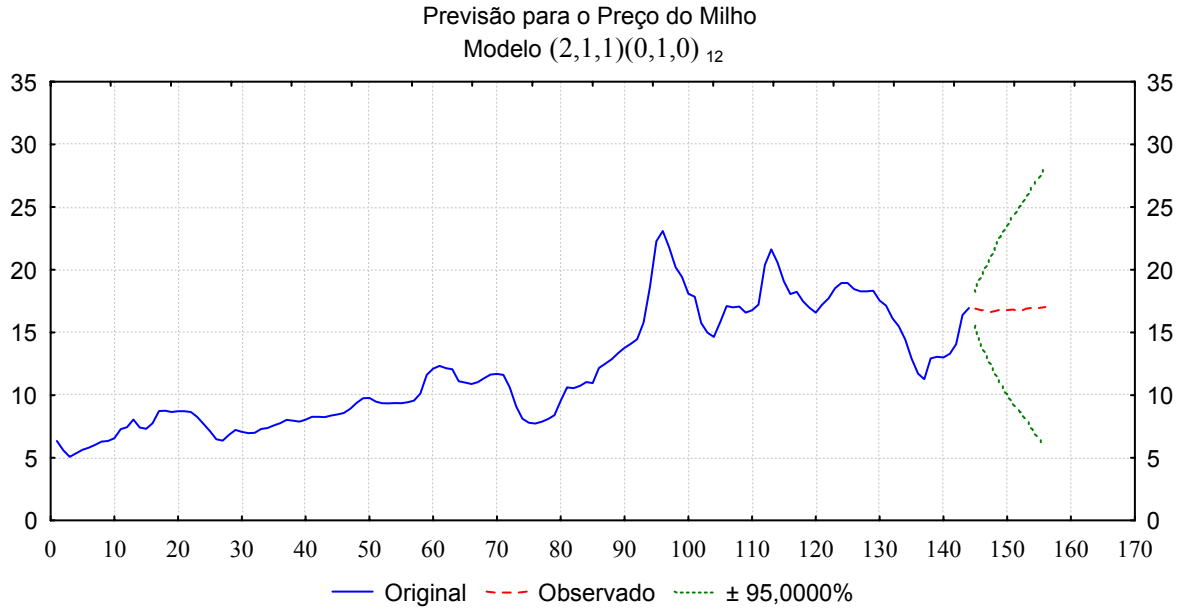


Figura 18 – Série Original e a Previsão para o ano de 2007.

A Figura 18 apresenta a Série Original com os dados previstos para o ano de 2007. Conclui-se que o modelo ajustado satisfez os critérios necessários para modelar uma Série através da Metodologia Box & Jenkins.

Através da Tabela 8 e da Figura 18, é possível visualizar as previsões mensais para o ano de 2007. Além disso, podem-se observar os valores reais até maio de 2007, o erro padrão e os limites com 95% de confiabilidade. Comparando os valores previstos com os reais, conclui-se que esta metodologia conseguiu captar o padrão encontrado na Série em estudo e fazer a previsão para o preço da soja no ano de 2007, através do modelo selecionado: $Z_t = 0{,}70394Z_{t-1} + 0{,}2065Z_{t-2} + a_t - 0{,}95807a_{t-1}$

5 CONCLUSÃO

Utilizando a Metodologia de Box & Jenkins, é possível analisar o comportamento do preço das principais Lavouras Temporárias do Rio Grande do Sul, sendo as culturas que apresentaram maior produção no período de 1995 a 2006 foram Soja, Arroz e Milho. A metodologia aplicada apresenta eficiência quando sua construção é baseada no Ciclo Iterativo, seguindo as fases de Identificação, Estimação, Verificação e Previsão. Dessas, a fase crítica do procedimento é a Identificação, porque se busca escolher Modelos Parcimoniosos (pequeno número de parâmetros e previsões precisas).

As análises realizadas permitem tecer alguns comentários a respeito das variáveis estudadas. É fundamental escolher-se bem o modelo para a Série, pois todo o processo de previsão depende dessa escolha. O melhor instrumento para tal escolha é a análise do gráfico da Série, que deve sempre antecipar-se a qualquer estudo.

Logo, através dessa metodologia, foi possível encontrar vários modelos significativos para cada Série e ter a opção de escolher o melhor modelo, por meio de critérios penalizadores, que buscam minimizar seus parâmetros. Foram encontrados os seguintes modelos para as Séries dos Preços dos Principais Grãos: Soja: *ARIMA (0,1,2)(0,1,0);* Arroz: *SARIMA(1,1,0)(1,0,1) e* Milho: *ARIMA (2,1,1)(0,1,0).*

Esses resultados são extremamente relevantes para a agricultura, pois ela está mostrando-se receptiva a tecnologias e a recursos que agregam valor ao produto, tornando a atividade cada vez mais competitiva. Este estudo permite aos agricultores analisarem a viabilidade da atividade com a obtenção de lucros satisfatórios, mesmo considerando que estes dependem de fatores controláveis e incontroláveis.

Com o conhecimento de tais previsões, os gaúchos, sejam eles da agricultura, de empresas ou do Estado, dispõem de informações para desenvolverem um planejamento, através das expectativas de lucros e benefícios esperados, agregando valor ao seu trabalho.

Deixa-se como sugestão serem desenvolvidos novos estudos envolvendo a aplicação de modelos ARFIMA, que são modelos com memória longa. São processos não-estacionários que possuem uma longa função de autocorrelação, com decaimento muito longo e cuja análise necessitará de uma diferença fracionária.

6 BIBLIOGRAFIA

6.1 Referências Bibliográficas

BARBANCHO, A. G. **Fundamentos e Possibilidades da Econometria**. Rio de Janeiro: Forum Editora, 1970. p. 18 - 32.

BARBIERO, C. C. M. **Séries Temporais: Um Estudo de Previsão para a Receita Operacional da ECT** – Empresa Brasileira de Correios e Telégrafos. 2003. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina. Disponível em: <http://www.qualimetria.ufsc.br/dissertacoes_arquivos/claudia.pdf>. Acesso em 15 mar. 2007.

BARROS, M. **Consultoria Ltda: Capítulo 5: Introdução às Séries Temporais e aos Modelos ARIMA.** Disponível em: <www.mbarros.com/sitebuildercontent/sitebuilderfiles/LivroProcessosCap5parte.pdf>. Acesso em 30 out. 2006.

EMATER/RS - **Associação Riograndense de Empreendimentos de Assistência Técnica e Extensão Rural.** Disponível em: <http://www.emater.tche.br>. Acesso em 02 mar. 2007.

EMBRAPA **- Empresa Brasileira de Pesquisa Agropecuária.** Disponível em: <http://www.embrapa.br>. Acesso em 02 mar. 2007.

FARIAS, E. R.; ROCHA, F. J. S; LIMA, R. C. **Critérios de seleção de modelos sazonais de séries temporais: uma aplicação usando a taxa de desemprego da região metropolitana de Recife**. III Encontro Regional de Estudos do Trabalho – ABET, 22 a 24 de novembro de 2000 – Recife, PE. Disponível em: <http://www.race.nuca.ie.ufrj.br/abet/3reg/39.DOC>. Acesso em: 20 nov. 2005.

GRANGER, C. W. J.; NEWBOLD, Paul. **Forecasting Economic Time Series**. New York: Academic Press, 1977.

IBGE – **Instituto Brasileiro de Geografia e Estatística.** Disponível em: <http://www.ibge.com.br>. Acesso em 05 out. 2006.

IRGA - **Instituto Rio Grandense do Arroz**. Disponível em: <http://www.irga.rs.gov.br>. Acesso em 02 mar. 2007.

MAKRIDAKIS, S; WHEELWRIGHT,S.; HYNDMAN,R J. **Forcasting methodos and applications.** 3 Ed. New York; John Wiley e Sons, 1998.

MORETTIN, P. A.; TOLOI, C. M. **Previsão de Séries Temporais**. 2ª edição, São Paulo: Atual, 1986.

MORETTIN, P. A.; TOLOI, C. M. **Previsão de Séries Temporais**. 2ª edição, São Paulo: Atual, 1987.

MORETTIN, P. A. & TOLOI, C. M. C., **Modelos de Função de Transferência**. Rio de Janeiro, **3ª Escola e séries Temporais Econometria,** 1989.

MORETTIN, P. A.; TOLOI, C. M. **Análise de Séries Temporais**. 1ª ed. São Paulo: E. Blücher, 2004.

MUELLER, A. **Uma Aplicação de Redes Neurais Artificiais na Previsão do Mercado Acionário**. Dissertação aprovada pelo Curso de Pós-graduação em Engenharia de Produção; 1996. Disponível em <http://www.eps.ufsc.br/disserta96/mueller/cap3>. Acesso em 29 mar. 2007.

MURTEIRA, B. J. F.; MÜLLER, D. A.; TURKMAN, K. F. **Análise de sucessões cronológicas**. Lisboa: McGraw-Hill, 1993.

NELSON, Charles R. **Applied Time Series Analysis**. San Francisco : Holden-Day, 1973.

SOUZA, R. C.; CAMARGO, M. E. **Análise e previsão de séries temporais:** Os modelos ARIMA. Ijuí, RS: Sedigraf, 1996.

SOUZA, Reinaldo Castro. **Modelos Estruturais para Previsão de Séries Temporais: Abordagens Clássica e Bayesiana**. In : 17° Colóquio Brasileiro de Matemática. Rio de Janeiro, 1989.

VASCONCELOS, M. A. S.; ALVES, D. **Manual de Econometria.** São Paulo: Atlas, 2000.

6.2 Bibliografias Consultadas

EHLERS, R.S. (2005) **Análise de Séries Temporais.** Departamento de Estatística,UFPR. Disponível em <http://www.est.ufpr.br/~ehlers>. Acesso em: 21 nov. 2006.

GRANGER, C. W. J. ***Forecasting in Business and Economics.*** 2nd Edition, Academic Press, 1989.

LATORRE, M.R.D.O. & CARDOSO, M.R.A; **Análise de Séries Temporais em epidemiologia: uma introdução sobre os aspectos metodológicos.** Revista Brasileira Epidemiologia Vol. 4, Nº 3, 2001. Disponível em: <http://www.scielo.br/pdf/rbepid/v4n3/02.pdf>. Acesso em 22 nov. 2006.

OLIVEIRA, M. A.; FAVERO, L. P.L. **Uma Breve Descrição de Algumas Técnicas Para Análise de Séries Temporais: Séries de Fourier, Wavelets, Arima, Modelos Estruturais para Séries de Tempo e Redes Neurais.** Apresentado no VI SEMEAD - SEMINÁRIOS EM ADMINISTRAÇÃO/2003. USP. Disponível em: <http://www.wad.fea.usp.br/semead/6semead>. Acesso em 15 mar. 2007**.**

SÁFADI, T. **Uso de Séries Temporais na Análise de Vazão de Água na Represa de Furnas**. Dra. em Estatística. Professora Adjunto IV. Departamento de Ciências Exatas – Universidade Federal de Lavras/UFLA. Disponível em: <http://www.editora.ufla.br/revista/28_1/art19.pdf >. Acesso em 15 mar. 2007.

SOUZA, F. M. **Modelos Box & Jenkins Aplicados a Demanda de Leitos Hospitalares** (Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, Santa Maria, 2006.

WERNER, L.; RIBEIRO, J. L. D. **Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais.** [S. l.]: Gestão & Produção, v. 10, n. 1, p. 47-67, abr. 2003.