

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

**ANÁLISE DAS SÉRIES DE DISTRIBUIÇÃO E
CONSUMO DE ÁGUA EM SANTA MARIA – RS
UTILIZANDO A METODOLOGIA DE BOX & JENKINS**

MONOGRAFIA DE ESPECIALIZAÇÃO

Juliana dos Santos Nunes

**Santa Maria, RS, Brasil
2007**

**ANÁLISE DAS SÉRIES DE DISTRIBUIÇÃO E CONSUMO DE
ÁGUA EM SANTA MARIA – RS UTILIZANDO A
METODOLOGIA DE BOX & JENKINS**

por

Juliana dos Santos Nunes

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. Ivanor Müller

Santa Maria, RS, Brasil
2007

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Programa de Pós-Graduação em Estatística e Modelagem
Quantitativa**

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

**ANÁLISE DAS SÉRIES DE DISTRIBUIÇÃO E CONSUMO DE ÁGUA
EM SANTA MARIA – RS UTILIZANDO A METODOLOGIA DE BOX &
JENKINS**

elaborada por

Juliana dos Santos Nunes

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

COMISSÃO EXAMINADORA:

**Ivanor Müller, Dr.
(Presidente/Orientador)**

Adriano Mendonça Souza, Dr. (UFSM)

Fernando de Jesus Moreira Junior, Ms. (UFSM)

Santa Maria, 17 de agosto de 2007.

AGRADECIMENTOS

À Universidade Federal de Santa Maria, pela oportunidade de novas visões para o crescimento pessoal, intelectual e profissional.

Ao Professor Doutor Ivanor Müller, meu orientador, pela orientação, segurança e confiança que soube transmitir e também pelo incentivo, sem o qual este trabalho não teria sido realizado.

Ao Programa de Pós-Graduação em Estatística e Modelagem Quantitativa, pela oportunidade de realização da Especialização.

Aos professores do Programa de Pós-Graduação em Estatística e Modelagem Quantitativa da Universidade Federal de Santa Maria, pela contribuição dada através de suas experiências e pelos conhecimentos transmitidos ao longo do curso.

Aos meus pais e a minha família, pelo incentivo e conforto que me proporcionaram.

Ao Robison, pela compreensão que teve durante a realização do curso, sempre me acompanhando em todas as horas.

Aos meus colegas, que, no decorrer do curso, estiveram sempre incentivando e proporcionando desafios para que aprimorássemos nossos conhecimentos.

À Companhia Riograndense de Saneamento, em nome da agente administrativa Andréia Zanini e do eletrotécnico Ivo Bertoldo, pelo fornecimento dos dados utilizados na elaboração de minha Monografia.

Enfim, a todas aquelas pessoas que direta ou indiretamente contribuíram para a realização desta Monografia, o meu carinho e agradecimento.

RESUMO

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

ANÁLISE DAS SÉRIES DE DISTRIBUIÇÃO E CONSUMO DE ÁGUA EM SANTA MARIA – RS UTILIZANDO A METODOLOGIA DE BOX & JENKINS

Autora: Juliana dos Santos Nunes

Orientador: Dr. Ivanor Müller

Data e Local de Defesa: Santa Maria, 17 de agosto de 2007.

Entre os recursos naturais, a água é o elemento mais importante para a subsistência das espécies. Com o crescimento da população, a demanda por água aumenta. Isso aliado ao desperdício e poluição dos mananciais, gera um problema de grandes proporções: por quanto tempo mais teremos água disponível na natureza. É preciso desenvolver tecnologias que permitam um melhor planejamento para a utilização dos recursos hídricos, investindo em políticas públicas de educação ambiental e conscientização para o cuidado com a água. Dentro desse contexto, o trabalho aqui apresenta uma análise das séries de Distribuição e Consumo de Água na cidade de Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2005, utilizando a Metodologia de Box & Jenkins, objetivando fazer previsões doze passos à frente. Os dados foram coletados pela Companhia Riograndense de Saneamento - CORSAN. A análise das séries mostrou que ambas possuem homogeneidade nos dados e apontou como melhor modelo para a série distribuição o $SARIMA(1,1,2)(1,0,0)_{12}$ e para a série consumo o $SARIMA(1,1,2)(0,0,1)_{12}$. Comparando os valores previstos com os reais, concluiu-se que as duas séries apresentaram variações. Isso se explica pelo fato de que, no ano de 2006, houve uma diminuição significativa no volume de água da Barragem do DNOS, uma das fontes de abastecimento da cidade. Por isso, a Corsan fez um planejamento diferenciado para esse ano, a fim de evitar racionamentos prolongados e transtornos para a população.

Palavras-chave: Séries Temporais; Metodologia Box & Jenkins; Água.

ABSTRACT

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

ANÁLISE DAS SÉRIES DE DISTRIBUIÇÃO E CONSUMO DE ÁGUA EM SANTA MARIA – RS UTILIZANDO A METODOLOGIA DE BOX & JENKINS

(WATER CONSUMPTION AND DISTRIBUTION SERIES ANALYSIS IN SANTA MARIA
– RS USING BOX & JENKINS' METHODOLOGY)

Author: Juliana dos Santos Nunes
Advisor: Dr. Ivanor Müller
Date and Place of Defense: Santa Maria, August 17th 2007.

Among the natural resources, water is the most important element for the species' survival. With the population growth, the demand for water increases. This, allied to the waste and the pollution of the water springs, generates a problem of great proportions: for how much time will water be available in nature. It is necessary to develop technologies that allow us a better planning for the use of water resources, investing in public policies of environmental education and awareness to the water issue. In this context, this work presents an analysis of the Distribution Series and Water Consumption in the city of Santa Maria – RS, in the period from January 2000 to December 2005, using Box and Jenkins Methodology, aiming to make forecasts twelve steps ahead. The data was collected by Companhia Riograndense de Saneamento – CORSAN. The series analysis showed that both have data homogeneity and pointed the SARIMA(1,1,2)(1,0,0)₁₂ as the best model for the distribution series and SARIMA(1,1,2)(0,0,1)₁₂ for the consumption series. Comparing the predicted values with the real ones, it was concluded that the two series presented variations. It is due to the fact that in the year of 2006, there was a significant decrease in the water volume of the Dam of DNOS, one of the city's of water supply. Considering this, CORSAN made a differentiated planning for this year in order to avoid extended water rationing, and disruptions to the population.

Key Words: Time Series; Box & Jenkins Methodology; Water.

LISTA DE FIGURAS

FIGURA 2.1 - Sistema dinâmico.....	22
FIGURA 2.2 - Fluxograma do ciclo iterativo de Box & Jenkins.....	32
FIGURA 4.1 - Série original distribuição de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2005.....	47
FIGURA 4.2 - Função de autocorrelação (FAC).....	48
FIGURA 4.3 - Função de autocorrelação parcial (FACP).....	48
FIGURA 4.4 - Série distribuição de água em Santa Maria original e dife- renciada.....	49
FIGURA 4.5 - Função de autocorrelação da série transformada.....	50
FIGURA 4.6 - Função de autocorrelação parcial da série transformada....	50
FIGURA 4.7 - Função de autocorrelação residual (FACr).....	52
FIGURA 4.8 - Função de autocorrelação parcial residual (FACPr).....	52
FIGURA 4.9 - Gráfico da série distribuição com os valores reais e valo- res previstos pelo modelo escolhido.....	53
FIGURA 4.10 - Série original consumo de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2005.....	54
FIGURA 4.11 - Função de autocorrelação (FAC).....	55
FIGURA 4.12 - Função de autocorrelação parcial (FACP).....	55
FIGURA 4.13 - Gráfico da série consumo de água em Santa Maria origi- nal e diferenciada.....	56
FIGURA 4.14 - Função de autocorrelação da série transformada.....	57
FIGURA 4.15 - Função de autocorrelação parcial da série transformada...	57
FIGURA 4.16 - Função de autocorrelação residual (FACr).....	59
FIGURA 4.17 - Função de autocorrelação parcial residual (FACPr).....	59
FIGURA 4.18 - Gráfico da série consumo com os valores reais e valores previstos pelo modelo escolhido.....	59

LISTA DE TABELAS

TABELA 4.1 - Estatística descritiva da série distribuição.....	48
TABELA 4.2 - Critérios de comparação e escolha do melhor modelo para a série distribuição.....	51
TABELA 4.3 - Valores reais e previstos da série distribuição para o ano de 2006, em m ³	53
TABELA 4.4 - Estatística descritiva da série consumo.....	54
TABELA 4.5 - Critérios de comparação e escolha do melhor modelo para a série consumo.....	58
TABELA 4.6 - Valores reais e previstos da série consumo para o ano de 2006, em m ³	60

LISTA DE ABREVIATURAS E SIGLAS

AIC – Akaike Information Criterion

ANA – Agência Nacional de Águas

AR – Auto-regressivo

ARMA – Auto-regressivo de Médias Móveis

CORSAN – Companhia Riograndense de Saneamento

DNOS – Departamento Nacional de Obras e Saneamento

EQM – Erro Quadrado Médio

ETA – Estação de Tratamento de Água

FAC – Função de Autocorrelação

FACP – Função de Autocorrelação Parcial

KW – Kruskal-Wallis

MA – Médias Móveis

ONU – Organização das Nações Unidas

PAC – Plano de Aceleração do Crescimento

SARIMA – Auto-regressivo de Médias Móveis Sazonal

SBC – Schwartz Bayesian Criterion

SQR – Soma do Quadrado dos Resíduos

UNESCO – Organização das Nações Unidas para a Educação, a Ciência e a Cultura.

LISTA DE APÊNDICES

APÊNDICE A - Dados originais da variável distribuição de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2006.....	67
APÊNDICE B - Dados originais da variável consumo de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2006.....	68

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Objetivos.....	13
1.2 Justificativa e importância do trabalho.....	13
1.3 Delimitação da pesquisa.....	14
1.4 Estrutura da Monografia.....	14
2 REVISÃO DE LITERATURA	16
2.1 Água e seus usos.....	16
2.2 Série temporal.....	20
2.3 Metodologia de Box & Jenkins.....	31
2.4 Comentários gerais do capítulo.....	43
3 MATERIAL E MÉTODOS	44
3.1 Metodologia.....	44
3.2 Caracterização da área de estudo.....	44
3.3 Comentários gerais do capítulo.....	46
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS	47
4.1 Série distribuição de água.....	47
4.2 Série consumo de água.....	54
4.3 Comentários gerais do capítulo.....	60
5 CONCLUSÕES E RECOMENDAÇÕES	61
5.1 Conclusões.....	61
5.2 Sugestões para trabalhos futuros.....	62
6 BIBLIOGRAFIA	64
APÊNDICES	66

1 INTRODUÇÃO

Entre os recursos naturais, a água é o elemento mais importante para a subsistência das espécies, pois dependem da disponibilidade da mesma para satisfazer suas necessidades. Quase todos os aspectos da vida do homem giram em torno da água, razão pela qual os povos se desenvolveram nas proximidades de fontes de água.

Para que possa ser consumida, sem apresentar riscos à saúde a água tem que ser tratada, limpa e descontaminada. A Companhia Riograndense de Saneamento – CORSAN - mantém captações em rios, lagos e barragens responsáveis por 80% do volume total produzido. Os 20% restantes - grande parte destinada a abastecer pequenas localidades - são buscados em mananciais subterrâneos. A água superficial é captada por meio de bombas e é conduzida, através das adutoras de água bruta, até as estações de tratamento de água, também chamadas ETAs.

Um serviço deficiente de abastecimento de água potável afeta a saúde das populações. Por isso, é importante contar com um sistema adequado de abastecimento. O sistema de água potável é um conjunto de estruturas, equipamentos e instrumentos destinados a produzir água de consumo humano a fim de entregá-la aos usuários em quantidade e qualidade adequadas, tendo um serviço contínuo a um custo razoável. Os sistemas de abastecimento de água geralmente contêm os seguintes componentes: obras de captação, estação de tratamento, redes de distribuição e conexões domiciliares.

É preciso desenvolver tecnologias que permitam um melhor planejamento para a utilização dos recursos hídricos, investindo em políticas públicas de educação ambiental e conscientização para o problema relacionado a este recurso tão escasso e necessário à vida do ser humano. Nesse sentido, a previsão da distribuição e do consumo de água numa determinada região é importante para apontar épocas que exigirão maior atenção dos responsáveis pelo serviço e, assim, orientar as ações de planejamento, visando evitar racionamentos e promover o uso consciente dos recursos hídricos.

1.1 Objetivos

Esta pesquisa será desenvolvida com a finalidade de fazer previsões para a Distribuição e o Consumo de Água em Santa Maria.

- Objetivo geral

Prever a distribuição e o consumo de água em Santa Maria – RS, como uma proposta que poderá ser implantada no sistema de gerenciamento da CORSAN a fim de orientar as ações de planejamento e pesquisa para uma melhor utilização dos recursos hídricos.

- Objetivos específicos

- Fazer as estatísticas descritivas das séries representativas da distribuição e do consumo de água em Santa Maria, no período de janeiro de 2000 à dezembro de 2005;
- Modelar as séries;
- Validar as previsões obtidas pelos modelos escolhidos.

1.2 Justificativa e importância do trabalho

A empresa pública responsável pelo abastecimento de água tratada no estado do Rio Grande do Sul - Companhia Riograndense de Saneamento (CORSAN) – vem buscando, nos últimos anos, implementar, em parceria com os Governos Federal e Estadual, projetos que tornem possível um melhor aproveitamento da água captada na natureza. Entre eles, um sistema de gerenciamento que controlará o destino da água distribuída pela empresa, identificando com maior rapidez e eficiência vazamentos e perdas na rede, possibilitando uma ação rápida que minimizará as perdas e aumentará a disponibilidade de recursos hídricos para um maior número de pessoas.

Este trabalho é importante como ferramenta desse sistema de gerenciamento, ajudando a eliminar racionamentos de água durante os meses de maior estiagem, pois através da previsão de séries temporais, é possível ter uma estimativa da quantidade de água que será necessária para a distribuição, e o quanto dessa água será consumida pela população de Santa Maria durante todo um ano, facilitando assim um planejamento mês a mês.

1.3 Delimitação da pesquisa

Este estudo delimita-se à previsão da distribuição e do consumo mensal de água tratada na cidade de Santa Maria, RS no período de janeiro à dezembro de 2006.

1.4 Estrutura da Monografia

Para atender aos objetivos propostos, a estrutura deste trabalho está organizada em seis capítulos.

No Capítulo 1, são apresentadas a introdução, os objetivos, a justificativa e importância do trabalho, a delimitação do tema e a estrutura do trabalho.

No Capítulo 2, apresenta-se o referencial teórico sobre o tema da pesquisa, ou seja, um estudo sobre a água e uma revisão de literatura sobre Séries Temporais e Metodologia de Box & Jenkins.

No Capítulo 3, apresentam-se os Materiais e Métodos com a caracterização da área de estudo.

No Capítulo 4 estão a identificação e apresentação dos dados coletados, bem como o desenvolvimento de sua análise e interpretação.

No Capítulo 5, apresentam-se as considerações finais a respeito da resposta ao problema de pesquisa proposto, e verifica-se o atendimento aos objetivos estabelecidos. Além disso, indicam-se novas oportunidades de pesquisa que poderão complementar e aprimorar o presente trabalho.

As Referências Bibliográficas encerram o presente trabalho, apresentando as obras bibliográficas que foram referenciadas e as que serviram de subsídio para a realização do mesmo.

2 REVISÃO DE LITERATURA

Este capítulo apresenta um estudo sobre a água e seus usos, a teoria sobre séries temporais e a metodologia de Box & Jenkins que servirá como suporte para o desenvolvimento do trabalho.

2.1 A água e seus usos

No ano do nascimento de Cristo, a população mundial estimada era de 200 milhões. Hoje passamos dos 6 bilhões de habitantes. Toda esta população necessita de água, de alimento e de espaço para sobreviver. Porém, a água não está disponível a todos de forma equilibrada. Em muitas regiões a população sofre com sua escassez, sendo ela motivo de conflito entre nações. Além disso, a água vem sendo contaminada indiscriminadamente, principalmente por conta de agrotóxicos, esgotos, resíduos industriais, lixo e pelos poluentes do ar que se misturam com a chuva.

Os países que atualmente não sofrem com a escassez têm hábitos de desperdício de água, especialmente nas residências.

A água é um elemento vital para os seres humanos e o ambiente. Em condições normais, perdemos em torno de 2,5 litros de água por dia, sendo 400 ml na respiração (durante a expiração); 1,2 litros ao urinar; 600 ml na transpiração e entre 100ml e 300 ml na evacuação. A hidratação regular é muito importante, pois quando a perda de água atinge um litro a sensação é de sede. Quando o prejuízo chega a 2 litros, além da sede, há cansaço e fadiga. Já quando temos 3 litros de água a menos, ocorre a desidratação e há riscos para a vida devido ao mau funcionamento dos rins.

Por ser um elemento essencial à vida, a água deve chegar à população com qualidade e quantidade adequadas para atender todas as suas necessidades. O homem, além de utilizar a água para o desempenho de suas atividades enquanto ser vivo, também a usa como elemento de desenvolvimento econômico, por exemplo, nos processos industriais; na irrigação das plantações; na geração de energia; na navegação e na manutenção dos ecossistemas. Portanto, a escassez de água além de pôr em risco a vida no planeta, também pode afetar diversas atividades econômicas,

entre elas a geração de energia elétrica. A crise de energia que afetou a população brasileira em 2001 foi o resultado mais claro da má gestão dos recursos hídricos do país.

De acordo com dados da UNESCO sobre recursos hídricos, o volume total de água no planeta é calculado em torno de 1,4 bilhões km³. Porém, a quase totalidade dessa água é salgada, apenas 2,7% é de água doce. Deste total de água doce, quase tudo está congelado, 0,7% encontra-se no subsolo e aproximadamente 0,007% está na forma de rios e lagos.

Devido à poluição e à degradação ambientais registradas ao longo dos anos, vários mananciais já não apresentam água apropriada para o consumo humano, portanto, está ficando cada vez mais difícil encontrar água com qualidade.

Segundo a ONU, o ser humano utiliza 54% da água doce acessível dos rios, lagos e aquíferos. Se esse consumo continuar aumentando no ritmo atual, dentro de 25 anos a humanidade absorverá 90% da água doce disponível no planeta, deixando apenas 10% para as outras espécies. Aproximadamente um bilhão de pessoas não tem acesso à água tratada. A falta desta causa a morte de quatro milhões de crianças por ano em todo o mundo, vítimas de doenças como a cólera.

Os principais fatores que estão causando essa escassez são os desflorestamentos, as queimadas, as práticas agrícolas inadequadas, o uso excessivo de água em irrigação, os grandes desperdícios domésticos e a poluição.

Poluição é a contaminação da água por substâncias que interferem na saúde das pessoas e animais, na qualidade de vida e no funcionamento dos ecossistemas. Alguns tipos de poluição têm causas naturais - erupções vulcânicas, por exemplo - mas a maioria é causada pelas atividades humanas, a chamada poluição antropogênica. À medida que a tecnologia foi se sofisticando, o risco de contaminação tornou-se maior.

As águas são poluídas, basicamente, por dois tipos de resíduos: os orgânicos e os inorgânicos. Os resíduos orgânicos normalmente têm origem animal ou vegetal, e provêm dos esgotos domésticos e de diversos processos industriais ou agropecuários. Os resíduos inorgânicos vêm de indústrias - principalmente das químicas e petroquímicas. Conforme sua composição e concentração, os poluentes hídricos têm a

capacidade de intoxicar e matar microorganismos, plantas e animais aquáticos, tornando a água imprópria para o consumo ou para o banho.

A ONU aponta duas sugestões básicas para diminuir a escassez de água: aumentar a sua disponibilidade e utilizá-la mais eficazmente.

Todo ser humano tem direito a consumir ou a usar a água para as suas necessidades individuais fundamentais. Esse consumo da água realiza-se diretamente através da sua captação dos cursos de água e lagos, ou pelo recebimento da água através dos serviços públicos ou privados de abastecimento. A existência do ser humano, por si só, já lhe garante o direito à água e ar. Negar água ao ser humano é negar-lhe o direito à vida. O direito à vida antecede os outros direitos.

Por tudo isso, é necessário que haja mudanças nos hábitos de consumo da população, e que políticas públicas eficientes para tratamento de esgoto e distribuição de água sejam implementadas.

Com a aprovação da Lei 9.433 (1997) que institui a Política Nacional de Recursos Hídricos, criou-se uma nova, importante e moderna estrutura para a gestão destes recursos, prevendo processos participativos e novos instrumentos econômicos que promovam o uso mais eficiente da água, como a cobrança pelo seu uso. O Governo Federal criou, em 2000, a Agência Nacional de Águas - ANA, responsável, entre outras coisas, por implementar a nova Lei.

Como exemplo de política pública voltada para a gestão de recursos hídricos, cito o programa Pró-Água Nacional do Governo Federal, criado para a região nordeste, que visa a atender ao abastecimento urbano de água e seus usos múltiplos em nível nacional. O programa inclui também questões de saneamento. Agora o programa foi ampliado para todo o Brasil. O Rio Grande do Sul é o estado modelo para o seu desenvolvimento, e os projetos estão inseridos no Plano de Aceleração do Crescimento (PAC).

Os principais problemas que enfrentamos são: a exploração de aquíferos e águas superficiais; o uso de tecnologias que demandam quantidades excessivas de água nas indústrias; o despejo de águas residuais provindas de indústrias sem tratamento prévio; as técnicas de irrigação que desperdiçam e contaminam a água; o

inadequado lançamento do esgoto doméstico, e a falta de esclarecimento dos usuários para fazer frente a esta situação.

Produzir água potável não é fácil. Requer investimento de grandes cifras para construir estações de tratamento e comprar os insumos necessários para purificá-la. A qualidade da água tratada depende do seu uso. É de vital importância para a saúde pública que a comunidade conte com um abastecimento seguro que satisfaça as necessidades domésticas tais como o consumo, a preparação de alimentos e a higiene pessoal. Para alcançar este propósito devem ser cumpridas uma série de normas de qualidade (física, química e microbiológica), de tal maneira que a água esteja livre de organismos capazes de originar enfermidades e de qualquer mineral ou substância orgânica que possa prejudicar a saúde.

Fases do processo de tratamento da água

- Floculação: é o processo onde a água recebe uma substância química chamada de sulfato de alumínio. Este produto faz com que as impurezas se aglutinem formando flocos para serem facilmente removidos.
- Decantação: como os flocos de sujeira são mais pesados do que a água caem e se depositam no fundo do decantador.
- Filtração: nessa fase, a água passa por várias camadas filtrantes onde ocorre a retenção dos flocos menores que não ficaram na decantação. A água então fica livre das impurezas.

Estas três etapas: floculação, decantação e filtração recebem o nome de clarificação. Nesta fase, todas as partículas de impurezas são removidas deixando a água límpida. Mas ainda não está pronta para ser usada. Para garantir a qualidade da água, após a clarificação é feita a desinfecção.

- Cloração: consiste na adição de cloro. Este produto é usado para destruição de microorganismos presentes na água.
- Fluoretação: é uma etapa adicional. O produto aplicado tem a função de colaborar para redução da incidência da cárie dentária.

Cada ETA possui um laboratório que processa análises e exames físico-químicos e bacteriológicos destinados à avaliação da qualidade da água desde o manancial até o sistema de distribuição. Além disso, existe um laboratório central que faz a aferição de todos os sistemas e também realiza exames especiais como: identificação de resíduos de pesticidas, metais pesados e plâncton. Esses exames são feitos na água bruta, durante o tratamento e em pontos da rede de distribuição, de acordo com o que estabelece a legislação em vigor.

2.2 Série temporal

Segundo Souza & Camargo (1996, p. 1), uma série temporal é um conjunto de observações de uma dada variável, ordenadas segundo o parâmetro tempo, geralmente em intervalos eqüidistantes.

Se Z_t representa o valor da variável aleatória Z no instante t , denotaremos a série temporal por Z_1, Z_2, \dots, Z_N onde N é o tamanho da série ou o número de observações seriais da variável.

Segundo os autores, as séries temporais podem ser classificadas em:

- a) **Discretas:** quando o conjunto de observações for finito ou infinito numerável, isto é, $N \equiv Z = \{1, 2, \dots, t\}$. Denota-se por Z_t .
- b) **Contínuas:** quando o conjunto for infinito não-numerável, isto é, $N = \{t : 0 < t < N\}$. Denota-se por $Z(t)$.
- c) **Determinísticas:** quando uma função matemática pode ser usada para estabelecer exatamente os valores futuros da série.
- d) **Estocásticas:** quando os valores futuros da série só podem ser estabelecidos em termos probabilísticos, pois o modelo compõe-se também de um fator aleatório.
- e) **Multivariadas (discretas ou contínuas):** quando a série temporal é representada por um vetor $\{Z(t), t \in N\}$ de ordem $r \times 1$.
- f) **Multidimensional:** quando se tem $\{Z(t), t \in N\}$ e t é um vetor de ordem $p \times 1$.

- Enfoques básicos

Souza & Camargo (1996, p. 2), destacam dois enfoques básicos aplicados na análise de séries temporais:

- Análise no domínio do tempo;
- Análise no domínio da frequência.

- Análise no domínio do tempo

É quando se considera a evolução temporal do processo. A análise mede a magnitude do processo do evento que ocorre em determinado instante de tempo. Ela é baseada em um modelo paramétrico. Nela são utilizadas as funções de autocovariância e autocorrelação.

- Análise no domínio da frequência

É uma análise da frequência com que certos eventos ocorrem. Esta representação é útil em dois casos: quando os componentes harmônicos da série, separados, têm uma significação física, ou quando os efeitos práticos do processo são mais convenientes se medidos separadamente por suas componentes harmônicas. A função associada é a densidade espectral.

- Objetivos da análise de séries temporais

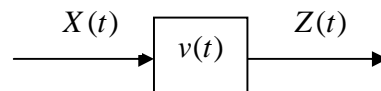
Para Morettin & Toloí (2004, p. 3-5), quando obtida a série temporal $Z(t_1), \dots, Z(t_N)$, observada nos instantes t_1, \dots, t_N , podemos estar interessados em:

- a) **investigar** o mecanismo gerador da série temporal; por exemplo, ao analisar uma série de alturas de ondas, podemos querer saber como estas ondas foram geradas;
- b) **fazer** previsões de valores futuros da série; estas podem ser a curto prazo, como para séries de vendas, produção ou estoque, ou a longo prazo, como para séries populacionais, de produtividade, etc.;

- c) **descrever** apenas o comportamento da série; neste caso, a construção do gráfico, a verificação da existência de tendências, de ciclos e variações sazonais, a construção de histogramas e diagramas de dispersão, etc., podem ser ferramentas úteis;
- d) **procurar** periodicidades relevantes nos dados; neste caso, a análise espectral, mencionada anteriormente, pode ser de grande utilidade.

Os autores ressaltam que, em todos os casos, *modelos probabilísticos* ou *estocásticos* são construídos no domínio temporal ou de frequências. Estes modelos devem ser simples e parcimoniosos (isto é, o número de parâmetros envolvidos deve ser mínimo) e quando possível, sua utilização não deve apresentar dificuldades às pessoas interessadas em manipulá-los.

Muitas situações em ciências físicas, engenharia, ciências biológicas e sociais envolvem o conceito de *sistema dinâmico*, caracterizado por uma série de entrada $X(t)$, uma série de saída $Z(t)$ e uma *função de transferência* $v(t)$ (Figura 2.1).



Fonte: Morettin & Toloí (1986, p. 5).

Fig. 2.1 – Sistema dinâmico.

Ainda, segundo os autores, de particular importância são os *sistemas lineares*, nos quais a saída é relacionada com a entrada através de um funcional linear envolvendo $v(t)$.

Problemas de interesse aqui são:

- a) **estimar** a *função de transferência* $v(t)$, conhecendo-se as séries de entrada e saída;
- b) **fazer** previsões da série $Z(t)$, com o conhecimento de observações da série de entrada $X(t)$;
- c) **estudar** o comportamento do sistema, simulando-se a série de entrada;
- d) **controlar** a série de saída $Z(t)$, de modo a trazê-la o mais próximo possível de um valor desejado, ajustando-se convenientemente a série de entrada $X(t)$; este controle é necessário devido às perturbações que normalmente afetam um sistema dinâmico.

- Análise dos componentes básicos de uma série temporal

Para Werner & Ribeiro (2003), séries temporais são compostas por quatro elementos:

- 1) **Tendência:** verifica o sentido de deslocamento da série ao longo de vários anos.
- 2) **Ciclo:** movimento ondulatório que ao longo de vários anos tende a ser periódico.
- 3) **Sazonalidade:** movimento ondulatório de curta duração, em geral, inferior a um ano; associada, na maioria dos casos, às mudanças climáticas.
- 4) **Ruído aleatório ou erro:** compreende a variabilidade intrínseca aos dados e não pode ser modelado.

Conforme Murteira et al. (1993, p. 6), a tendência ou *trend* tem significado intuitivo e pode ser descrita como inércia da sucessão, marcha principal, variação “em média” ao longo do tempo, ou ainda, mudança de nível. Compreende os movimentos que se manifestam suave e consistentemente durante períodos longos.

A componente estacional ou sazonal descreve as variações em relação à tendência que ocorrem, em geral, no decorrer de um ano; numa sucessão trimestral, por exemplo, os valores do primeiro podem ser sistematicamente mais elevados do que os valores do segundo, etc.

Segundo os autores, os movimentos sazonais são oscilações que se repetem todos os anos (ou com uma periodicidade mais curta), nem sempre seguindo um padrão rígido; podem ter causas naturais ou causas sociais. As primeiras estão quase sempre associadas às estações do ano, cujo efeito é intenso em atividades tais como: agricultura, turismo, consumo de energia, etc. As segundas estão associadas a usos, costumes e tradições sociais, disposições fiscais, etc.

Os movimentos oscilatórios ou ciclos, nomeadamente os que se verificam nas séries de produções, preços, etc., associam-se às fases alternadas de expansão e depressão que afetam o sistema econômico e, embora recorrentes, não apresentam periodicidade definida. Os ciclos longos são dificilmente separáveis da tendência.

Ainda segundo os autores, a componente aleatória ou ruído, onde os métodos menos sofisticados colocam tudo o que não conseguem definir ou modelar, tem papel extremamente relevante nos modernos modelos probabilísticos.

- Teste da Seqüência (Wald-Wolfowitz)

Para Morettin & Tolo (1986), a construção do gráfico de uma série temporal pode revelar características importantes como: tendência, sazonalidade, variabilidade, observações aberrantes e outras características. Além da inspeção gráfica, podem-se utilizar testes de hipóteses estatísticos para verificar se realmente existe tendência na série.

Estes testes podem ser realizados antes da estimação da tendência ou depois da estimação da tendência,

As hipóteses a serem testadas são:

H_0 : não existe tendência na série.

H_1 : existe tendência na série.

Os testes para a tendência só podem ser efetuados se a série estiver livre de sazonalidade e ciclos.

O teste da seqüência, que será abordado a seguir, é para ser efetuado antes da estimação da tendência.

Procedimento do teste:

- Calcula-se a mediana m e atribuem-se as letras:

$$\begin{cases} A, \text{ se } Z_t \geq m \\ B, \text{ se } Z_t < m \end{cases}$$

Sejam, $N = (n_1 \text{ pontos } A) + (n_2 \text{ pontos } B)$ e $T_1 =$ número total de seqüências
(troca de seqüências de letras)

- Rejeita-se H_0 se existem poucas seqüências, isto é, se T_1 é pequeno.

Para se tomar uma decisão correta em relação à rejeição de H_0 , deve-se consultar o valor tabelado para esta estatística que diz que: deve-se rejeitar H_0 se $T_1 < W_\alpha$, onde W_α é o α quantil da distribuição T_1 . Para n_1 e n_2 maior que 20, pode-se utilizar $T_1 \sim N(\mu; \sigma^2)$.

$$\mu = \frac{2 \cdot n_1 \cdot n_2}{N} + 1$$

Onde:
$$\sigma^2 = \frac{2 \cdot n_1 \cdot n_2 \cdot (2 \cdot n_1 \cdot n_2 - N)}{N^2 (N - 1)}$$

$$Z_{calc} = \frac{T_1 - \mu}{\sigma}$$

- Teste de Kruskal-Wallis

Pode-se testar também, a existência da componente sazonal, antes e depois de sua estimação. Em ambos os casos testam-se as hipóteses:

H_0 : não existe sazonalidade determinística.

H_1 : existe sazonalidade determinística.

Para maiores detalhes do teste e o procedimento para a sua utilização, ver Morettin & Tolo (1986, p. 46-47).

- Modelos para séries temporais

Souza & Camargo (1996, p. 26-27) destacam que por um modelo matemático entende-se uma construção abstrata de um fenômeno real, que deveria comportar-se em concordância com a situação real em todos os aspectos relevantes. No modelo perdem-se detalhes para reduzir ao essencial.

O objetivo de um modelo é clarificar a estrutura e a dinâmica do fenômeno e ajudar a fazer previsões e controlar o sistema.

Diz-se que um modelo concorda com fenômeno natural quando suas previsões concordam com a realidade.

A construção de um modelo se faz em três etapas:

- a) **Indução**: ou seja, construção do modelo a partir das observações;
- b) **Dedução**: o estudo das implicações do modelo;
- c) **Verificação**: experimentação do modelo e testes das previsões.

Conforme Morettin & Tolo (1987, p. 16), os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas.

- Processos estocásticos

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

Nestas 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 $Z = \{0, \pm 1, \pm 2, \dots\}$ ou o conjunto dos reais \mathfrak{R} . Também para cada $t \in T$, $Z(t)$ será uma variável aleatória real.

Para Souza & Camargo (1996, p. 16-18) um processo estocástico está estatisticamente determinado quando se conhece suas funções de distribuição até a N -ésima ordem. Na prática, ocorrem duas situações problemáticas: não se conhece todas as funções de distribuição até a N -ésima ordem e, comumente tem-se apenas uma realização do processo estocástico em questão, a partir da qual se deseja inferir todas as características do mecanismo gerador da série. Para superar essas dificuldades assumem-se duas restrições: **Estacionariedade e Ergodicidade**.

- Estacionariedade

Se o processo estocástico que gerou a série de observações é invariante em relação ao tempo, diz-se que o mesmo é estacionário. Se as características do processo se alteram no decorrer do tempo, diz-se que é não estacionário. Um processo estacionário pode ser classificado em:

a) **Estritamente estacionário:** quando suas características não são afetadas por variações devido à escolha da origem dos tempos, ou seja, quando as séries Z_t e Z_{t+k} estão distribuídas identicamente qualquer que seja “ k ”.

b) **Estritamente estacionário de ordem finita:** diz-se que um processo é estritamente estacionário de ordem “i” se a estacionariedade item (a) não é válida para todo $t_j \in N$, mas somente para $j \leq i$.

c) **Estacionário em sentido lato (ou fracamente estacionário ou, ainda, estacionário de segunda ordem):** quando a sua função valor médio é constante e sua função de covariância depende somente da diferença em valor absoluto, $t_s - t_j$. Observa-se que o processo estacionário em sentido lato requer condições envolvendo somente os momentos de primeira e segunda ordem, ao invés de sua função finito-dimensional. Daí porque é chamado por muitos autores de estacionário de segunda ordem.

Assim, um processo estacionário em sentido lato pode não ser estritamente estacionário, pois os momentos até segunda ordem não garantem condições sobre a estacionariedade da função de probabilidade. Porém, se o processo é estritamente estacionário de ordem dois, então ele é estacionário em sentido lato.

Conforme Morettin & Tolo (1987, p. 6), uma das suposições mais freqüentes que se faz a respeito de uma série temporal é a de que ela é estacionária, ou seja, ela se desenvolve 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*, sendo o caso mais simples aquele em que a série flutua ao redor de uma reta, com inclinação positiva ou negativa (tendência linear).

Podemos ter, também, uma forma de não estacionariedade explosiva, como o crescimento de uma colônia de bactérias.

Para Souza & Camargo (1996, p. 19) a importância do conhecimento da série ser ou não estacionária reside no fato de que, quando se trabalha com uma série estacionária, está-se em presença de uma função amostral do processo que tem a mesma forma em todos os instantes do tempo $t \in N$, o que acarreta possibilidades de obtenção de estimativa das características do processo de forma bastante simples, o que, em caso contrário, não seria tarefa fácil.

Morettin & Toloí (2004, p. 4-5) destacam que a classe dos modelos ARIMA será capaz de descrever de maneira satisfatória séries estacionárias e séries não estacionárias, desde 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, 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.

Os autores lembram que, como a maioria dos procedimentos de análise estatística de séries temporais supõe que estas sejam estacionárias, será necessário transformar os dados originais, se estes não formam uma série estacionária. 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

$$\Delta Z_t = Z_t - Z_{t-1}, \quad (2.1)$$

a segunda diferença é

$$\Delta^2 Z_t = \Delta[\Delta Z_t] = \Delta[Z_t - Z_{t-1}], \quad (2.2)$$

ou seja,

$$\Delta^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2}. \quad (2.3)$$

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

$$\Delta^n Z_t = \Delta[\Delta^{n-1} Z_t] \quad (2.4)$$

Segundo os autores, em situações normais, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária.

- Ergodicidade

Para Souza & Camargo (1996, p. 22-24), um processo estocástico é dito ergódico se apenas uma realização do mesmo é o suficiente para se obter todas as estatísticas do mesmo. Esse resultado é conhecido como teorema ergódico.

De maneira similar ao que definimos para os processos estacionários, podemos também ter processos ergódicos **estritos** e **amplios**.

Finalmente, segundo os autores, pelas definições apresentadas, pode-se facilmente verificar que todo o processo ergódico também é estacionário, visto que uma realização de um processo não estacionário não poderá conter todas as informações necessárias para a especificação do processo.

Portanto, algumas das propriedades dos processos estacionários que também se aplicam aos processos ergódicos (ambos no sentido amplo) seriam:

- i) média e variância constantes;
- ii) funções de autocovariância e autocorrelação independentes da origem dos tempos.

- Ruído branco

Seja uma seqüência $\{a_t, t \in N\}$ de variáveis aleatórias independentes e identicamente distribuídas, com distribuição normal de média zero e variância constante σ_a^2 . Tais variáveis são chamadas de choques aleatórios e a seqüência é chamada de ruído branco. Assim, o processo $a_t, t = 0, 1, 2, \dots$ é um ruído branco se:

$a_t \sim N[0, \sigma_a^2]$, com:

Média: $E[a_t] = 0$.

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

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

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

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

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

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

Densidade espectral: $f(\omega) = \frac{1}{2\pi}$

Densidade espectral (não normalizada): $h(\omega) = \frac{\sigma_a^2}{2\pi}$.

Para os autores, o termo ruído branco advém de aplicações em engenharia; é uma analogia com “luz branca”, na qual todas as frequências (cores) estão presentes em quantidades iguais.

Para Farias et al. (2000), a análise dos resíduos de modelos alternativos (concorrentes) ajustados é de extrema importância na escolha final do(s) modelo(s) que melhor explica(m) 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(s) com esta característica. Uma análise da existência (ou não) da autocorrelação serial de resíduos é feita com base nas Funções de Autocorrelação dos Resíduos (FACr) e Autocorrelação Parcial dos Resíduos (FACPr) e seus respectivos correlogramas.

Box & Pierce (1970) apud Morettin & Tolo (2004) sugeriram um teste para as autocorrelações dos resíduos estimados, que apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muitos altos. Uma variação desse teste foi sugerida por Ljung & Box (1978).

Se o modelo for apropriado, a estatística

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \quad (2.5)$$

terá aproximadamente uma distribuição χ^2 com $K - p - q$ graus de liberdade. 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 .

- Operadores

Souza & Camargo (1996, p. 25-26) apresentam os seguintes operadores:

a) **Operador de retardo ou de translação para o passado:** representa uma defasagem de k períodos de tempo para trás. É denotado por B , e definido por:

$$B^k Z_t = Z_{t-k} \quad (2.6)$$

b) **Operador de avanço ou de translação para o futuro:** representa uma defasagem de k períodos de tempo para frente. É denotado por F , e definido por:

$$F^k Z_t = Z_{t+k} \quad (2.7)$$

c) **Operador diferença:** é uma transformação nos dados que consiste em tomar diferenças sucessivas da série original. É denotado por ∇ , e definido por:

$$\nabla_s^d Z_t = (1 - B^s)^d Z_t \quad (2.8)$$

Para $s = 1$ e $d = 1$, temos a primeira diferença

$$\nabla Z_t = (1 - B)Z_t = Z_t - Z_{t-1} = w_t \quad (2.9)$$

$$\nabla^2 Z_t = (1 - B)^2 Z_t = w_t - w_{t-1} = Z_t - 2Z_{t-1} + Z_{t-2} \quad (2.10)$$

então, $\nabla = (1 - B)$.

d) **Operador soma ou inverso:** é denotado por S , e definido por:

$$S_s^d = (\nabla_s^d)^{-1} = \nabla^{-1} \quad (2.11)$$

Para $s = 1$ e $d = 1$, temos:

$$S_1^1 Z_t = \nabla^{-1} Z_t = (1 - B)^{-1} Z_t = (1 + B + B^2 + \dots) Z_t = \sum_{j=0}^{\infty} Z_{t-j} \quad (2.12)$$

Assim,

$$S_1^1 Z_t = \sum_{j=0}^{\infty} Z_{t-j} \quad (2.13)$$

2.3 Metodologia de Box & Jenkins

Segundo Souza & Camargo (1996, p. 27-28), uma das características fundamentais da metodologia Box & Jenkins é interpretar uma dada série temporal como sendo uma realização de um vetor aleatório multivariado, cuja dimensão é a da série temporal disponível. A partir de uma única realização do processo e, com os argumentos de estacionariedade e ergodicidade do processo subjacente, procura-se detectar o sistema gerador da série, através das informações contidas na série.

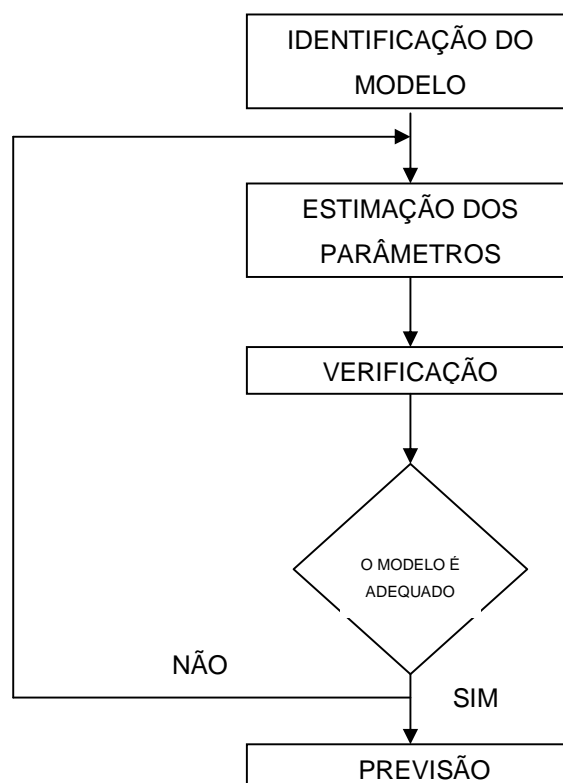
A filosofia da metodologia de Box & Jenkins utiliza-se de duas idéias: o princípio da parcimônia e a construção de um ciclo iterativo.

Segundo os autores, o princípio da parcimônia estabelece que devemos escolher um modelo com o menor número de parâmetros possíveis para uma representação matemática adequada. O ciclo iterativo é uma estratégia de seleção de modelos a ser empreendida até que tenhamos um modelo satisfatório.

Para Müller (1991, p. 26-27), as etapas do ciclo iterativo são:

- a) **identificação** da estrutura do modelo, feita através da análise da função de autocorrelação, função de autocorrelação parcial e outros critérios;
- b) **estimação**, feita através da maximização da função de verossimilhança condicional, ou através do método de mínimos quadrados;
- c) **verificação** do modelo ajustado, através dos resíduos a_t . Se este for ruído branco o modelo é adequado, caso contrário, repete-se o ciclo iterativo;
- d) **previsão** dos valores futuros utilizando o modelo adequado.

Na Figura 2.2, vemos o fluxograma do ciclo iterativo de Box & Jenkins.



Fonte: Müller (1991, p. 27)

Figura 2.2. Fluxograma do ciclo iterativo de Box & Jenkins.

- Modelos estacionários

Segundo Morettin & Tolo (2004, p. 111), este modelo supõe que a série temporal seja gerada através de um filtro linear (ou sistema linear), cuja entrada é ruído branco.

- Modelo auto-regressivo (AR)

Para Vasconcelos & Alves (2000, p. 205-206), de acordo com esse modelo, Z_t é descrito apenas por seus valores passados e pelo ruído branco a_t . A versão mais simples de um modelo AR é aquela em que Z_t depende somente de Z_{t-1} e de a_t . Diz-se, nesse caso, que o modelo é auto-regressivo de ordem 1, o que se indica compactamente por AR(1). A representação algébrica desse modelo é a seguinte:

$$Z_t = \phi_1 Z_{t-1} + a_t \quad (2.14)$$

onde ϕ_1 é um parâmetro e $E(a_t) = 0$; e $E(a_t^2) = s^2$; $E(a_t a_s) = 0$ para $t \neq s$.

Segundo Werner & Ribeiro (2003, p. 49), para o modelo ser estacionário é necessário que $|\phi_1| < 1$ (condição de estacionariedade) e que as autocovariâncias γ_k sejam independentes. No caso do modelo AR(1), as autocovariâncias são dadas por:

$$\gamma_k = \phi_1^k \gamma_0 \quad (2.15)$$

e as autocorrelações são dadas pela equação 2.16:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k \quad k = 0, 1, 2, \dots \quad (2.16)$$

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 (MA)

Segundo os autores, em um modelo de médias móveis (do inglês *moving average*), a série Z_t resulta da combinação linear dos ruídos brancos a_t , do período atual com aqueles ocorridos em períodos anteriores.

Morettin & Tolo (2004, p. 123) apresentam o modelo de médias móveis de primeira ordem – MA(1) – dado por:

$$Z_t = a_t - \theta_1 a_{t-1} \quad (2.17)$$

onde θ_1 é um parâmetro.

Werner & Ribeiro (2003, p. 50) apresentam as autocorrelações ρ_k , que nada mais são do que as autocovariâncias divididas pela variância, dadas por:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 \sigma_\varepsilon^2}{(1 + \theta_1^2) \sigma_\varepsilon^2} = \frac{-\theta_1}{(1 + \theta_1^2)} \quad \text{e} \quad (2.18)$$

$$\rho_k = 0 \quad k > 1 \quad (2.19)$$

A função de autocorrelação do modelo MA(1) apresenta apenas a primeira autocorrelação não nula, e as demais iguais à zero. A primeira autocorrelação será positiva se θ_1 for menor que zero, e negativa se θ_1 for maior que zero.

Para obter a condição de invertibilidade, isto é, para transformar um modelo MA(1) em um modelo AR(∞), é preciso impor a restrição de que $|\theta_1| < 1$.

- Modelo auto-regressivo de médias móveis (ARMA)

Morettin & Tolo (2004, p. 128) ressaltam que para muitas séries encontradas na prática, se quisermos um modelo com um número não muito grande de parâmetros, a inclusão de termos auto-regressivos e de médias móveis é a solução adequada.

O modelo ARMA(p,q) mais simples, segundo Vasconcelos & Alves (2000, p. 209), é o ARMA(1,1), dado por:

$$Z_t = \phi_1 Z_{t-1} + a_t - \theta_1 a_{t-1} \quad (2.20)$$

A função de autocorrelação do modelo ARMA (1,1), segundo Werner & Ribeiro (2003, p. 50), é dada por:

$$\rho_1 = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 + 2\phi_1 \theta_1} \quad (2.21)$$

$$\rho_k = \phi_1 \rho_{k-1} \quad \text{para } k > 1 \quad (2.22)$$

A função de autocorrelação do modelo ARMA(p,q) apresenta características da função MA(q) para as defasagens $k < q$, pelo fato de a “memória” do componente de médias móveis durar apenas q períodos. Para defasagens maiores que $k+1$ as características são iguais às de um modelo AR(p).

- Modelos não-estacionários

Segundo Vasconcelos & Alves (2000, p. 210), a metodologia de Box & Jenkins aplica-se a um caso específico de séries não estacionárias, ou seja, séries que se tornam estacionárias após a aplicação de diferenças. O número de diferenças necessário para tornar uma série estacionária é denominado ordem de integração.

Uma série não estacionária apresenta tendência estocástica, em contraposição à tendência determinística, que é expressa como função do tempo.

Werner & Ribeiro (2003, p. 51) ressaltam que, para detectar a não estacionariedade de uma série, o comportamento temporal pode ser analisado graficamente, ou então, através da aplicação de testes estatísticos de raiz unitária. O teste de raiz unitária mais usado é o Dickey-Fuller. Outros detalhes sobre esse teste podem ser vistos em Makridakis et al. (1998), Enders (1995) e Hamilton (1994).

Ainda, segundo os autores, o modelo aplicado a séries não estacionárias pode ser genericamente formulado da seguinte maneira: se Z_t tornar-se estacionária após a aplicação de d diferenças e a série resultante for representada por um modelo ARMA(p,q), diz-se que Z_t é descrita por um modelo ARIMA(p,d,q) representado por:

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.23)$$

onde $w_t = \Delta^d Z_t$.

Utilizando o operador de defasagem tem-se:

$$(1 - B)^d \phi(B) \gamma_t = \theta(B) a_t \quad (2.24)$$

- Modelos sazonais

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 que 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 $(P,D,Q)_s$. O modelo SARIMA $(p,q)(P,Q)_s$ é dado por:

$$\begin{aligned} & (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_{PS} B^{Ps})Z_t \\ & = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_{QS} B^{Qs})a_t \end{aligned} \quad (2.25)$$

Para Farias et al. (2000), se a série em estudo for não estacionária com sazonalidade, temos um SARIMA $(p,d,q)(P,D,Q)_s$, onde d representa a ordem de diferenciação não sazonal e D a ordem de diferenciação sazonal. A diferenciação não sazonal já foi tratada anteriormente. A diferenciação sazonal pode ser realizada da seguinte forma:

$$\Delta^D Z = Z_t - Z_{t-s} \quad (2.26)$$

- Diagnóstico

Segundo Santos (2006, p. 27), após a identificação e estimação do modelo, o último passo é diagnosticar sua adequação. Geralmente essa aplicação é feita examinando-se a série de resíduos gerada pelo modelo, buscando alguma estrutura de autocorrelação. Caso seja encontrada, significa que o modelo não deteve toda a informação contida na série original. O que se deve fazer primeiro é examinar a FAC e FACP estimadas da série de resíduos, como se esta fosse uma série temporal qualquer. Se houver estrutura identificável, o modelo é rejeitado.

Segundo Farias et al. (2000), existem vários critérios de seleção de modelos que permitem um “trade-off” entre uma redução na soma do quadrado dos resíduos estimados e um modelo mais parcimonioso. Os critérios mais usados são o Critério

“Akaike Information Criterion” (AIC) e o “Schwartz Bayesian Criterion” (SBC), cujas fórmulas são dadas por:

$$AIC = T \ln(SQR) + 2n \quad (2.27)$$

$$SBC = T \ln(SQR) + n \ln(T) \quad (2.28)$$

Onde,

SQR = soma do quadrado dos resíduos;

n = número de parâmetros estimados;

T = número de observações utilizadas.

Geralmente, quando se trabalha com variáveis defasadas perdem-se informações sobre a série temporal em estudo. Logo, para se comparar modelos alternativos (ou concorrentes) deve-se manter fixo o número de informações utilizadas, T, para todos os modelos em comparação.

Para os autores, a situação ideal é: quanto menor o AIC e o SBC melhor o ajustamento do modelo. Porém, faz-se necessário comparar os AIC's e os SBC's de modelos alternativos para se saber qual o modelo que melhor explica a dinâmica da série temporal em estudo. Porém, como a metodologia de Box & Jenkins leva em consideração outros aspectos do modelo estimado, dando grande ênfase na parcimônia, deve-se utilizar os critérios AIC e SBC juntamente com outros aspectos (características do modelo estimado) importantes na hora de escolher o melhor modelo.

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 $T \ln(SQR) > n \ln(T)$.

- Teste da normalidade dos resíduos

Segundo Jacobi (200?), o teste para uma amostra foi introduzido por Kolmogorov em 1933. Em 1939, Smirnov introduziu o procedimento do teste para dados de duas amostras.

É um teste de aderência que determina se os valores da amostra podem, razoavelmente, ser considerados como provenientes de uma população com aquela distribuição teórica e testa a bondade do ajuste para dados contínuos.

Em síntese, procura-se especificar a distribuição de frequência acumulada teórica e compará-la com a distribuição de frequência acumulada observada. Determina-se o ponto no qual as duas distribuições, teórica e observada, apresentam maior diferença e se essa diferença pode ser atribuída ao acaso.

O teste de Kolmogorov-Smirnov admite uma função de distribuição específica, com média e variância conhecidas.

Para testar normalidade, Lilliefors em 1967, introduziu uma modificação no teste de Kolmogorov-Smirnov, ampliando o seu uso aos casos em que a média e a variância não são previamente conhecidas, mas sim, estimadas através dos dados da amostra, conforme as seguintes equações, respectivamente:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.29)$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1} \quad (2.30)$$

Estrutura-se o teste analogamente ao de Kolmogorov-Smirnov (ver Siegel, p. 52), a partir dos Z_i , ao invés da variável original.

$$Z_i = \frac{(x_i - \bar{X})}{S}, \quad i = 1, 2, \dots, n \quad (2.31)$$

Onde Z_i é a variável padronizada.

Hipóteses:

H_0 : as observações seguem uma distribuição normal.

H_1 : as observações não seguem uma distribuição normal.

Regra de decisão:

$p > \alpha$, aceita-se H_0 .

$p \leq \alpha$, rejeita-se H_0 .

- Previsão

A previsão é uma das principais razões da popularidade da Metodologia de Box & Jenkins. Em muitos casos as previsões, principalmente de curto prazo, obtidas com base em Box & Jenkins são melhores que as obtidas com base nos modelos econométricos tradicionais. As previsões podem ser de dois tipos: “ex-ante” e “ex-post”. A previsão “ex-ante” é feita para calcular valores futuros, de curto prazo, da variável em estudo. Por outro lado, a previsão “ex-post” é feita para gerar valores dentro do período amostral. Logo, este tipo de previsão pode ser utilizado como um dos critérios de escolha entre modelos alternativos. Pois, quanto melhor forem essas previsões, melhor será o modelo estimado. O Erro Quadrado Médio da previsão (EQM), que é igual à média do quadrado da diferença entre cada valor previsto “ex-post” e o valor real observado na amostra, é uma medida formal da qualidade das previsões “ex-post”. Pois, quanto menor o EQM, melhor será o grau de ajustamento do modelo aos dados da série temporal em estudo.

Segundo Souza & Camargo (1996, p. 173-180), a previsão torna-se o cálculo do valor esperado de uma futura observação. Esse cálculo é condicionado aos valores passados e ao valor presente da variável. Ou seja, chamando $\hat{Z}_t(l)$ o valor previsto para um horizonte de “ l ” períodos de tempo futuro e “ t ” o período de origem da previsão, então,

$$\hat{Z}_t(l) = E(Z_{t+l} | Z_t, Z_{t-1}, \dots),$$

e o valor de Z no tempo $t+l$ é obtido pela equação do modelo ARIMA.

- Formas explícitas para o modelo de previsão

Considerando o modelo escrito em sua forma geral, ou seja,

$$\phi(B) \underbrace{\nabla^d Z_t}_{W_t} = \theta(B) a_t \quad (2.32)$$

Pode-se assim obter três formas possíveis para escrever as equações do modelo 2.32, nas quais se tenta explicar a série original Z_t . Para uma melhor compreensão das

equações que representamos a seguir, vamos considerar em 2.32 o polinômio $\phi'(B)$, onde:

$$\phi'(B) = \phi(B)\nabla^d,$$

que fornece para 2.32:

$$(1 - \xi_1 B - \xi_2 B^2 - \dots - \xi_{p+d} B^{p+d})Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)a_t \quad (2.33)$$

Os autores apresentam a seguir as três formas de explicitação de Z_t , todas obtidas através da formulação do modelo escrito pela equação 2.33.

- Forma de equações de diferenças

$$Z_t = f(Z_{t-j}; a_{t-i}); \quad j = 1, 2, \dots, p + q; \quad i = 0, 1, \dots, p$$

“O valor atual da série (Z_t) é a função de valores anteriores (Z_{t-j}) e valores atuais e anteriores dos resíduos (a_t)”.

Esta forma é obtida diretamente da equação 2.33, que fornece:

$$Z_t = \xi_1 Z_{t-1} + \xi_2 Z_{t-2} + \dots + \xi_{p+d} Z_{t-(p+d)} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.34)$$

- Forma de choques aleatórios

$$Z_t = f(a_{t-j}); \quad j = 0, 1, 2, \dots$$

“O valor atual da série é função dos valores dos resíduos a_t atual e anteriores”.

Esta expressão pode ser escrita como:

$$Z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B)a_t \quad (2.35)$$

Onde: $\psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots$

ou seja, Z_t é a saída de um filtro linear $\psi(B)$, cuja entrada é a seqüência de ruídos brancos a_t .

Os parâmetros ψ_1, ψ_2, \dots do polinômio $\psi(B)$ podem ser obtidos diretamente de 2.33 e da própria definição de $\psi(B)$, isto é:

$$\psi(B) = \xi^{-1}(B)\theta(B)$$

obtendo os parâmetros ψ_1, ψ_2, \dots , igualando-se os coeficientes de mesmo grau em B na igualdade acima.

- Forma invertida

$$Z_t = f(Z_{t-j}, a_t); j = 1, 2, \dots$$

“Valor atual da série (Z_t) em função dos valores anteriores (Z_{t-j}) e do valor atual do resíduo (a_t)”.

Esta forma é também conhecida como forma “**inversível**” do modelo, pois é obtida diretamente da equação 2.35 explicitando a expressão para a_t , ou seja:

$$a_t = \psi^{-1}(B)Z_t$$

ou,

$$a_t = \pi(B)Z_t$$

onde: $\pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$

Desta forma, explicitando para Z_t obtém-se:

$$Z_t = (\pi_1 B + \pi_2 B^2 + \dots)Z_t + a_t \quad (2.36)$$

Os coeficientes π'_j s de $\pi(B)$ podem ser obtidos da mesma forma que os ψ'_j s da forma de choques aleatórios, lembrando que pela equação 2.32 tem-se, neste caso:

$$\pi(B) = \phi(B)\theta^{-1}(B)$$

onde π_1, π_2, \dots são obtidos igualando-se os coeficientes de mesmo grau em B na igualdade acima.

- Previsão com erro quadrático mínimo

Seja $\hat{Z}_t(l)$ (previsão de $Z_{t+l}; l \geq 1$ feita na origem dos tempos t) uma função linear das observações presentes e passadas da série: Z_t, Z_{t+1}, \dots , que por sua vez pode também ser escrita como função linear dos resíduos presentes e passados a_t (conforme mostrado na equação 2.35 da forma de choques aleatórios).

Desta forma, pode-se escrever para $\hat{Z}_t(l)$:

$$\hat{Z}_t(l) = \psi_1^* a_t + \psi_{1+1}^* a_{t-1} + \psi_{1+2}^* a_{t-2} + \dots \quad (2.37)$$

Queremos determinar os coeficientes ψ_{1+j}^* ; $j = 0, 1, \dots$; tal que $\hat{Z}_t(l)$ seja ótimo, isto é, queremos encontrar os ψ_{1+j}^* tal que o erro médio quadrático da previsão seja mínimo.

Analicamente formula-se o seguinte problema:

Determinar ψ_{1+j}^* ; $j = 0, 1, \dots$ tal que:

$E\{e_t^2(l)\}$ seja mínimo, onde:

$$e_t(l) = Z_{t+l} - \hat{Z}_t(l)$$

Os autores recomendam inicialmente determinar a expressão analítica do erro $e_t(l)$.

Considerando a equação 2.35 para $t \rightarrow t+l$ tem-se:

$$Z_{t+l} = \sum_{j=0}^{\infty} \psi_j a_{t+l-j} = \psi_0 a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l+1} a_{t-1} + \dots \quad (2.38)$$

Substituindo as expressões 2.37 e 2.38 na equação do erro $e_t(l)$ tem-se:

$$\begin{aligned} e_t(l) &= [\psi_0 a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} + \psi_l a_t + \psi_{l+1} a_{t-1} + \dots] - [\psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \dots] \\ &= \psi_0 a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} + [\psi_l - \psi_l^*] a_t + [\psi_{l+1} - \psi_{l+1}^*] a_{t-1} + \dots \end{aligned}$$

Conseqüentemente, o erro médio quadrático será:

$$E\{e_t^2(l)\} = E\{\psi_0 a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} + [\psi_l - \psi_l^*] a_t + [\psi_{l+1} - \psi_{l+1}^*] a_{t-1} + \dots\}^2$$

Tomando-se as esperanças acima e levando em consideração que os a'_i s são descorrelatados, com média nula e variância σ_a^2 , tem-se:

$$E\{e_t^2(l)\} = \sum_{j=0}^{l-1} \psi_j^2 \sigma_a^2 + \sum_{i=0}^{\infty} (\psi_{l+i} - \psi_{l+i}^*)^2 \sigma_a^2 \quad (2.39)$$

Fica então claro de 2.39 que o $\min E\{e_t^2(l)\}$ ocorre quando

$$\psi_{l+i} = \psi_{l+i}^*; \quad i = 0, 1, \dots \text{ já que } \sum_{j=0}^{l-1} \psi_j^2 \sigma_a^2 \geq 0$$

Isto implica que: $\text{Min } E\{e_t^2(l)\} = \sum_{j=0}^{l-1} \psi_j^2 \sigma_a^2$.

Em outras palavras, ressaltam os autores, a previsão “ótima” (no sentido do erro médio quadrático mínimo) é aquela em que os ψ_{l+j}^* ; $j = 0, 1, \dots$ da equação 2.37 são iguais aos ψ_{l+j} ; $j = 0, 1, \dots$.

Pode-se então, reescrever a equação 2.37 da seguinte forma:

$$\hat{Z}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \psi_{l+2} a_{t-2} + \dots \quad (2.40)$$

e o erro da previsão l - passos à frente $[e_t(l)]$ é:

$$e_t(l) = \psi_0 a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} \quad (2.41)$$

2.4 Comentários gerais do capítulo

Neste capítulo, foram desenvolvidos itens sobre a água e seus usos, séries temporais e sobre a metodologia de Box & Jenkins que será utilizada para a previsão das séries temporais em estudo.

No próximo capítulo, apresentar-se-á a metodologia aplicada ao conjunto de dados com a caracterização da área geográfica.

3 MATERIAL E MÉTODOS

Este capítulo apresenta uma descrição da área de estudo e os procedimentos e técnicas utilizados para a realização deste trabalho.

3.1 Metodologia

Os dados foram coletados da planilha de controle operacional, gerada através do Sistema de Informações Operacionais da Diretoria de Operações da Companhia Riograndense de Saneamento – CORSAN, e correspondem ao registro mensal do volume total de água distribuída e consumida em Santa Maria, no período de Janeiro de 2000 à Dezembro de 2006, sendo que os resultados do ano de 2006 não entraram na modelagem da série, sendo utilizados para comparação com os resultados gerados pela previsão do modelo escolhido.

Foram constituídas duas séries de setenta e duas (72) observações mensais, de janeiro de 2000 a dezembro de 2005. Para a formatação e tratamento dos dados, foi utilizado o pacote Microsoft Office, por meio da Planilha Eletrônica Excel 2000. Em seguida, os dados foram exportados para a planilha do *software* estatístico Statistica versão 7, no qual foram realizadas as análises e modelagem dos dados, utilizando-se a metodologia de Box & Jenkins conforme citada no item 2. Foi seguido o fluxograma proposto para essa metodologia, englobando desde a análise inicial dos dados até a validação dos modelos encontrados por meio da análise dos resíduos e a realização de previsão das variáveis em estudo. Após a estimação dos parâmetros dos modelos, foi realizada a comparação dos valores estimados com os valores reais no período previsto.

3.2 Caracterização da área de estudo

Os dados para o estudo foram coletados no município de Santa Maria – RS pela Unidade de Saneamento de Santa Maria, pertencente à Superintendência Regional Central da Corsan, com sede neste município.

A cidade foi criada a partir de acampamentos de uma comissão demarcadora de limites entre terras de domínio espanhol e português que passavam pela região em 1797.

Santa Maria também é conhecida como Cidade Cultura (pela presença de inúmeras universidades e faculdades) e Cidade Coração do Rio Grande do Sul por situar-se na região central deste Estado. Sua população estimada em 2006 era de 270.073 habitantes.

Também é conhecida como *Santa Maria da Boca do Monte* por situar-se em uma região cercada por morros originados no final do derreamento basáltico ocorrido no Pleistoceno.

Coordenadas Geopolíticas:

- Área do Município: 1823,1 Km²;
- Altitude Mínima: 41 m acima do nível do mar;
- Altitude Média: 113 m acima do nível do mar;
- Temperatura Média: 19,2° C;
- Clima: Subtropical Úmido;
- Precipitação Pluviométrica Média Anual: 1700 mm;
- Latitude do distrito sede do município: -29,68417;
- Longitude do distrito sede do município: -53,80694.

Limites geográficos:

- Norte: Itaara, Júlio de Castilhos e São Martinho da Serra.
- Sul: São Gabriel e São Sepé.
- Leste: Silveira Martins, Restinga Seca e Formigueiro.
- Oeste: São Pedro do Sul, Dilermando de Aguiar.

A Companhia Riograndense de Saneamento (CORSAN) abastece hoje mais de seis milhões de gaúchos, o que corresponde a dois terços da população do Estado. A CORSAN está distribuída em nove regiões responsáveis pela administração das diversas localidades que recebem água tratada.

As chamadas superintendências regionais são formadas pelas regiões Sul, Litoral Norte, Metropolitana, Nordeste, Planalto Médio, Missões, Fronteira Oeste, Central e Sinos. As sedes das regionais localizam-se nas cidades de Rio Grande, Osório, Canoas, Bento Gonçalves, Passo Fundo, Santo Ângelo, Uruguaiana, Santa Maria e Canoas, respectivamente.

As 367 localidades atendidas pela Corsan nas nove regiões do Estado têm ainda 126 gerências que cuidam das necessidades locais. Hoje, 95% da população urbana das localidades abastecidas pela Companhia contam com água tratada.

As principais fontes de captação de água em Santa Maria são: a Barragem Rodolfo Costa e Silva (responsável por, aproximadamente, 65% da água captada) e a Barragem do DNOS (Departamento Nacional de Obras e Saneamento), hoje extinto. A Corsan conta ainda, com um poço artesiano que capta água subterrânea, localizado na Vila Bilibio. Esse poço abastece somente a comunidade da referida vila e, segundo a Corsan, será desativado ainda este ano.

3.3 Comentários gerais do capítulo

Neste capítulo, apresentou-se a metodologia aplicada ao conjunto de dados com a caracterização da área de estudo.

No próximo capítulo, desenvolver-se-á a análise e discussão dos resultados obtidos na pesquisa.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo, apresentam-se as informações obtidas durante a elaboração deste estudo, bem como suas análises e interpretações, relacionando-as com a fundamentação teórica.

4.1 Série distribuição de água

Inicialmente foi construído o gráfico da série original “distribuição de água em Santa Maria - RS, de janeiro de 2000 a dezembro de 2005”, conforme dados do Apêndice A. Como ressaltado anteriormente, os dados do ano de 2006 serão reservados para serem comparados aos resultados obtidos na previsão doze passos à frente gerada pelo modelo escolhido, a fim de validá-lo. Desta forma, ter-se-á uma série de 72 observações mensais, representada graficamente pela Figura 4.1.

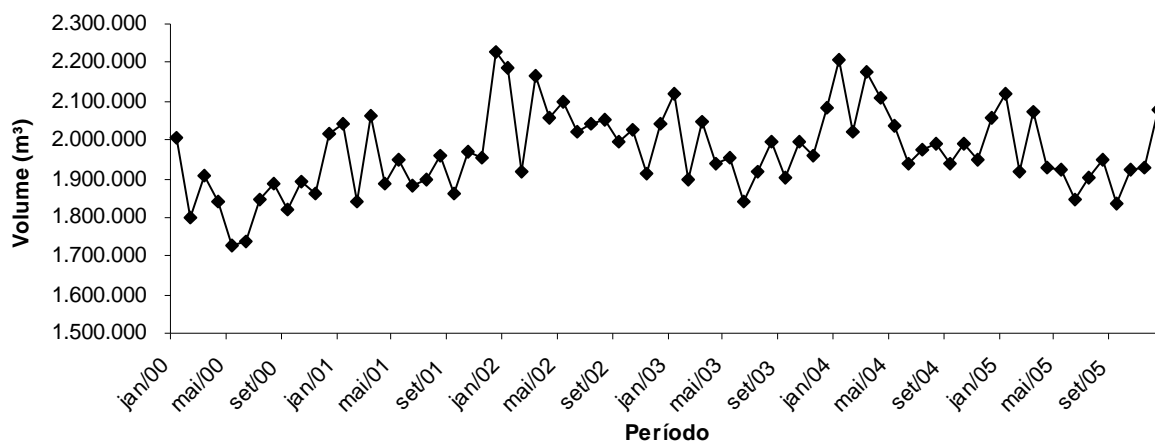


Figura 4.1 - Série original distribuição de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2005.

Através da análise do gráfico da série original, não se pode afirmar que a série é estacionária. Na Tabela 4.1, encontram-se as estatísticas descritivas da série original “distribuição de água”.

Tabela 4.1 – Estatística descritiva da série distribuição.

Estatísticas	Valores (m³)
Média	1.970.009
Mínimo	1.724.631
Máximo	2.226.780
Desvio Padrão	106.039
Coefficiente de Variação (%)	5%

A distribuição mensal média de água em Santa Maria no período foi de 1.970.009 m³, com um coeficiente de variação de 5%. Como o coeficiente de variação é menor que 20%, a média é representativa do conjunto de dados, assim como os dados apresentam uma pequena variabilidade no seu conjunto, o que indica homogeneidade. O desvio padrão foi de 106.039 m³ e o valor mínimo de distribuição ocorreu no mês de maio de 2000 e o valor máximo no mês de dezembro de 2001.

As Funções de Autocorrelação e Autocorrelação Parcial são apresentadas nas Figuras 4.2 e 4.3, respectivamente.

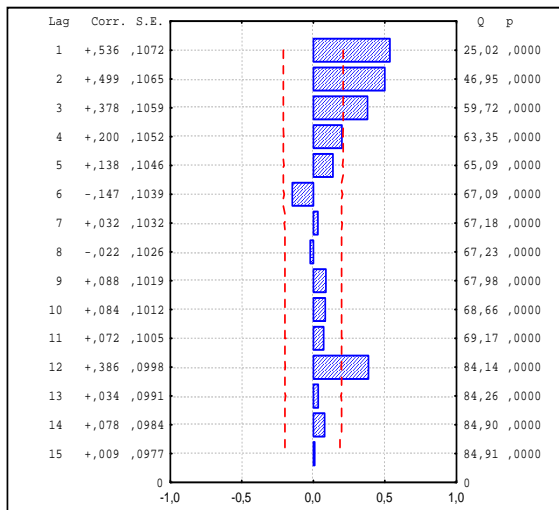


Figura 4.2 - Função de autocorrelação (FAC).

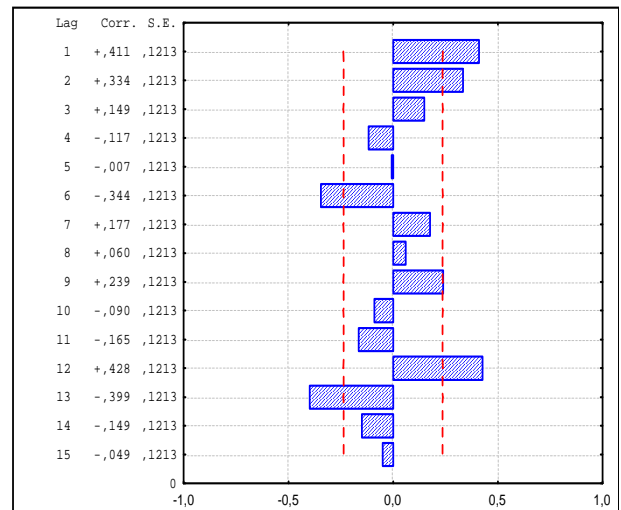


Figura 4.3 - Função de autocorrelação parcial (FACP).

Observa-se que a função de autocorrelação decresce para zero a partir do lag 4, o que pode indicar a estacionariedade da série. Observa-se também que o lag 12 é significativo, o que indica a presença de sazonalidade. Para verificar a existência de tendência na série, aplicaremos o Teste de Seqüências (Wald-Wolfowitz). Como o número de pontos A e o número de pontos B são maiores que 20, aproximamos o W_α , com $\alpha = 5\%$, pela Distribuição Normal, ou seja, $T_1 \sim N(\mu, \sigma^2)$. Como $Z_{calc} = -6,77 < Z_{tab} = -1,96 \Rightarrow T_1 < W_\alpha$, rejeitamos a hipótese H_0 que afirmava ser a seqüência aleatória, o que implicaria na ausência de tendência na série. Logo, pelo Teste da Seqüência, existe tendência na série distribuição de água.

Identificada a presença de tendência, o próximo passo, segundo a Metodologia de Box & Jenkins, é a transformação da série com a primeira diferença. O gráfico da série diferenciada e da série original encontra-se na Figura 4.4.

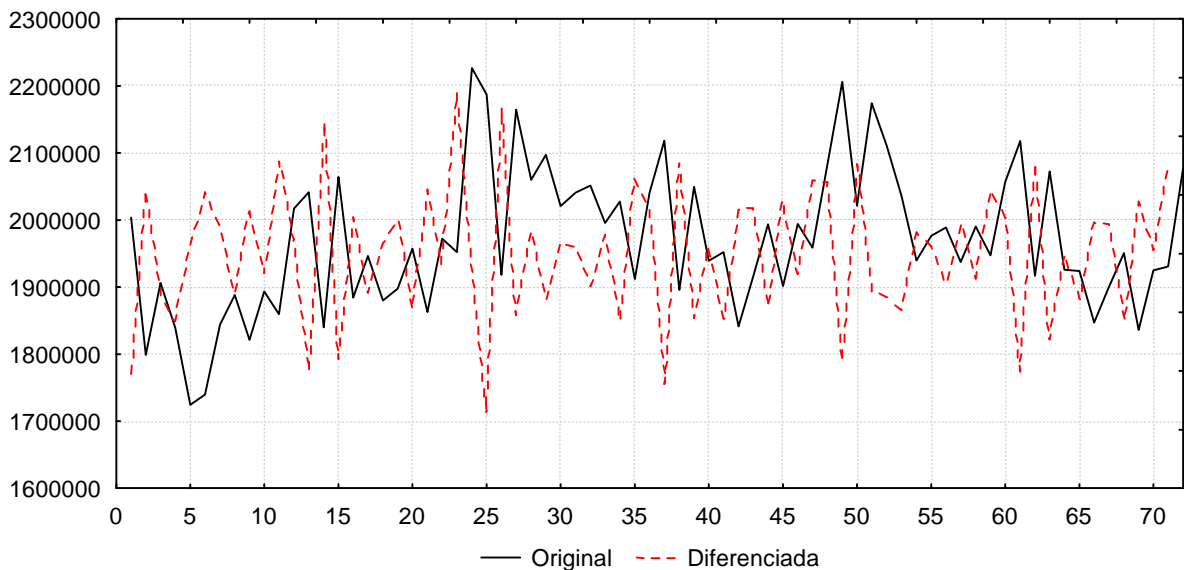


Figura 4.4 – Série distribuição de água em Santa Maria original e diferenciada.

As Funções de Autocorrelação e Autocorrelação Parcial da série transformada com diferença simples são apresentadas nas Figuras 4.5 e 4.6, respectivamente.

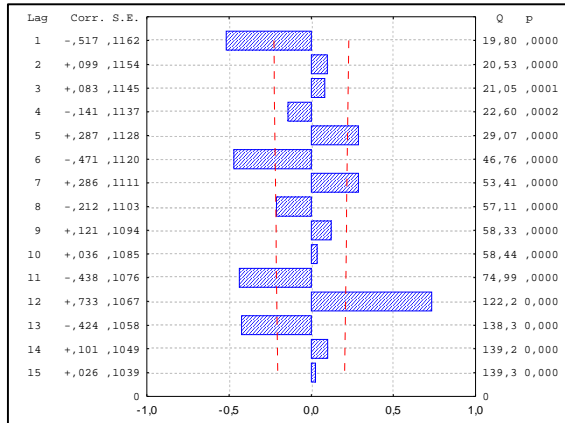


Figura 4.5 – Função de autocorrelação da série transformada.

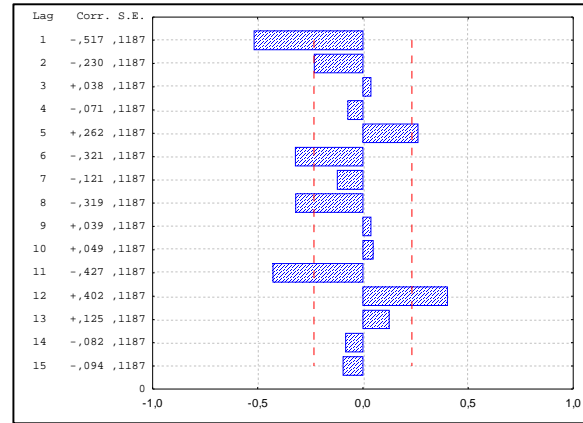


Figura 4.6 – Função de autocorrelação parcial da série transformada.

A aplicação do Teste de Seqüências na série diferenciada, com $-1,96 < Z_{calc} = 0,64 < 1,96$, o que implica na aceitação da hipótese H_0 que afirma ser a seqüência aleatória, ou seja, a série não apresenta tendência, comprova ser desnecessário continuar diferenciando a série.

Após deixar a série livre de tendência, aplicou-se o Teste de Kruskal-Wallis para verificar a existência de sazonalidade. A hipótese H_0 que diz não existir sazonalidade na série foi rejeitada, pois $T_1 = 19,63$, aproximado pela χ^2 com 11 graus de liberdade, é menor que $T_{1c} = 56,52$, o que corresponde à Região de Rejeição de H_0 . Logo, pelo Teste KW, existe sazonalidade na série distribuição de água.

A Tabela 4.2 apresenta os modelos encontrados com estimativa dos parâmetros e estatística “p” dos mesmos, assim como os critérios AIC e SBC que, juntamente com o erro quadrado médio (média do quadrado da diferença entre o valor estimado e o valor real em cada mês), ajudarão na identificação do melhor modelo.

Pelos critérios AIC e SBC, o modelo que melhor explicaria a série seria o SARIMA(1,1,0)(1,0,0)₁₂. Porém, ao observar o Erro Quadrado Médio (EQM), detecta-se que o modelo que possui melhor grau de ajustamento aos dados da série em estudo é o SARIMA(1,1,2)(1,0,0)₁₂ e, por isso, será o modelo utilizado para prever a série distribuição de água. A utilização do EQM juntamente com os critérios AIC e SBC, deve-se ao fato de os mesmos “punirem” os modelos com maior número de

parâmetros, o que poderia levar a uma desconsideração de modelos que explicariam melhor a série em estudo, apesar de serem menos parcimoniosos.

O modelo escolhido é representado pela equação:

$$(1 + 0,778297B)(1 - 0,851409B^{12})(1 - B)Z_t = (1 + 0,599813B - 0,276069B^2)a_t \quad (4.1)$$

$$\begin{matrix} (-5,99208) & (9,16670) & & (-4,11792) & (2,65841) \end{matrix}$$

Os valores entre parênteses abaixo da equação são os t_{calc} para os parâmetros do modelo.

Tabela 4.2 - Critérios de comparação e escolha do melhor modelo para a série distribuição.

Modelo	Parâmetros	Estatística p dos parâmetros	Critério AIC	Critério SBC	EQM (km ³)
SARIMA(1,1,0)(1,0,0) ₁₂	$\phi = -0,267690$ $\Phi = 0,819263$	0,032 0,000	2.165,41	2.486,71	30,224
SARIMA(0,1,1)(1,0,0) ₁₂	$\theta = 0,258834$ $\Phi = 0,822727$	0,024 0,000	2.165,44	2.486,74	28,438
SARIMA(1,1,2)(1,0,0)₁₂	$\phi = -0,778297$ $\theta_1 = -0,599813$ $\theta_2 = 0,276069$ $\Phi = 0,851409$	0,000 0,000 0,009 0,000	2.447,01	3.089,61	26,951
SARIMA(1,1,0)(0,0,1) ₁₂	$\phi = -0,427987$ $\Theta = -0,545816$	0,000 0,000	2.191,57	2512,87	44,776
SARIMA(2,1,1)(1,0,0) ₁₂	$\phi_1 = -1,11616$ $\phi_2 = -0,34757$ $\theta = -0,87378$ $\Phi = 0,83541$	0,000 0,005 0,000 0,000	2.444,91	3.087,51	28,969
SARIMA(2,1,2)(1,0,0) ₁₂	$\phi_1 = -1,68351$ $\phi_2 = -0,84461$ $\theta_1 = -1,59615$ $\theta_2 = -0,67933$ $\Phi = 0,83607$	0,000 0,000 0,000 0,005 0,000	2.585,10	3.388,35	31,089

A análise dos resíduos será feita através das Funções de Autocorrelação e Autocorrelação Parcial, apresentadas nas Figuras 4.7 e 4.8, respectivamente.

Analisando as Funções de Autocorrelação e Autocorrelação Parcial dos resíduos gerados pelo modelo SARIMA(1,1,2)(1,0,0)₁₂, verifica-se que 95% das observações estão dentro dos limites superior e inferior, o que é razoável para se admitir que não existe autocorrelação nos resíduos. Antes de afirmar que os resíduos formam um ruído branco, será aplicado o teste de Lilliefors para a normalidade dos mesmos, onde temos as hipóteses:

H_0 : as observações seguem uma distribuição normal. (aceitação: $p > \alpha$)

H_1 : as observações não seguem uma distribuição normal. ($p \leq \alpha$)

Pelo teste de Lilliefors, $p > 0,2 \Rightarrow P > 0,05$, ou seja, $p > \alpha$, logo aceita-se H_0 e os resíduos seguem uma distribuição normal.

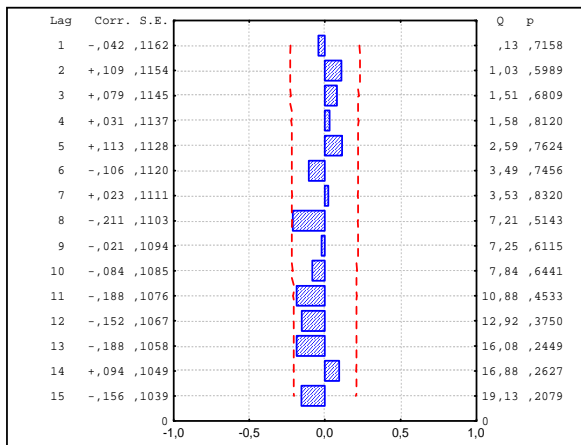


Figura 4.7 – Função de autocorrelação residual (FACr).

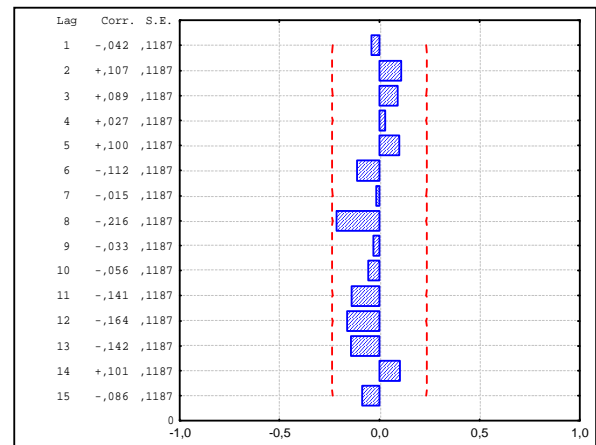


Figura 4.8 – Função de autocorrelação parcial residual (FACPr).

A Figura 4.9 mostra o gráfico da série distribuição com os valores observados e valores previstos pelo modelo escolhido e a Tabela 4.3 traz esses valores, bem como seus limites inferiores e superiores e o erro de previsão.

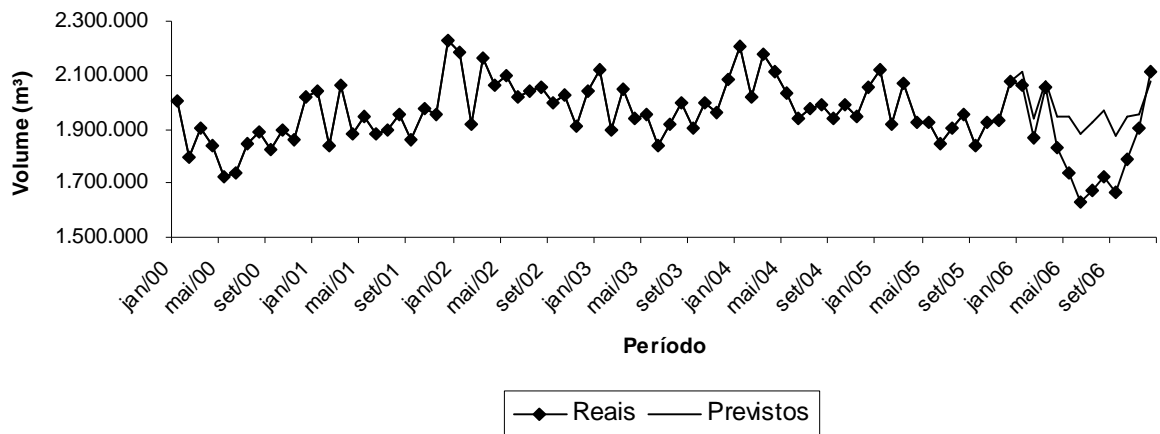


Figura 4.9 – Gráfico da série distribuição com os valores reais e valores previstos pelo modelo escolhido.

No ano de 2006, a distribuição de água na cidade de Santa Maria sofreu variação em função da estiagem ocorrida na região, que acarretou na diminuição de 50% da capacidade da Barragem do DNOS. Conseqüentemente, os valores reais diferem do restante da série, não sendo possível validar o modelo por comparação dos mesmos com os valores previstos.

Tabela 4.3 - Valores reais e previstos da série distribuição para o ano de 2006, em m³.

Período	Reais	Previstos	Limite inferior 95%	Limite superior 95%	Erro padrão de estimativa
jan/06	2.060.155	2.113.750	2.000.009	2.227.491	68.193
fev/06	1.869.956	1.937.159	1.789.958	2.084.360	88.254
mar/06	2.051.619	2.073.925	1.907.411	2.240.440	99.834
abr/06	1.832.059	1.945.941	1.756.670	2.135.213	113.478
mai/06	1.735.780	1.946.743	1.741.051	2.152.434	123.323
jun/06	1.626.434	1.879.442	1.655.761	2.103.123	134.108
jul/06	1.670.176	1.926.196	1.687.903	2.164.490	142.869
ago/06	1.726.590	1.967.760	1.714.213	2.221.308	152.014
set/06	1.669.272	1.871.443	1.604.609	2.138.276	159.980
out/06	1.785.068	1.946.257	1.665.953	2.226.561	168.056
nov/06	1.904.813	1.951.643	1.659.095	2.244.190	175.397
dez/06	2.112.562	2.076.708	1.771.956	2.381.461	182.715

4.2 Série consumo de água

Inicialmente foi construído o gráfico da série original “consumo de água em Santa Maria - RS, de janeiro de 2000 a dezembro de 2005”, conforme dados do Apêndice B. Assim como na série distribuição de água, os dados do ano de 2006 serão reservados para comparação com os resultados obtidos na previsão doze passos à frente gerada pelo modelo escolhido. Como o período é o mesmo, a série consumo de água também será de 72 observações mensais representada graficamente pela Figura 4.10.

Pela análise gráfica da série original, não se pode concluir que a série é estacionária, portanto será aplicado o Teste da Seqüência (Wald-Wolfowitz) para verificar se há necessidade de realizar diferenças na série. Na Tabela 4.4, encontram-se as estatísticas descritivas da série original consumo de água.

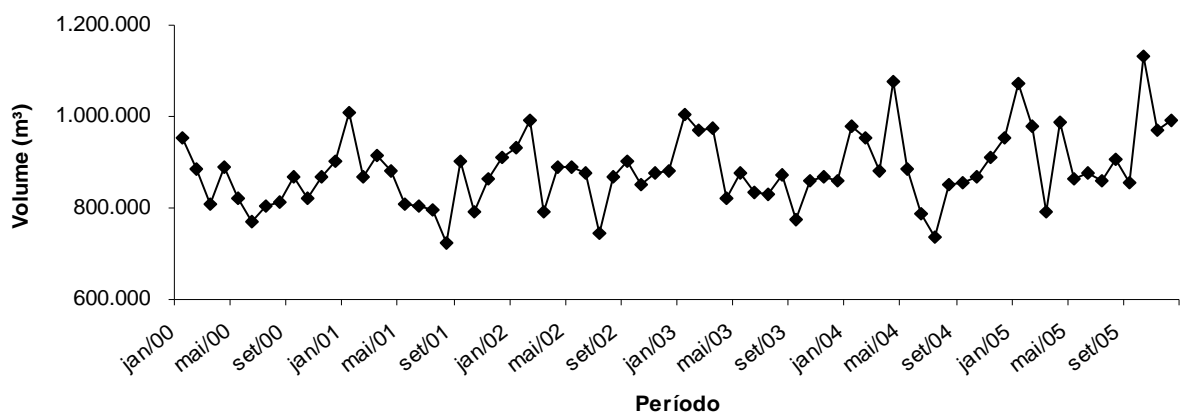


Figura 4.10 - Série original consumo de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2005.

Tabela 4.4 – Estatística descritiva da série consumo.

Estatísticas	Valores (m³)
Média	882.469
Mínimo	722.570
Máximo	1.132.127
Desvio Padrão	80.281
Coeficiente de Variação (%)	9%

O consumo mensal médio de água em Santa Maria no período foi de 882.469 m³, com um coeficiente de variação de 9%, logo a média é representativa do conjunto de dados e há homogeneidade nos mesmos. O desvio padrão foi de 80.281 m³ e o valor mínimo de consumo ocorreu no mês de agosto de 2001, enquanto que o valor máximo ocorreu no mês de outubro de 2005.

As Funções de Autocorrelação e Autocorrelação Parcial, apresentadas nas Figuras 4.11 e 4.12, respectivamente.

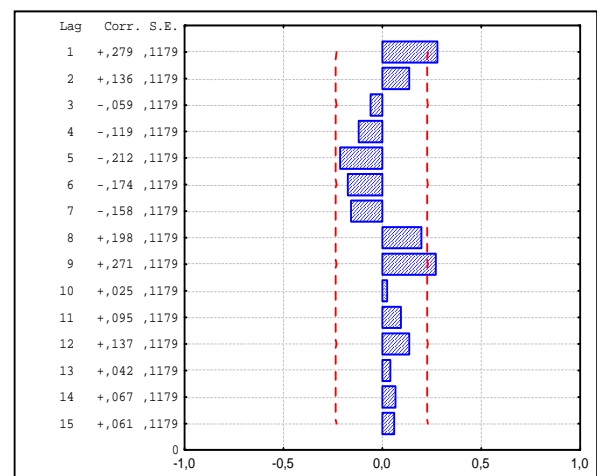
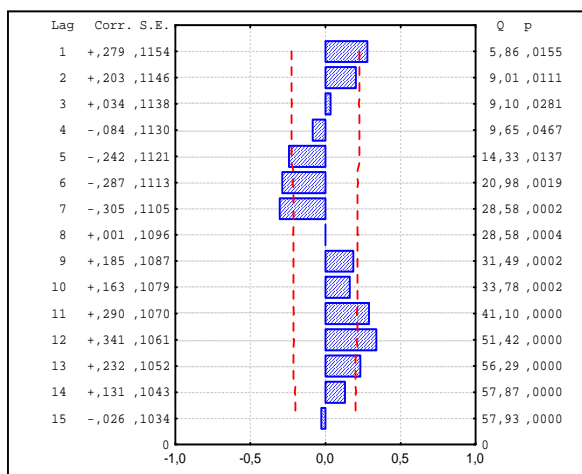


Figura 4.11 - Função de autocorrelação (FAC).

Figura 4.12 - Função de autocorrelação parcial (FACP).

Observa-se que a função de autocorrelação não decresce para zero, logo a série não apresenta estacionariedade. Para verificar a existência de tendência na série, será aplicado o Teste de Seqüências (Wald-Wolfowitz). Como o número de pontos A e o número de pontos B são maiores que 20, aproximamos o W_α , com $\alpha = 5\%$, pela Distribuição Normal, ou seja, $T_1 \sim N(\mu, \sigma^2)$. Como $-1,96 < Z_{calc} = 0,52 < 1,96$, aceita-se a hipótese H_0 que afirma ser a seqüência aleatória, o que implica na ausência de tendência na série.

Como a série não é estacionária, resta investigar a presença de sazonalidade através do Teste de Kruskal-Wallis (KW). A hipótese H_0 que diz não existir sazonalidade na série foi rejeitada, pois $T_1 = 19,63$, aproximado pela χ^2 com 11 graus

de liberdade, é menor que $T_{1c} = 34,91$, o que corresponde à Região de Rejeição de H_0 . Logo, pelo Teste KW, existe sazonalidade na série consumo de água.

Apesar da série não apresentar tendência significativa no teste aplicado, fez-se necessário a transformação da mesma por meio de uma diferença simples, pois não foi possível encontrar modelos adequados que descrevessem a série em sua forma original e satisfizessem a análise dos resíduos. As séries original e diferenciada encontram-se na Figura 4.13.

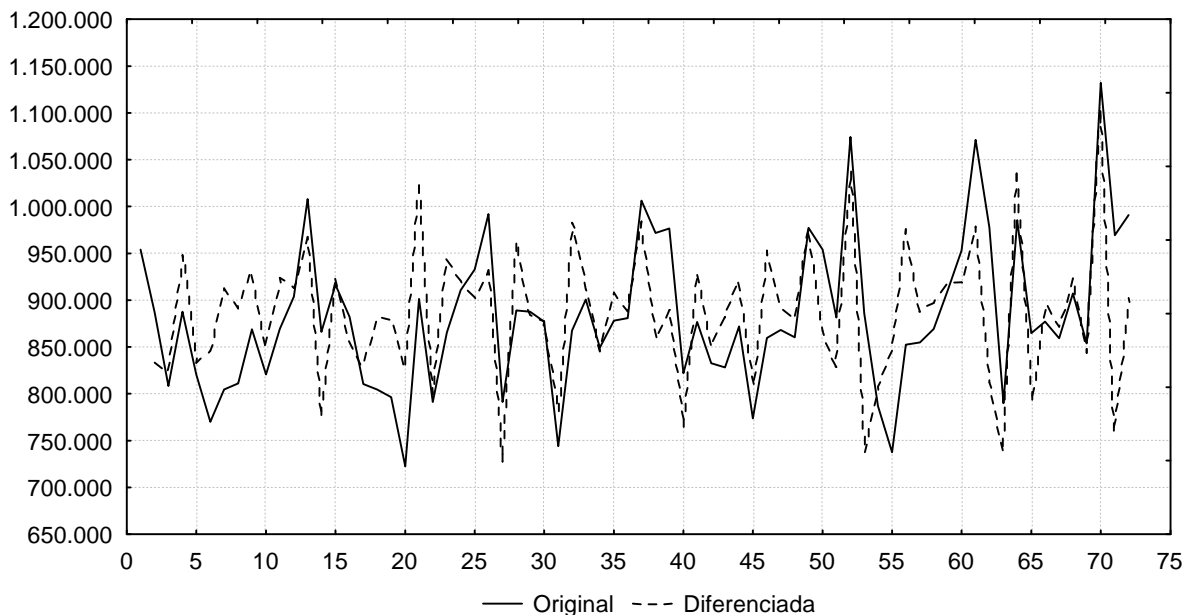


Figura 4.13 – Gráfico da série consumo de água em Santa Maria original e diferenciada.

As Funções de Autocorrelação e Autocorrelação Parcial da série transformada com diferença simples são apresentadas nas Figuras 4.14 e 4.15, respectivamente.

A Tabela 4.5 apresenta os modelos encontrados com a estimativa dos parâmetros e estatística “ p ” dos mesmos, assim como os critérios AIC e SBC que, juntamente com o erro quadrado médio (média do quadrado da diferença entre o valor estimado e o valor real em cada mês), guiarão na identificação do melhor modelo.

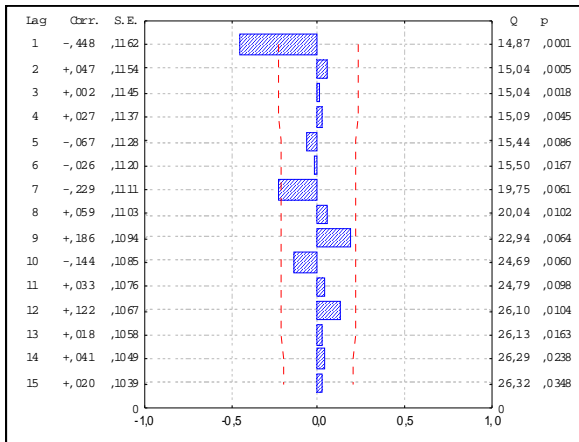


Figura 4.14 – Função de autocorrelação da série transformada.

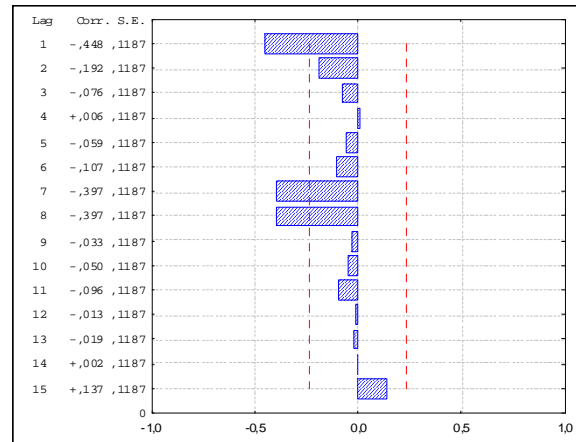


Figura 4.15 – Função de autocorrelação parcial da série transformada.

O modelo que apresentou menor valor para os critérios AIC, SBC e menor valor para o erro quadrado médio (EQM) foi o SARIMA(0,1,1)(1,0,0)₁₂. Portanto, pelos critérios, esse é o modelo que possui melhor grau de ajustamento aos dados da série “consumo de água”. Porém, ao realizar o teste de Lilliefors para a normalidade dos resíduos, os mesmos não seguem fortemente a distribuição normal, ou seja, $p < 0,05$, logo não formam um ruído branco e o modelo foi descartado. O mesmo acontece com o modelo SARIMA(0,1,1)(1,0,1)₁₂. O próximo modelo a ser considerado, pelos critérios estabelecidos, é o SARIMA(1,1,2)(0,0,1)₁₂.

O modelo escolhido é representado pela seguinte equação:

$$(1 + 0,982464B)(1 - B)Z_t = (1 + 0,467155B - 0,532661B^2)(1 + 0,300833B^{12})a_t \quad (4.2)$$

(-23,8002)
(-2,9108)
(3,3279)
(-2,4248)

Os valores entre parênteses abaixo da equação são os t_{calc} para os parâmetros do modelo.

A análise dos resíduos será feita através das Funções de Autocorrelação e Autocorrelação Parcial, apresentadas nas Figuras 4.16 e 4.17, respectivamente.

Tabela 4.5 - Critérios de comparação e escolha do melhor modelo para a série consumo.

Modelo	Parâmetros	Estatística p dos parâmetros	Critério AIC	Critério SBC	EQM (km ³)
SARIMA(2,1,0)(2,0,0) ₁₂	$\varphi_1 = -0,661825$	0,000	2.465,86	3.108,46	8,574
	$\varphi_2 = -0,315732$	0,013			
	$\Phi_1 = 0,295405$	0,036			
	$\Phi_2 = 0,420363$	0,017			
SARIMA(0,1,1)(1,0,1) ₁₂	$\theta = 0,630391$	0,000	2.320,73	2.802,68	4,640
	$\Phi = 0,999684$	0,000			
	$\Theta = 0,675014$	0,031			
SARIMA(1,1,2)(0,0,1)₁₂	$\varphi = -0,982464$	0,000	2.473,09	3.115,70	4,814
	$\theta_1 = -0,467155$	0,005			
	$\theta_2 = 0,532661$	0,001			
	$\Theta = -0,300833$	0,018			
SARIMA(0,1,1)(1,0,0) ₁₂	$\theta = 0,812066$	0,000	2.182,87	2.504,17	4,487
	$\Phi = 0,440352$	0,001			

Analisando as Funções de Autocorrelação e Autocorrelação Parcial dos resíduos gerados pelo modelo SARIMA(1,1,2)(0,0,1)₁₂, verifica-se que 95% das observações estão dentro dos limites superior e inferior, o que é razoável para se admitir que não existe autocorrelação nos resíduos. Antes de afirmar que os resíduos formam um ruído branco, será aplicado o teste de Lilliefors para a normalidade dos mesmos, onde temos as hipóteses:

H_0 : as observações seguem uma distribuição normal.

H_1 : as observações não seguem uma distribuição normal.

Pelo teste de Lilliefors, $p > 0,2 \Rightarrow P > 0,05$, ou seja, $p > \alpha$, logo aceita-se H_0 e os resíduos seguem uma distribuição normal.

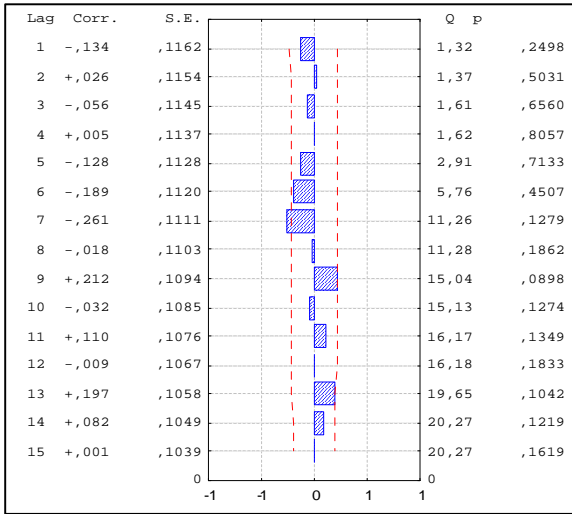


Figura 4.16 – Função de autocorrelação residual (FACr).

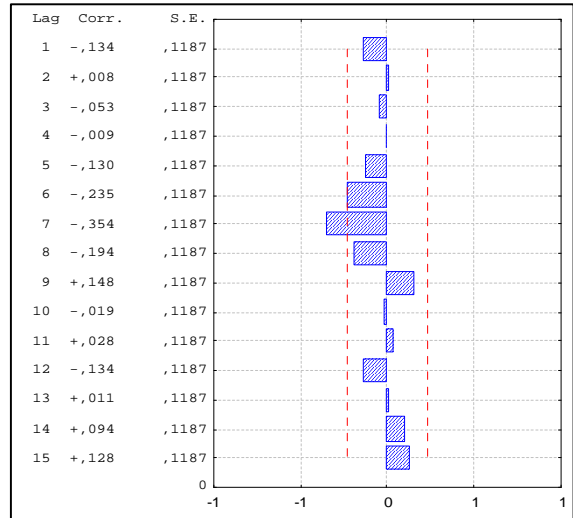


Figura 4.17 – Função de autocorrelação parcial residual (FACPr).

A Figura 4.18 mostra o gráfico da série consumo com os valores observados e valores previstos pelo modelo escolhido e a Tabela 4.6 traz esses valores, bem como seus limites inferiores e superiores e o erro de previsão.

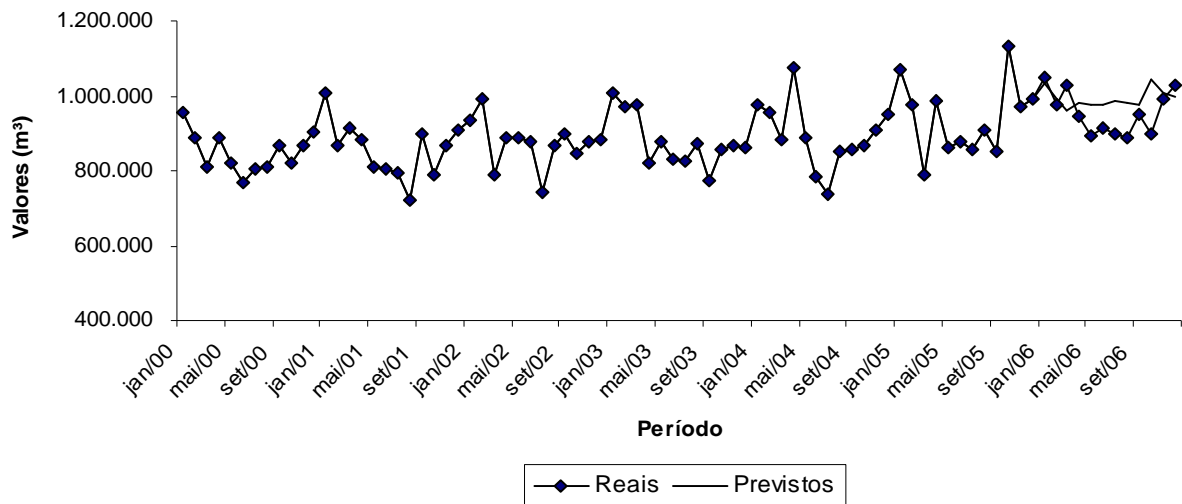


Figura 4.18 – Gráfico da série consumo com os valores reais e valores previstos pelo modelo escolhido.

Tabela 4.6 - Valores reais e previstos da série consumo para o ano de 2006, em m³.

Período	Reais	Previstos	Limite inferior 95%	Limite superior 95%	Erro padrão de estimativa
jan/06	1.048.618	1.035.544	898.866	1.172.222	81.945
fev/06	975.866	993.541	841.654	1.145.428	91.064
mar/06	1.028.723	959.586	795.290	1.123.883	98.504
abr/06	944.698	982.115	804.989	1.159.241	106.196
mai/06	891.159	977.697	789.801	1.165.594	112.653
jun/06	913.721	976.034	776.842	1.175.225	119.425
jul/06	896.969	989.099	780.253	1.197.944	125.213
ago/06	889.000	980.328	761.283	1.199.373	131.329
set/06	948.934	974.045	746.170	1.201.920	136.623
out/06	901.219	1.046.629	809.387	1.283.871	142.239
nov/06	993.734	1.006.235	760.802	1.251.668	147.149
dez/06	1.028.052	996.784	742.644	1.250.923	152.370

4.3 Comentários gerais do capítulo

Neste capítulo, desenvolveu-se a análise e discussão dos resultados obtidos na pesquisa relacionados com a fundamentação teórica.

No próximo capítulo, desenvolver-se-á a conclusão e sugestão para trabalhos futuros.

5 CONCLUSÕES E RECOMENDAÇÕES

Neste capítulo, apresenta-se uma síntese daquilo que foi mais significativo no estudo realizado, procurando apresentar os argumentos utilizados em torno dos objetivos propostos e as sugestões de temas potenciais para futuros estudos na área.

5.1 Conclusões

Os avanços tecnológicos e científicos melhoraram o padrão de vida da população em geral e também aumentaram a expectativa de vida. Com isso, houve um aumento indiscriminado da população que resultou, já no final do século XX, em uma superpopulação com 6 bilhões de habitantes.

Toda essa população necessita fundamentalmente de recursos naturais, como a água, para sobreviver. Além de ela ser limitada, sua distribuição é deficiente, pois em muitas regiões há escassez. O mais grave é que, além de a consumirmos desordenadamente, estamos contaminando-a com o nosso lixo, dejetos industriais, esgotos e agrotóxicos.

Como os governantes não se comprometem seriamente com investimentos e políticas públicas eficientes, cabe à população dar o exemplo, exigindo soluções e fazendo a sua parte. É necessário que as pessoas percebam que todas as suas ações causam impactos no meio ambiente, e, portanto, é fundamental que modifiquem seus hábitos de consumo, principalmente em relação à água.

Neste trabalho foi realizada uma análise das séries de distribuição e consumo de água na cidade de Santa Maria – RS no período de janeiro de 2000 a dezembro de 2005, utilizando a Metodologia de Box & Jenkins, com o objetivo de fazer previsões doze passos à frente, isto é, para o ano de 2006.

Após a análise gráfica das séries, identificamos que a série distribuição possui tendência crescente, isto é, aumenta no decorrer do tempo. Isto pode ser explicado, além de outros fatores, pelo aumento populacional que ocasiona uma demanda maior por água tratada. Porém, como a série consumo não apresentou a mesma tendência

crescente, podemos pensar em um aumento de ligações clandestinas, pois essas ligações não são medidas no consumo ou, aumento das perdas, o que não ocorreu, segundo os dados da Corsan. Também concluímos que, tanto a série distribuição quanto a série consumo, apresentam sazonalidade.

Através das estatísticas descritivas, concluímos que a média é representativa dos dados e a variabilidade é pequena, logo há homogeneidade em ambas as séries. Também se concluiu que os maiores valores de distribuição e consumo ocorrem nos meses de verão, e os menores valores se dão nos meses de inverno.

Com relação à modelagem, para a série distribuição, concluiu-se que o melhor modelo foi o SARIMA(1,1,2)(1,0,0) com sazonalidade de 12 meses. E, para a série consumo, o modelo que prevaleceu após o teste da normalidade dos resíduos foi o SARIMA(1,1,2)(0,0,1) também com sazonalidade de 12 meses.

A contribuição desta pesquisa para a melhoria do abastecimento de água tratada na cidade de Santa Maria vem do conhecimento da necessidade futura de distribuição em relação ao consumo da população. Assim, é possível pensar com antecedência nas soluções para futuros problemas de abastecimento apontados pela previsão e, desta forma, implementar as medidas necessárias para solucionar esses problemas.

Na seqüência, apresentam-se algumas sugestões para trabalhos futuros relacionados ao tema estudado que podem vir a complementar os resultados deste trabalho.

5.2 Sugestões para trabalhos futuros

Como o problema da escassez de água própria para consumo está se agravando cada dia mais em todo o planeta, inclusive em Santa Maria, é importante manter um controle eficiente dos gastos e captação de água, a fim de evitar racionamentos e garantir a qualidade de vida das futuras gerações. Pensando nisso, a Corsan está tomando uma série de medidas para evitar o desperdício e a perda de água em vazamentos, além de regularizar as ligações clandestinas da cidade.

Após a Corsan implantar todos os mecanismos de controle da distribuição e consumo da água na cidade, seria relevante um acompanhamento das séries

estudadas neste trabalho nos próximos anos, monitorando seu comportamento com a aplicação de testes de comparação para verificar se houve mudança significativa na rede de Santa Maria. Além disso, sugiro um estudo dos fatores que se relacionam com o consumo da água em Santa Maria, utilizando análise multivariada.

6 BIBLIOGRAFIA

COMPANHIA RIOGRANDENSE DE SANEAMENTO. **Empresa**. Disponível em <<http://www.corsan.com.br/empresa/empresa.htm>>. Acesso em 13 jun. 2007.

COMPANHIA RIOGRANDENSE DE SANEAMENTO. **Ambientais**. Disponível em <<http://www.corsan.com.br/ambientais>>. Acesso em 13 jun. 2007.

CONCEIÇÃO, T. C. **Água**: essencial para a vida no planeta. Disponível em <<http://www.cenedcursos.com.br/agua-essencial-para-a-vida-no-planeta.html>>. Acesso em 13 jun. 2007.

DIAS, A. B. **Estimação dos parâmetros do modelo de regressão linear simples pelo método não-paramétrico**. 2005, 49 p. Monografia (Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, Santa Maria, 2005.

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: 10 dez. 2006.

JACOBI, L. F. **Apostila da disciplina estatística não-paramétrica**. Universidade Federal de Santa Maria, Santa Maria, 200?, 69 p.

MARIN, S. R.; WAQUIL, P. D. **Política Cambial Brasileira nas Décadas de 1980 e 1990**: Impacto Sobre o Setor de Grãos no Rio Grande do Sul. In: 1º Encontro de Economia Gaúcha, 16 e 17 de maio de 2002. [S.l.]: Disponível em: <http://www.fee.rs.gov.br/sitefee/download/eeg/1/mesa_5_marin_waquil.pdf>. Acesso em: 10 dez. 2006.

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

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

MORETTIN, P. A.; TOLOI, C. M. **Series temporais**. Coleção Métodos Quantitativos. São Paulo: Atual, 1986.

MÜLLER, I. **Análise da causalidade entre moeda e inflação**. 1991, 87 p. Monografia (Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, Santa Maria, 1991.

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

ONG WWF BRASIL. **Natureza Brasileira**. Disponível em <http://www.wwf.org.br/natureza_brasileira/meio_ambiente_brasil/agua/index.cfm>. Acesso em 13 jun. 2007.

PREFEITURA MUNICIPAL DE SANTA MARIA. **Dados gerais de Santa Maria**. Disponível em <<http://www.santamaria.rs.gov.br/>>. Acesso em 13 jun. 2007.

RUSSO, S. L.; ILGNER, N.; MIRANDA, E. **Análise Através de Séries Temporais do Número de Pessoas que Realizam a Travessia dos Portos - Porto Xavier/San Xavier e Porto Mauá/Alba Posse. Pesquisa Operacional**, Curitiba, v. 1, n. 1, 1998. Disponível em: <<http://www.urisan.tche.br/~gep/1997/analiseseriestemporais.pdf>>. Acesso em 10 dez. 2006.

SANTOS, R. O. **Previsão da demanda de exames de direção veicular no Estado do Rio Grande do Sul**. 2006, 74 p. Monografia (Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, Santa Maria, 2006.

SIEGEL, S. **Estatística não-paramétrica**. São Paulo: McGraw-Hill, 19??.

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

UNIVERSIDADE FEDERAL DE SANTA MARIA. Centro de Ciências Naturais e Exatas. **Metodologia Box-Jenkins**. Santa Maria, [200-], 35 p.

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

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.

APÊNDICES

Apêndice A – Dados originais da variável distribuição de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2006.

Tabela 1 – Dados originais da série distribuição de água no período de janeiro de 2000 a dezembro de 2006, em m³.

Período	Dados	Período	Dados	Período	Dados
Jan/00	2.003.909	Jan/03	2.118.281	Jan/06	2.060.155
Fev/00	1.798.732	Fev/03	1.896.007	Fev/06	1.869.956
Mar/00	1.906.174	Mar/03	2.049.505	Mar/06	2.051.619
Abr/00	1.839.668	Abr/03	1.939.450	Abr/06	1.832.059
Mai/00	1.724.631	Mai/03	1.952.300	Mai/06	1.735.780
Jun/00	1.739.906	Jun/03	1.841.520	Jun/06	1.626.434
Jul/00	1.844.425	Jul/03	1.916.488	Jul/06	1.670.176
Ago/00	1.888.366	Ago/03	1.993.781	Ago/06	1.726.590
Set/00	1.821.135	Set/03	1.901.532	Set/06	1.669.272
Out/00	1.893.341	Out/03	1.994.188	Out/06	1.785.068
Nov/00	1.859.855	Nov/03	1.959.153	Nov/06	1.904.813
Dez/00	2.016.924	Dez/03	2.084.220	Dez/06	2.112.562
Jan/01	2.041.333	Jan/04	2.206.359		
Fev/01	1.840.117	Fev/04	2.021.698		
Mar/01	2.064.332	Mar/04	2.174.469		
Abr/01	1.884.536	Abr/04	2.110.812		
Mai/01	1.946.781	Mai/04	2.036.386		
Jun/01	1.880.148	Jun/04	1.939.957		
Jul/01	1.897.868	Jul/04	1.976.448		
Ago/01	1.957.415	Ago/04	1.989.066		
Set/01	1.862.796	Set/04	1.937.558		
Out/01	1.972.104	Out/04	1.990.445		
Nov/01	1.952.401	Nov/04	1.947.486		
Dez/01	2.226.780	Dez/04	2.056.721		
Jan/02	2.187.295	Jan/05	2.117.751		
Fev/02	1.918.401	Fev/05	1.916.722		
Mar/02	2.164.765	Mar/05	2.072.391		
Abr/02	2.059.945	Abr/05	1.925.936		
Mai/02	2.097.550	Mai/05	1.923.869		
Jun/02	2.020.931	Jun/05	1.847.164		
Jul/02	2.040.898	Jul/05	1.900.256		
Ago/02	2.051.660	Ago/05	1.950.492		
Set/02	1.995.957	Set/05	1.836.261		
Out/02	2.027.880	Out/05	1.924.991		
Nov/02	1.912.141	Nov/05	1.930.648		
Dez/02	2.041.147	Dez/05	2.078.061		

Apêndice B – Dados originais da variável consumo de água em Santa Maria – RS, no período de janeiro de 2000 a dezembro de 2006.

Tabela 2 – Dados originais da série consumo de água no período de janeiro de 2000 a dezembro de 2006, em m³.

Período	Consumo	Período	Consumo	Período	Consumo
Jan/00	953.941	Jan/03	1.006.315	Jan/06	1.048.618
Fev/00	887.054	Fev/03	971.805	Fev/06	975.866
Mar/00	808.240	Mar/03	976.485	Mar/06	1.028.723
Abr/00	887.801	Abr/03	822.042	Abr/06	944.698
Mai/00	820.632	Mai/03	876.762	Mai/06	891.159
Jun/00	770.134	Jun/03	832.477	Jun/06	913.721
Jul/00	804.547	Jul/03	828.067	Jul/06	896.969
Ago/00	811.250	Ago/03	871.997	Ago/06	889.000
Set/00	868.758	Set/03	773.811	Set/06	948.934
Out/00	820.562	Out/03	859.577	Out/06	901.219
Nov/00	869.205	Nov/03	868.271	Nov/06	993.734
Dez/00	903.811	Dez/03	860.504	Dez/06	1.028.052
Jan/01	1.008.260	Jan/04	977.487		
Fev/01	866.163	Fev/04	954.131		
Mar/01	916.324	Mar/04	881.467		
Abr/01	882.196	Abr/04	1.074.476		
Mai/01	810.521	Mai/04	886.062		
Jun/01	804.578	Jun/04	786.825		
Jul/01	796.360	Jul/04	737.681		
Ago/01	722.570	Ago/04	852.431		
Set/01	901.236	Set/04	854.865		
Out/01	791.423	Out/04	869.288		
Nov/01	865.681	Nov/04	910.621		
Dez/01	910.634	Dez/04	953.048		
Jan/02	932.707	Jan/05	1.071.171		
Fev/02	991.943	Fev/05	977.553		
Mar/02	791.182	Mar/05	789.864		
Abr/02	889.261	Abr/05	985.589		
Mai/02	887.496	Mai/05	864.793		
Jun/02	877.203	Jun/05	877.127		
Jul/02	743.966	Jul/05	859.403		
Ago/02	867.473	Ago/05	907.309		
Set/02	900.528	Set/05	853.834		
Out/02	849.272	Out/05	1.132.127		
Nov/02	878.098	Nov/05	969.464		
Dez/02	881.079	Dez/05	990.938		