



UNIVERSIDADE FEDERAL DO AMAPÁ
CURSO DE LICENCIATURA EM MATEMÁTICA

**MODELOS DE SÉRIES TEMPORAIS APLICADOS AO CONSUMO
DE ENERGIA ELÉTRICA DO ESTADO DO AMAPÁ**

Macapá-AP
2012

JANAÍNA LIMA DA SILVA
JOELMA DE FIGUEIREDO SOEIRO

**MODELOS DE SÉRIES TEMPORAIS APLICADOS AO CONSUMO
DE ENERGIA ELÉTRICA DO ESTADO DO AMAPÁ**

Trabalho de conclusão de curso apresentado ao
colegiado de Matemática da Universidade Fede-
ral do Amapá, como parte das exigências para
a obtenção do título de Licenciatura em Ma-
temática, sob a orientação da Prof^a. Ms. Vânia
de Fátima Lemes de Miranda.

Área de Concentração: Inferência Estatística

Orientadora: *Prof^a.Msc.Vânia de Fátima Lemes de Miranda*

Co-orientadora: *Prof^a.Msc.Simone de Almeida Delphim Leal.*

Macapá-AP
2012

JANAÍNA LIMA DA SILVA
JOELMA DE FIGUEIREDO SOEIRO

MODELOS DE SÉRIES TEMPORAIS APLICADOS AO CONSUMO DE ENERGIA ELÉTRICA DO ESTADO DO AMAPÁ

Este Trabalho de Conclusão de Curso foi julgado e aprovado pela comissão avaliadora do Colegiado de Matemática da Universidade Federal do Amapá. Composta pelos integrantes abaixo-relacionados:

AVALIADORES:

Co-Orinetadora: Prof.Msc.Simone de Almeida Delphim Leal
Unifap

Prof.Esp.João Socorro Pinheiro Ferreira
Unifap

Prof.Dr.Erasmo Senger
Unifap

Avaliado em: ____/____/____

Macapá-AP
2012

Dedicamos este trabalho às nossas famílias, especialmente as nossas mães e aos nossos pais. À nossa orientadora, professora Ms. Vânia de Fátima Lemes de Miranda, que nos ajudou muitíssimo contribuindo para a conclusão do mesmo.

Agradecemos, primeiramente , à Deus por ter nos dado força nesta jornada. Aos professores do colegiado de matemática pelo apoio e dedicação no decorrer do curso. À Prof^a.Dra.Maria Imaculada Sousa pelo suporte nas dificuldades encontradas na elaboração deste trabalho. À Companhia de Eletricidade do Amapá-CEA por disponibilizar os dados históricos de consumo de energia do estado do Amapá. Ao professor Edson que contribuiu com o nosso trabalho.

“A sabedoria de um ser humano não está no quanto ele sabe, mas no quanto ele tem consciência de que não sabe.”

(Augusto Cury)

Resumo

Este trabalho tem o intuito de descrever e analisar o comportamento do consumo de energia elétrica do Estado do Amapá através de séries temporais. Os setores analisados foram o residencial, industrial, comercial e outros, sendo descritos o consumo anual de energia, por meio da análise gráfica, utilizando-se o modelo de regressão linear simples para descrever a tendência do consumo. Considerando o consumo total de energia, os dados são mensais, compreendendo o período de janeiro de 1995 a dezembro de 2010. O método utilizado para constatar a presença das componentes sazonal e tendência foi o de regressão linear múltipla. Os modelos de séries temporais ajustados foram os de Box e Jenkins, e as previsões obtidas foram utilizadas para identificar o modelo mais adequado de acordo com o método do erro quadrático médio.

Palavras-chave: consumo de energia, séries temporais, tendência, componente sazonal, regressão linear, método do erro quadrático médio.

Abstract

This paper aims to describe and analyze the behavior of the electricity consumption in the State of Amapá through time series. The analyzed setors were the residential, industrial, commercial and other, and described the annual energy consumption, through of the graphical analysis, using the linear regression model to describe the trend of consumption. Considering the total energy consumption, the information are monthly, comprising the period from January 1995 to December 2010. The method used to assess the presence of seasonal and trend components was the multiple linear regression. The time series models adjusted were the Box and Jenkins,, and the predictions obtained were used to identify the most suitable according to the method of the mean square error.

Key words: electricity consumption, time series, trend, seasonal component, linear regression, method of the mean square error.

Lista de Figuras

2.1	Exemplos de séries temporais.	20
2.2	Temperatura do ar, de dado local, durante 24 horas.. . . .	20
2.3	Um processo estocástico interpretado como uma família de variáveis aleatórias	21
2.4	Série não estacionária quanto ao nível e inclinação.	31
3.1	Gráfico de séries temporais para o consumo de energia nos diferentes setores.	39
3.2	Gráfico de linhas da série de consumo total e da série diferenciada.	40
3.3	Gráfico da série de consumo anual (em 1000 MVH) com reta de tendência ajustada.	41
3.4	Função de autocorrelação para os resíduos dos modelos ajustados.	43
3.5	Função de autocorrelação para os resíduos dos modelos ajustados.	43
3.6	Função de autocorrelação para os resíduos dos modelos ajustados.	44
3.7	Comparações dos valores observados com os valores previstos para o ano de 2010 com o modelo SARIMA $(0, 1, 2)(0, 1, 1)_{12}$	45

Lista de Tabelas

3.1	Parâmetros da análise de regressão múltipla da variável consumo de energia como função dos meses e dos anos.	41
3.2	Parâmetros da análise de regressão múltipla da variável consumo de energia como função dos meses e dos anos.	42
3.3	Parâmetros da análise de regressão múltipla da variável consumo de energia como função dos meses e dos anos.	44
3.4	Consumo total mensal, observado e previsto para o ano de 2010, com os modelos M1: SARIMA(1,1,1)(1,1,0)12; M2: SARIMA(0,1,2)(0,1,1)12; M3: SARIMA(0,1,1)(0,1,1)12.	44
3.5	Valores do erro quadrático médio de previsão (MSE) para cada modelo. . .	45

Índice

Resumo	vi
Abstract	vii
Lista de Figuras	viii
Lista de Tabelas	ix
1 Introdução	12
1.1 Energia Elétrica no Amapá	13
1.1.1 Hidrelétrica Coaracy Nunes	13
1.1.2 Usina Termoelétrica de Santana	14
1.1.3 Sistema de Transmissão	14
1.1.4 Consumo atual	15
1.2 Problemática	15
1.3 Justificativa	15
1.4 Objetivos	16
1.4.1 Objetivo geral	16
1.4.2 Objetivos específicos	16
1.5 Metodologia	16
1.5.1 Testes sobre os resíduos	17
2 Revisão Bibliográfica	19
2.1 O que são séries Temporais?	19
2.2 Processos Estocásticos	21
2.3 Classificação de Série Temporal	21
2.3.1 Determinística	22
2.3.2 Estocástica	22
2.4 Componentes ou Movimentos	22
2.4.1 Sistemáticas	23
2.4.1.1 Tendência	23
2.4.1.1.1 Testes para a componente tendência	23

2.4.1.2	Sazonalidade	25
2.4.1.2.1	Testes para componente sazonal	25
2.4.1.3	Ciclo	27
2.4.2	Não-sistemáticas	27
2.4.2.1	Componente Irregular (ou aleatória)	28
2.5	Estacionariedade	28
2.5.1	Diferença	29
2.5.2	Estacionariedade Não Homogênea	30
2.5.3	Transformação Não linear de Z_t	30
2.5.4	Estacionariedade Homogênea	30
2.6	Previsão de Séries Temporais	31
2.6.1	Critérios para Avaliar a Adequação de Modelos a Séries Temporais	31
2.6.2	Modelos de Previsão	32
2.6.2.1	Modelo ARMA	32
2.6.2.2	Modelo de Regressão	33
2.6.2.3	Modelo ARIMA	34
2.6.2.3.1	Estruturação do Modelo ARIMA	35
2.6.2.4	Modelos Sazonais	36
2.6.2.4.1	Sazonalidade Determinística	38
2.6.2.4.2	Sazonalidade Estocástica	38
3	Resultados	39
3.1	Aplicação do Modelo de Regressão Múltipla	39
3.2	Ajuste dos Modelos SARIMA	42
	Conclusão	46
	Referências Bibliográficas	47

Prefácio

O presente trabalho propõe o estudo de séries temporais baseado no consumo de energia elétrica no Estado do Amapá e está estruturado da seguinte forma.

No capítulo 1, é apresentada uma introdução sobre energia elétrica, um histórico da Energia no Estado do Amapá, a justificativa de escolha do tema, os objetivos a serem alcançados e a metodologia onde são expostos os métodos de séries temporais utilizados na elaboração deste trabalho.

No capítulo 2, são apresentados de modo resumido os principais conceitos, modelos e métodos para a análise e previsão de séries temporais através de uma revisão bibliográfica.

No capítulo 3, tem-se a aplicação dos métodos de Regressão Múltipla e SARIMA na Série do Consumo de Energia Elétrica do Estado do Amapá.

Em sequência são apresentadas as conclusões dos métodos aplicados na série trabalhada.

Capítulo 1

Introdução

A energia elétrica é indispensável para o funcionamento das indústrias, comércio, instituições públicas e particulares, sobretudo para o consumo da própria população que a utiliza para vários fins, sendo de suma importância para o desenvolvimento da sociedade em geral. Quando se fala em energia elétrica, de imediato lembra-se das usinas hidrelétricas, isso na realidade, porque no Brasil, cerca de 90% da energia produzida provém das usinas hidrelétricas, enquanto que somente uma pequena parcela da eletricidade é produzida por usinas termoeletricas, usinas nucleares e termonucleares (ANDRADE e RIBEIRO, 2001), soluções caras e prejudiciais ao meio ambiente. Entretanto, outra solução se oferece: a economia de energia por meio de bons hábitos por parte dos consumidores e do desenvolvimento e difusão de novas tecnologias mais eficientes.

O aumento do consumo de energia nas últimas décadas, devido ao crescimento industrial e populacional, sobretudo, as mudanças climáticas e os baixos níveis fluviométricos, tiveram por consequência a crise energética em 2001, em que vários fatores foram expostos sobre o consumo de energia no país. Um fator bastante comum que contribuiu para esse problema foi o desperdício de energia elétrica no país devido à falta de informação por parte da grande maioria dos consumidores, não administrando com cautela o seu consumo (ANDRADE e RIBEIRO, 2001).

Em decorrência do consumo exorbitante de energia, novas usinas com altíssimo potencial elétrico estão sendo construídas para atender a demanda, como por exemplo, as hidrelétricas do rio Madeira (RO) e Belo Monte (PA), e ainda a usina atômica de Angra 3. Quando estiverem em funcionamento, o governo possibilitará a oferta de 12,3 mil megawatts adicionais sobre a eletricidade gerada atualmente. Novas soluções estão sendo estudadas e aplicadas, para que não haja novos racionamentos futuramente, como o que aconteceu em junho de 2001 (<http://wwwp.fc.unesp.br/lavarda/procie/dez14/-marcos/index.htm>).

1.1 Energia Elétrica no Amapá

O sistema elétrico do Estado do Amapá é um sistema isolado composto pela Usina Hidrelétrica Coaracy Nunes, pela Usina Termoelétrica de Santana e pelo Sistema de Transmissão que abastece 81% do Estado.

1.1.1 Hidrelétrica Coaracy Nunes

Um decreto assinado em 02 de março de 1956, pelo presidente Juscelino Kubistcheck de número 2.740, autorizou o Governo do então Território do Amapá a organizar a Companhia de Eletricidade do Amapá - CEA, sociedade de economia mista destinada a construir e explorar sistemas de produção, transmissão e distribuição de energia elétrica e serviços correlatos, bem como promover tudo o que fosse necessário para a expansão do mercado de energia elétrica no Território, de forma a estimular a criação de um parque industrial, porém seu objetivo principal seria a construção da Usina Hidroelétrica Coaracy Nunes nos termos do decreto número 35.701 de 23 de junho de 1954. No projeto frisava-se que a construção da obra pouco afetaria o orçamento do Governo Federal, pois 84% do capital gasto na construção viriam do arrendamento das jazidas de manganês existentes no então Território do Amapá.

Com a constituição das Centrais Elétricas do Norte do Brasil S/A - Eletronorte foi encampada os bens e instalações vinculadas à concessão da produção de energia elétrica da Usina Hidroelétrica Coaracy Nunes e sistema de transmissão associado, cabendo a mesma, promover o desenvolvimento das obras e arrematar e preparar o quadro de pessoal para a manutenção e operação do sistema Coaracy Nunes.

Em setembro de 1975, era energizado em 69 kV, o sistema de transmissão, incluindo a Subestação de Macapá, o que permitiu finalmente suprir a cidade com energia da Usina Hidroelétrica Coaracy Nunes a partir de 12 de outubro, em caráter experimental, até novembro quando teve início a operação comercial, com uma potência instalada de 40 MW. A inauguração oficial da Usina ocorreu no dia 13 de janeiro de 1976.

A Usina Hidrelétrica Coaracy Nunes atualmente teve sua capacidade ampliada para 78 MW, e foi implantada a jusante da cachoeira do Paredão, no rio Araguari, a cerca de 15km a montante da cidade de Ferreira Gomes, foi construída com uma potência instalada de 40 MW. Hoje, com a implantação da unidade 03, com a potência de 30 MW e a recapacitação das unidades 01 e 02, a oferta está em 78 MW, representando uma economia de 42 milhões de reais por ano, com a redução no consumo de derivados de petróleo utilizados nas usinas termelétricas.

O sistema não compõe os dois grandes sistemas elétricos interligados brasileiros, sendo sua área de influência restrita ao próprio Estado.

1.1.2 Usina Termoelétrica de Santana

Situado no município de Santana a 17 km de distância de Macapá, a Usina Termoelétrica de Santana é o principal parque de geração de energia elétrica do estado do Amapá, atendendo 13 dos 16 municípios do estado, o equivalente a 81% da população, beneficiando aproximadamente 500 mil habitantes.

Atendendo Gestões da Diretoria da Eletronorte junto ao Governo Federal e com o objetivo de suprir a demanda de energia elétrica do estado e livrar assim a população dos constantes e graves racionamentos, teve início em 1993 a implantação do Parque Termoelétrico de Santana com a instalação de três unidades geradoras tipo a gás LM2500, provenientes do Parque Termoelétrico de Camaçari-BA, com entrada em funcionamento da primeira unidade em abril de 1993. E as duas seguintes em julho de 1993 (TG-03) e novembro de 1993 (TG-02), perfazendo 54 megawatts instalados nesta primeira etapa.

Com o crescimento da demanda, houve a necessidade de acréscimo na disponibilidade de geração o que foi feito em 1997 com a instalação de unidades tipo motor diesel Wärt-silä 18V46 com o primeiro motor sendo instalado em novembro de 1997 (UGD-06), o segundo e terceiro motores instalados em dezembro de 1997 (UGD-04 e UGD-05) e fevereiro de 2000 instalou-se o quarto motor (UGD-07). Perfazendo uma disponibilidade total do Parque Térmico de 116 megawatts.

Em 2004, no intuito de suprir a deficiência de geração devido o aumento de demanda, foram instalados na Usina Termoelétrica de Santana, 23 motores de 1 megawatt cada.

Em 2005 foram instalados na Usina Térmica de Santana, 32 motores de 1,6 megawatts cada, para atender o Sistema Amapá, durante os serviços de melhorias nas turbinas hidráulicas de Coaracy Nunes. A Eletronorte contratou a empresa GEBRA por dois anos para disponibilizar 40 MW.

Portanto, existem 51,2 MW instalados para garantir um contrato de 40MW.

Em 2009, com o encerramento do contrato da GEBRA (Geradora Brasileira de Energia Elétrica), foi realizado processo de licitação e outra empresa saiu vencedora a SOENERGY - SISTEMAS INTERNACIONAIS DE ENERGIA, para atender um contrato de 45 MW.

1.1.3 Sistema de Transmissão

O Sistema Elétrico do Amapá compreende 9 subestações em níveis de tensão de 138, 69 e 13,8 quilovolts. A subestação Central está situada na Usina Hidrelétrica Coaracy Nunes eleva a tensão de 13,2 quilovolts das unidades geradoras para 138 quilovolts fazendo com a energia chegue as Subestações de Santana e Tartarugalzinho, onde a tensão é rebaixada para 69 e 13,8 quilovolts; as Subestações de Santana e Portuária que atende o município de Santana e Ilha de Santana; as subestações Santa Rita, Macapá II e Equatorial que fazem

o suprimento de energia da capital do estado; além das Subestações de Tartarugalzinho, Amapá e Calçoene que atendem os respectivos municípios.

1.1.4 Consumo atual

No Estado cresce gradativamente o consumo de energia, pois novas residências do interior do estado estão sendo beneficiadas pelo Programa Luz para Todos do Governo Federal, que desde 2003 possibilita a chegada de energia às comunidades mais longínquas do Estado. Em oito anos apenas duas mil ligações foram efetuadas por falta de verbas, porém em 2011 com a retomada do programa no Estado, mais domicílios receberam energia elétrica principalmente na área rural. No Amapá, a meta do Programa Luz Para Todos é o abastecimento de energia 24h nos 16 municípios do Estado. ([http://agenciaamapa.com.br/noticia/25946Programa-Luz para Todos avança no Estado do Amapá](http://agenciaamapa.com.br/noticia/25946Programa-Luz-para-Todos-avanca-no-Estado-do-Amapa)).

O Amapá atualmente não é auto-suficiente no abastecimento energético, todos os anos passa por racionamento em decorrência do baixo nível do Rio Araguari. Visando buscar melhorias nesse contexto, iniciou-se a construção da usina hidrelétrica em Ferreira Gomes e também se deu início a construção da Usina Hidrelétrica de Santo Antônio no vale do Jarí, esta terá a capacidade de produzir 300 MW de energia elétrica e proporcionará a indústria de celulose maior competitividade com outros mercados. Além disso, ofertará energia limpa a três milhões de habitantes, o equivalente a cinco vezes a população do estado do Amapá (http://www.amapadigital.net/noticia_view.php?ID=7741). Com isso o Estado do Amapá será auto-suficiente e ainda poderá exportar energia.

1.2 Problemática

A série do consumo de energia elétrica do Estado do Amapá foi analisada a partir de métodos de séries temporais. Busca-se então, quais os métodos mais eficientes para identificar e descrever os componentes da série, e quais os modelos da classe SARIMA fornecem as melhores previsões de consumo de energia para períodos futuros.

1.3 Justificativa

A energia elétrica tornou-se um dos fatores essenciais para o desenvolvimento da economia mundial, seja ela eólica, nuclear, hidráulica, etc. Isto significa que o crescimento sócio-econômico e industrial está diretamente ligado ao aumento da demanda energética, logo se não existe energia, não se pode expandir o setor industrial. Isso vem ocorrendo atualmente no Amapá com a construção de novas usinas hidrelétricas, para que o setor industrial cresça no Estado, sobretudo a indústria de celulose.

Usando os métodos de séries temporais, objeto de estudo desse trabalho, foi possível fazer a análise do histórico de consumo de energia no Amapá, fazendo inicialmente uma descrição do comportamento da série, por classe de consumo (residencial, industrial, comercial e outros), pois cada tipo de consumidor tem comportamento específico, além de tarifas próprias.

Portanto este trabalho é importante por que mostra o crescimento do consumo de energia elétrica no decorrer dos anos no Estado e as soluções encontradas para atender a crescente demanda energética no Amapá.

1.4 Objetivos

1.4.1 Objetivo geral

Estudar o comportamento da série do consumo de energia elétrica do Estado do Amapá usando séries temporais.

1.4.2 Objetivos específicos

- a) Identificar a componente tendência e a componente sazonal, presente na série;
- b) Descrever o comportamento das componentes na série;
- c) Ajustar a série a um modelo adequado para análise dos resultados;
- d) Fazer previsões de valores futuros;

1.5 Metodologia

Os dados utilizados para análise referem-se ao consumo residencial, industrial, comercial e outros (poder público, rural, serviços públicos e iluminação pública) cedidos pela Companhia de Eletricidade do Amapá (CEA) através da assessoria de planejamento empresarial.

Foram analisados os dados históricos do consumo anual de energia elétrica referentes aos anos de 1980 à 2010.

A série de consumo total de energia elétrica do Estado do Amapá compõe-se de 192 observações, correspondentes aos dados mensais, isto equivale ao período de janeiro de 1995 a dezembro de 2010. Os dados observados no ano de 2010 foram excluídos da análise e reservados para a comparação com as previsões dos modelos. Assim, a análise contou com uma série de 180 observações.

A série em estudo foi analisada inicialmente considerando o modelo de regressão linear múltipla para verificar hipótese da existência de tendência ou sazonalidade na série. Neste modelo, os anos e os meses são considerados como variáveis independentes. Havendo significância para meses, tem-se a presença do componente sazonal, e havendo significância para anos tem-se a componente tendência.

Para eliminar a componente tendência, foi aplicada uma diferença na série, usando a expressão:

$$\Delta Z(t) = Z(t) - Z(t - 1) \quad (1.1)$$

Triola (1998) apresenta a regressão linear múltipla como um modelo matemático que associa a variável dependente Y com as variáveis independentes X_i , sendo \hat{Y} o valor predito da variável dependente y e K o número de variáveis independentes:

$$\hat{Y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k \quad (1.2)$$

O método de estatística clássica da análise de regressão pode ser normalmente usado para estimar os parâmetros de modelos comuns de tendência não constante (Cryer e Chan, 2008). Logo, a análise de regressão linear múltipla foi utilizada para verificar a significância estatística desses componentes, sazonalidade e tendência.

1.5.1 Testes sobre os resíduos

Ao invés de olhar para as autocorrelações residuais individualmente pode-se testar se um grupo de autocorrelações é significativamente diferente de zero através das chamadas estatísticas Q . O uso do teste de Box-Pierce para as hipóteses

$$H_0 : \rho(1) = \dots = \rho(m) = 0$$

$$H_1 : \rho(k) \neq 0, \text{ para algum } k \in 1, \dots, m.$$

sendo a estatística de teste dada por

$$Q = n \sum_{k=1}^m r_k^2 \quad (1.3)$$

Na prática o número m de autocorrelações amostrais é tipicamente escolhido entre 15 e 30. Se o modelo ajustado for apropriado então Q terá distribuição aproximadamente qui-quadrado com $m - p - q$ graus de liberdade. Assim, valores grandes de Q fornecem indicação contra a hipótese de que as autocorrelações são todas nulas, em favor da hipótese de que ao menos uma delas é diferente de zero.

Os modelos SARIMA(p,d,q)(P,D,Q) analisados para as previsões foram SARIMA(1,1,1)(1,1,0), SARIMA(0,1,2)(0,1,1) e SARIMA(0,1,1)(0,1,1) e assumem a seguinte forma para o ajustamento

$$(1 - B)(1 - B^{12})Z_t = (1 - \theta B)(1 - \theta B^{12})a_t \quad (1.4)$$

As análises dos dados e os cálculos necessários foram feitos utilizando o software R (R Development Core team).

Capítulo 2

Revisão Bibliográfica

Neste capítulo serão expostos os principais conceitos concernentes a séries temporais e seus métodos de previsão.

2.1 O que são séries Temporais?

Série Temporal ou cronológica é uma série estatística e como tal é um conjunto de dados ordenados segundo suas características comuns, as quais servirão futuramente para fazer análises e inferências, mais especificamente é uma série cujos dados são dispostos em correspondência com o tempo, isto significa que o tempo varia, no entanto o fato e local conservam-se constantes.

Morettin e Tolo (1987) definem série temporal como um conjunto de observações ordenadas no tempo, de maneira que essas observações são ordenadas em intervalos de tempo comumente iguais. Em séries temporais a ordem dos dados é fundamental. De maneira objetiva Silva et al. descrevem uma série temporal como sendo um conjunto de observações discretas, realizadas em períodos equidistantes e que apresentam uma dependência serial entre essas observações. De forma ilustrativa, a Figura 2.1 exhibe exemplos de series temporais.

Fonte:Ribeiro, Goldschmidt e Choren (2009)

Na verdade, o que se conhece por série temporal é parte de uma trajetória dentre muitas que poderiam ser observadas, isto é, o conjunto de dados para análise.

Trajectoria é definida como uma curva do processo físico que está em observação, podendo ser expressa por:

$$Z^1(t), Z^2(t), Z^3(t), \dots \quad (2.1)$$

para cada t constante, obtém-se o valor de uma variável aleatória $Z(t)$, o que significa que terá uma certa distribuição de probabilidade(MORETTIN e TOLOI,1987).

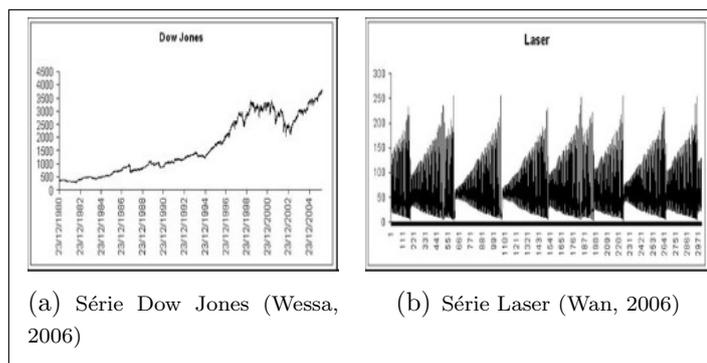


Figura 2.1: Exemplos de séries temporais.

Cada curva do gráfico abaixo é chamada de trajetória ou série temporal ou função amostral.

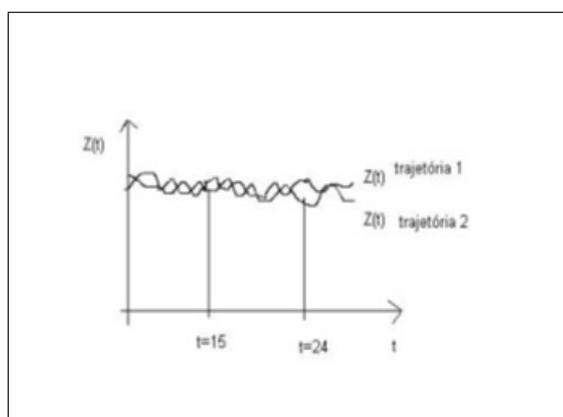


Figura 2.2: Temperatura do ar, de dado local, durante 24 horas..

Fonte: Morettin e Toloí (1987)

No gráfico acima $Z^j(t)$ é o valor da temperatura no instante t , para a j -ésima trajetória (j -ésimo dia de observação). Para cada t constante, teremos os valores de uma variável aleatória $Z(t)$ que terá certa distribuição de probabilidade (MORETTIN e TOLOI, 1987).

Dessa forma conclui-se que série temporal pode ser explicada como a parte de uma trajetória ou ainda de uma conversão parcial do processo estocástico. De acordo com Dias (2008), um processo estocástico está estatisticamente determinado quando se conhece suas funções de distribuição até a n -ésima ordem. O autor ressalta que na prática, não é possível conhecer todas as funções de distribuição até a n -ésima ordem tendo em vista uma única observação do processo estocástico, a partir do qual se deseja inferir características do mecanismo gerador da série. Porém, normalmente é possível construir um modelo simplificado da série temporal que explica sua natureza aleatória de uma maneira útil para fazer previsões.

2.2 Processos Estocásticos

Um processo estocástico é um modelo de série temporal que oferece uma descrição da natureza aleatória do processo que gerou a amostra de observações em estudo. Ou seja, admite-se que a amostra obtida foi retirada de uma função de distribuição de probabilidade Z e que cada Z_t é uma variável aleatória (DIAS, 2008).

Para Prado (2011) um processo estocástico é uma coleção de variáveis aleatórias seqüenciadas no tempo e definidas em um conjunto de pontos T podem ser contínuos ou discretos.

Considere 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 (v.a.) definida num espaço de probabilidades (Ω, A, P) .

O conjunto T é normalmente tomado como o conjunto dos inteiros $Z = 0, \pm 1, \pm 2, \dots$ ou o conjunto dos reais IR . Como, para $t \in T$, Z_t é uma v.a. definida sobre Ω , então Z_t é uma função de dois argumentos, $Z(t, \omega), t \in T, \omega \in \Omega$ (BEZERRA, 2006).

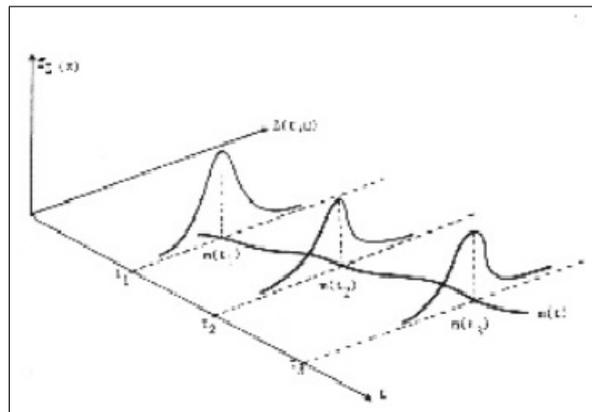


Figura 2.3: Um processo estocástico interpretado como uma família de variáveis aleatórias

Fonte: MORETTIN e TOLOI(1987)

Um processo estocástico é caracterizado por uma família de variáveis aleatórias que descrevem a evolução de algum fenômeno de interesse. Processos estocásticos que caracterizam os estudos de séries temporais descrevem a evolução temporal de um fenômeno de interesse (PELLEGRINE, 2000).

2.3 Classificação de Série Temporal

De acordo com Milone (2006) uma Série Temporal pode ser classificada como determinística ou estocástica.

2.3.1 Determinística

Uma série é dita determinística se seus valores podem ser representados por uma função univariada (em que a série temporal é explicada apenas por seus valores passados) do tipo:

$$y = f(t) \tag{2.2}$$

em que,

y é o valor da variável em estudo;

t a data a que ela se refere;

f a regra que vincula a variável a um ponto específico da linha do tempo.

Os modelos univariados de séries temporais consideram como essencial o comportamento das variáveis no tempo. Dessa forma, esses modelos não se preocupam com as variáveis explicativas, e sim com o estudo da evolução no tempo e da elaboração de uma função, na qual a variável dependente aparece em função dela mesma, regredida no tempo e/ou com os termos de erros, também defasados (BARROS, 2003 citada por RODRIGUES, SILVA e LINDEN, 2006).

2.3.2 Estocástica

Diz-se que uma série temporal é estocástica, quando a série envolve além de uma função matemática do tempo também um termo aleatório, representada por:

$$y = f(t, \epsilon) \tag{2.3}$$

2.4 Componentes ou Movimentos

Uma série temporal é descrita em função de suas componentes, sendo $Z(t)$ o valor observado da série no tempo t , T_t a componente tendência, S_t a componente sazonal e a_t o erro aleatório (Morettin e Tolo, 1987). O modelo citado por Morettin e Tolo (1987) é classificado como modelo aditivo, que pressupõe componentes não correlacionadas atuando de modo absoluto e independente; sendo aplicado às séries cujas taxas de variação são constantes (MILONE, 2006). Isto é, usam-se quando a componente sazonal (S_t) não depende das outras componentes, como por exemplo, a tendência (T_t).

As séries temporais são analisadas por meio de seus principais movimentos, conhecidas como tendência, sazonalidade, ciclo e aleatória sendo assim subdivididas em componentes sistemáticas, e, componentes não-sistemáticas (MILONE, 2006).

2.4.1 Sistemáticas

São componentes que apontam movimentos regulares.

2.4.1.1 Tendência

Milone (2006) expõe a componente tendência como uma componente sistemática que é a indicadora da direção global dos dados, portanto, relaciona-se ao incremento ou ao decréscimo dos valores da mesma com o decorrer do tempo. Dias (2010) compartilha do mesmo conceito ao afirmar que tendência é o componente mais importante de uma série temporal; consiste no movimento de direção geral que cobre um longo período de tempo e reflete as mudanças ocorridas nos dados de modo bastante consistente e gradual. Pode ser visualizado pela linha em torno da qual os dados flutuam.

Segundo Campos (2008), tendência em uma série temporal é a mudança gradual observada por meio da variação dos valores da série ao longo do tempo e que se mantém ao se remover os componentes de ciclos, sazonalidade e fatores aleatórios. E ainda, quando aplicada em consumo de energia elétrica, esse conceito, normalmente está relacionado ao comportamento ao longo do tempo devido à alteração correspondente na quantidade de consumidores ou a mudanças do perfil de clientes percebidas em longo prazo.

2.4.1.1.1 Testes para a componente tendência Para fazer a estimação da componente tendência T_t na série é necessário aplicar testes específicos. Caso exista outra componente (como sazonalidade S_t) na série, além da tendência, teríamos de eliminá-la antes de testar a presença da tendência (MORETTIN e TOLOI, 1987).

Os testes podem ser realizados de duas maneiras, antes da estimação de T_t ou depois que se obtém uma estimativa de T_t .

Existem três testes não-paramétricos para serem realizados antes da estimação, que geralmente baseiam-se em hipóteses que podem não estar verificadas para o caso de uma série temporal: Teste do sinal (Cox-Stuart), Teste de sequências (Wald-Wolfowitz) e Teste baseado no coeficiente de correlação de Spearman.

Morettin e Toloí (1987) alertam para sempre ter precaução ao utilizar tais testes, por serem pouco eficientes na detecção de alternativas de interesse.

Teste do sinal (Cox-Stuart)

As observações são agrupadas em pares $(Z_1, Z_{1+c}), (Z_2, Z_{2+c}), \dots, (Z_{N-c}, Z_N)$ onde $c = \frac{N}{2}$, se N é par e $c = \frac{N+1}{2}$, se N é ímpar. A cada par associamos o sinal + se $Z_i < Z_{i+c}$ e o sinal - se $Z_i > Z_{i+c}$, eliminando os empates. Considere n número de pares em que $Z_i \neq Z_{i+c}$.

Serão testadas as hipóteses

$$H_0 : P(Z_i < Z_{i+c}) = P(Z_i > Z_{i+c}) \forall i \quad (2.4)$$

$$H_1 : P(Z_i < Z_{i+c}) \neq P(Z_i > Z_{i+c}) \forall i \quad (2.5)$$

tem-se então que este é um teste bilateral.

A estatística do teste é T_2 , que corresponde ao número de pares com sinal positivo (+). Para $n \leq 20$, a regra de decisão é baseada na distribuição binomial com parâmetros n e $p = \frac{1}{2}$. Para um dado nível de significância α , rejeitamos H_0 se $T_2 \leq b$ ou $T_2 \geq n-b$ sendo b igual a $\alpha/2$ -quantil da binomial (n,p) .

Teste de sequências (Wald-Wolfowitz)

Supondo m a mediana da série; atribuímos a cada valor Z_t os símbolos A, se $Z_t \geq m$ e B se $Z_t < m$. Teremos então $N = (n_1 \text{ pontos A}) + (n_2 \text{ pontos B})$.

A estatística do teste é T_1 o número total de sequências (isto é, grupos de símbolos iguais).

Rejeitamos H_0 se há poucas sequências, ou seja, se T_1 é pequeno. Para um dado nível de significância α rejeitamos H_0 se $T_1 < w_\alpha$, onde w_α é o α -quantil da distribuição de T_1 , que é tabelado (Tabela 23 de Conover, 1971).

Para n_1 ou n_2 maior que 20 podemos usar a aproximação normal, isto é, $T_1 \sim N(\mu, \sigma^2)$ sendo que

$$\mu = \frac{2n_1n_2}{N} + 1 \quad (2.6)$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2) - N}{N^2(N-1)} \quad (2.7)$$

Teste baseado no coeficiente de correlação de Spearman

Este é um teste mais poderoso para identificar variáveis de interesse do que o teste de Cox-Estuart. Considere R_t o posto de Z_t dentre as N observações; se houver empates utilize postos médios.

A estatística de teste é

$$T_3 = \sum_{t=1}^N (R_t - t)^2 \quad (2.8)$$

sendo, $t = 1, \dots, N$ são os postos "naturais" dos instantes de tempo.

Rejeitamos H_0 se $T_3 \leq c_1$ ou $T_3 \geq c_2$, os valores de c_1 e c_2 são tabelados (Tabela 9 de Conover, 1971).

O teste pode ser feito também em função do coeficiente de correlação de Spearman dado por:

$$\rho = 1 - \frac{6T_3}{N(N^2 - 1)} \quad (2.9)$$

(tabela 10 de Conover,1971)

2.4.1.2 Sazonalidade

Esse movimento é classificado como sistemático, causados por fenômenos não econômicos. A sazonalidade aparece quando as observações são intra-anuais, isto é, registradas mensalmente, trimestralmente ou diariamente (MORETTIN e TOLOI, 1987), em geral quando o intervalo de tempo for inferior a um ano.

Nas séries temporais de energia elétrica é comum observar a existência de comportamentos recorrentes que caracterizam a sazonalidade devido à influência de fatores exógenos (fatores externos) em diferentes horizontes de observação (CAMPOS, 2008). Dias (2010) diz que variações sazonais são variações cíclicas de curto prazo (menos de um ano). Referem-se a movimentos de padrões regulares ou repetidos em torno da linha de tendência, em que os dados de uma série temporal parecem obedecer a períodos como meses, semanas, dias, etc.

Há dois tipos de sazonalidade, a aditiva quando a série apresenta flutuações sazonais mais ou menos constantes não importando o nível global da série e a multiplicativa quando tamanho das flutuações sazonais varia havendo dependência do nível global da série (EHLERS, 2009).

2.4.1.2.1 Testes para componente sazonal Para verificar a existência da componente sazonal os testes podem ser realizados antes e depois da estimação, para isso é necessário testar as seguintes hipóteses.

H_0 : não existe sazonalidade determinística

H_1 : existe sazonalidade determinística

A eliminação da tendência deve ser feita caso esteja presente na série. Posteriormente aplicar o teste aos resíduos (valores) $Y_t = Z_t - \hat{T}_t$, isso para o modelo aditivo. Os testes utilizados podem ser paramétricos e não-paramétricos.

Testes Paramétricos

Utiliza-se o teste F usual da análise de variância, antes da estimação de S_t . O modelo subjacente é

$$Y_{ij} = S_j + e_{ij}, i = 1, \dots, n_j; j = 1, \dots, s \quad (2.10)$$

supondo $e_{ij} \sim N(0, \sigma^2)$, independentes.

Serão testadas as hipóteses

$$H_0 : S_1 = S_2 = \dots = S_s \text{ (não existe sazonalidade)}$$

$$H_1 : S_i \neq S_j, \text{ para algum } i \text{ e } j \text{ (existe sazonalidade)}$$

A estatística de teste é:

$$T_3 = \frac{N - s}{s - 1} \cdot \frac{\sum_{j=1}^s n_j (\bar{Y}_{.j} - \bar{Y})^2}{\sum_{j=1}^s \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2} \quad (2.11)$$

que tem, sob H_0 , uma distribuição $F(s - 1, N - 1)$.

Testes Não-Paramétricos

i) Teste de Kruskal-Wallis

Tomando uma amostra de tamanho n_j de uma população com período s , em que as observações são expressas

$$Y_{ij}, j = 1, \dots, s; i = 1, \dots, n_j; N = \sum_{j=1}^s n_j \quad (2.12)$$

De maneira que as observações Y_{ij} são substituídas por seus postos R_{ij} , sendo obtidos a partir de todas as N observações após ordená-las. Considere $R_{.j}$ a soma dos postos associados a j -ésima amostra.

$$R_{.j} = \sum_{i=1}^{n_j} R_{ij}, j = 1, \dots, s. \quad (2.13)$$

A estatística de teste é

$$T_1 = \frac{s}{N(N+1)} \sum_{j=1}^s \frac{R_{.j}^2}{n_j} - 3(N+1) \quad (2.14)$$

H_0 será rejeitada se $T_1 \geq T_{1c}$, em que T_{1c} é o valor crítico tal que $P_{H_0}(T_1 \geq T_{1c}) = \alpha$, α é o nível de significância do teste.

Para n_j suficientemente grande, ou $s \geq 4$, a distribuição de T_1 , sob H_0 , pode ser aproximada por uma variável χ^2 com grau de liberdade $s - 1$ graus de liberdade (11, no caso de dados mensais).

ii) Teste de Friedman

Enquanto a ordenação no teste de Kruskal-Wallis é feita a partir de todas as N observações para obtenção de R_{ij} , a ordenação no teste de Friedman é feita dentro de cada ano. Este teste considera os anos independentes, em outras palavras, as observações de um ano são independentes das observações de outro ano qualquer.

A estatística do teste é

$$T_2 = \frac{s}{ps(s+1)} \sum_{j=1}^s (R_{.j}^*)^2 - 3p(s+1) \quad (2.15)$$

sendo,

p : número de anos;

s : período de sazonalidade ($s=12$),

$R_{.j}^*$: soma dos pontos no j -ésimo mês, isto é, $R_{.j}^* = \sum_{i=1}^p R_{ij}^*$, R_{ij} é o posto de Y_j dentro do i -ésimo ano, $j = 1, \dots, s$;

T_2 pode ser aproximado por um χ^2 com $s - 1$ graus de liberdade.

2.4.1.3 Ciclo

A componente cíclica é entendida como a parte da série temporal que apresenta um movimento em torno da tendência ao longo do tempo; pode ser identificada no intervalo entre as observações maiores ou igual a um ano (FONSECA, MARTINS e TOLEDO, 1985). Conforme Fávero et. al. (2003), o ciclo é um movimento de aparência quase periódica com fases de pico e fases de vale alternadamente. A decomposição de uma série em componentes cíclicas é chamada de análise de Fourier, ou análise espectral.

Variações cíclicas são conhecidas como variações que representam movimentos regulares em torno da reta ou da curva de tendência e referem-se às oscilações ao longo prazo. Os ciclos podem ser ou não periódicos, isto é, podem ou não seguir padrões análogos, depois de intervalos de tempos iguais (DIAS, 2010).

2.4.2 Não-sistemáticas

São componentes que apresentam movimentos irregulares, dispares, tendo por referência os parâmetros dos períodos anteriores.

2.4.2.1 Componente Irregular (ou aleatória)

A componente irregular ou aleatória relaciona-se com os movimentos imprevistos gerados aleatoriamente dentro de uma série (CAMPOS, 1991 citado por FÁVERO et.al.). Essa componente é uma mistura brusca de perturbações irregulares e esporádicas nos movimentos das séries que tipificam os fenômenos, isto é, a resultante dos efeitos de múltiplas causas (naturais, sociais, econômicas, etc.), as quais não há como saber quando ocorrerão nem suas intensidades pois, são desconhecidas e imprevisíveis (MILONE, 2006).

Reis (2005) se refere a aleatória como movimento de uma série temporal que se desloca esporadicamente, provocados por eventos casuais. O autor ainda chama a atenção para os movimentos dessa natureza, pois requerem atenção especial, tendo em vista que podem causar novos movimentos cíclicos.

A relevância da componente aleatória está diretamente ligada ao poder de alterar a direção da tendência, assim como a amplitude dos ciclos existentes. Disso decorre a precisão nas estimativas relativas a fenômenos em que a componente aleatória é irrelevante em relação as componentes sistemáticas.

2.5 Estacionariedade

Uma das suposições mais freqüentes que se faz a respeito de uma série temporal é a de que ela seja estacionária, ou seja, ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável (MORETTIN e TOLOI, 1987). Isto significa que a média da série temporal e as variâncias entre as suas observações não variam com o tempo.

Uma série temporal pode ser estacionaria em um período muito longo, mas também em um período muito curto mudando de nível ou inclinação. Neste ultimo caso, a série é não estacionária homogênea (REIS, 2005), mais adiante será descrito que a presença de não estacionariedade é uma das razões para transformar os dados originais.

Pode-se ainda dizer que é uma série que não apresenta tendência, somente sazonalidade, por conseqüência uma série não estacionária é aquela que apresenta as duas componentes ou nenhuma componente, isto é, tendência e sazonalidade (SÁFADI, 2004 citada por BOTOLINI et. al., 2011), essa é a mais comum de ser encontrada, devido a presença da componente tendência, como por exemplo as séries econômicas.

Iquiapaza et. al. (2005) define uma variável estacionária como aquela cujo valor não é afetado permanentemente pelos termos de erro contidos nas observações passadas. Um exemplo simples é o processo de ruído branco:

$$y_t = \epsilon_t \quad (2.16)$$

em que t representa o tempo, y é a variável dependente e ϵ uma variável aleatória não correlacionada serialmente com média zero e variância constante $var(\epsilon_t) = \delta^2$. Também pode ser estacionária ao redor de uma tendência.

$$y_t = T_t + \epsilon_t \quad (2.17)$$

em que T é o coeficiente que caracteriza a tendência no tempo. Em ambos os casos, é independente dos valores anteriores de $\epsilon_t(\epsilon_{t+1}, \epsilon_{t-2}, \dots)$.

Uma variável não estacionária é aquela cujo valor corrente é permanentemente afetado pelos termos de erro contidos nas observações passadas. O exemplo clássico é o caminho aleatório (random walk)

$$y = y_{t-1} + \epsilon_t \quad (2.18)$$

Existem diferentes testes para identificação de não estacionariedade, sendo os mais conhecidos os de Dickey-Fuller e Phillips-Perron. Porém estes foram desenvolvidos para analisar as séries individualmente, isto é, apresenta uma limitação ao assumir implícita ou explicitamente que as unidades sendo analisadas se comportam de forma independente (IQUIAPAZA et. al., 2005).

2.5.1 Diferença

Pode-se ainda considerar um processo estacionário caso o processo estocástico subjacente que gerou a série seja invariante em relação ao tempo. No entanto, geralmente a maior parte das séries encontradas são não estacionárias, e, para transformá-las em estacionárias é necessário utilizar uma das propriedades de série temporal, a capacidade de se tornar estacionária a partir de sucessivas diferenciações na série original, um exemplo bem comum desse tipo de série são as séries econômicas. De acordo com Morettin e Tolo (1987), a primeira diferença de $Z(t)$ é definida por:

$$\Delta Z(t) = Z(t) - Z(t - 1) \quad (2.19)$$

A segunda diferença é:

$$\Delta^2 Z(t) = \Delta[\Delta Z(t)] = \Delta[Z(t) - Z(t - 1)] \quad (2.20)$$

No geral, a d-ésima diferença de $Z(t)$ é:

$$\Delta^d Z(t) = \Delta[\Delta^{d-1} Z(t)] \quad (2.21)$$

2.5.2 Estacionaridade Não Homogênea

Uma série pode apresentar várias formas de não estacionariedade e, a partir de um número finito de diferenças a série se torna estacionária, a qual é chamada de série não estacionária homogênea. Encontra-se nas literaturas dois casos comuns de não estacionariedade homogênea:

1. Quanto ao Nível: oscila ao redor de um nível médio durante algum período de tempo e depois saltam para outro nível temporariamente, sendo que para torná-las estacionárias é suficiente tomar uma diferença;
2. quanto à inclinação: oscila em uma direção e depois mudam para outra direção temporária. Para torná-las estacionárias é necessário tomar a segunda diferença.

2.5.3 Transformação Não linear de Z_t

A diferenciação nem sempre é suficiente para tornar uma série estacionária. Assim Morettin diz que será necessário atentar para certa transformação não-linear de Z_t e $Z^* = \log Z_t$, sendo na maioria das vezes suficiente para alcançar a homogeneidade. De maneira, que possa modelar (séries econômicas).

$$\Delta \log Z_t = \log Z_t - \log Z_{t-1} \quad (2.22)$$

2.5.4 Estacionaridade Homogênea

Outra classe que será capaz de descrever de maneira satisfatória as séries estacionárias e não-estacionárias, mas que não apresentam comportamento explosivo são os modelos ARIMA (autoregressivos-integrados-médias-móveis). Este tipo de 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 (BEZERRA, 2006). A figura 2.4 ilustra esta forma de não estacionariedade.

Fonte: MORETTIN e TOLOI (1987)

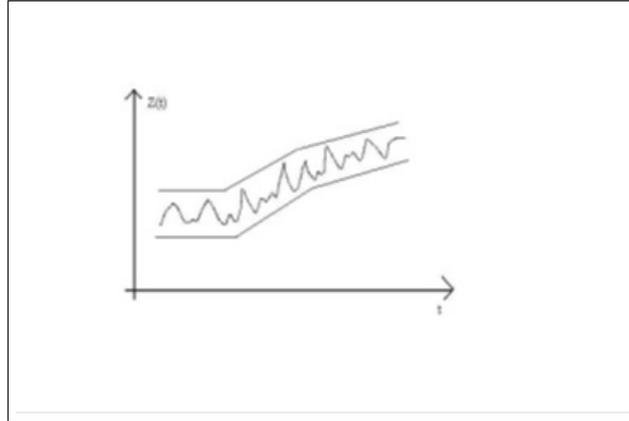


Figura 2.4: Série não estacionária quanto ao nível e inclinação.

2.6 Previsão de Séries Temporais

Um dos objetivos da análise de séries temporais é avaliar o modelo nos termos de previsão de valores futuros da série observada, ou seja, predizer o futuro possibilita fazer planos futuros a longo, médio e curto prazo e, tomar decisões apropriadas. Morettin e Tolo (1987) salientam que não necessariamente um modelo conduz a um procedimento de previsão, sendo necessário especificar uma função-perda.

Ainda segundo Morettin (1987) quando se deseja fazer previsões em alguma área há dois enfoques a considerar, no primeiro o analista se baseia na teoria para construir um modelo em quanto no segundo o modelo é obtido diretamente dos dados disponíveis, sem recorrer a uma possível teoria subjacente. Este segundo é chamado de séries temporais.

A análise de uma série temporal se dá a partir da construção no qual pode se observar características importantes como tendência e sazonalidade, variabilidade, observações atípicas, etc. Entretanto sempre acompanhada de uma análise gráfica (PRADO, 2011).

Para que seja possível prever os valores futuros de uma série com base em valores passados, Ribeiro et. al.(2009) destacam ser necessário que se disponha de uma memória histórica de dados ocorridos anteriormente. Todavia, o conjunto de dados, por si só, não permite a previsão dos valores futuros. Para isso, é necessária a utilização de algoritmos, técnicas ou métodos de previsão de series temporais, que podem envolver cálculos simples ou procedimentos complexos.

2.6.1 Critérios para Avaliar a Adequação de Modelos a Séries Temporais

Pellegrini (2000) diz que vários modelos podem ser empregados na previsão de valores futuros da série temporal, depende do comportamento da série a qual se deseja analisar. A escolha do modelo mais apropriado é feita a partir do somatório dos erros gerados por cada

modelo ($e_t = z_t - \hat{z}_t$). Uma vez que o cálculo dos erros pode resultar em valores positivos e negativos, zerando assim o seu somatório, diferentes formas de cálculo para o somatório dos erros podem ser empregadas. Estas diferentes formas de cálculo constituem-se em critérios para escolha de modelos mais apropriados a séries temporais. Os critérios mais utilizados são:

- Média do quadrado dos erros (MQE): $\frac{1}{n} \sum_{t=1}^n e_t^2$
- Média absoluta dos erros (MAE): $\frac{1}{n} \sum_{t=1}^n |e_t|$
- Média absoluta percentual dos erros (MAPE): $\frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{z_t} \times 100 \right|$

Portanto, usando um dos critérios de cálculo, o modelo adequado será aquele que tiver o menor erro associado.

2.6.2 Modelos de Previsão

Existem dois enfoques utilizados na análise de séries temporais. Em ambos, o objetivo é construir modelos para estas séries. No primeiro enfoque, a análise é feita no domínio temporal e os modelos propostos são modelos paramétricos (com um número finito de parâmetros). No segundo, a análise é conduzida no domínio de frequências e os modelos propostos são modelos não-paramétricos (BEZERRA, 2006). Dentre os modelos paramétricos temos, por exemplo, os modelos ARMA, Erro ou Regressão e ARIMA. No domínio de frequências temos a análise espectral, que tem inúmeras aplicações em ciências físicas e engenharia, principalmente na engenharia elétrica, e que consiste em decompor a série dada em componentes de frequências e onde a existência do espectro é a característica fundamental (BEZERRA, 2006).

Existem diversos métodos para auxiliar na tarefa de previsão de séries temporais, como por exemplo: modelos de Suavização Exponencial, modelos auto-regressivos (AR), de medias moveis (MA) e Modelos ARIMA. Tecnologias de inteligência computacional, tais como redes neurais, lógica nebulosa e algoritmos genéticos, proporcionam a criação de metodologias avançadas de previsão (RIBEIRO et. al, 2009).

2.6.2.1 Modelo ARMA

É um modelo utilizado para ajustar funções com base em valores passados e na média da série. O nome ARMA vem do inglês autoregressive moving models que significa modelo autoregressivo de médias móveis.

Um modelo ARMA é definido por parâmetros p e q , o primeiro representa o número de termos autoregressivos e o segundo a ordem de médias móveis.

O ARMA nos dá os seguintes operadores para manipulação dos modelos.

- Operador translação para o passado (B)

$$BZ_t = Z_{t-1}, B^m Z_t = Z_{t-m} \quad (2.23)$$

- Operador translação para o futuro (F)

$$FZ_t = Z_{t+1}, F^m Z_t = Z_{t+m} \quad (2.24)$$

- Operador diferença (Δ)

$$\Delta Z_t = Z_t - Z_{t-1} = (1 - B)Z_t \quad (2.25)$$

- Operador soma (S)

$$SZ_t = \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + \dots = (1 + B + B^2 + \dots)Z_t \quad (2.26)$$

do que segue

$$SZ_t = (1 - B)^{-1}Z_t = \Delta^{-1}Z_t \quad (2.27)$$

isto é

$$S = \Delta^{-1} \quad (2.28)$$

2.6.2.2 Modelo de Regressão

Nos modelos usuais de regressão, uma variável dependente é estimada através de uma ou mais variáveis independentes, acrescidos de um termo de resíduo. Supõe-se, via de regra, resíduos normalmente distribuídos, com média zero, variância σ_a^2 e componentes não correlacionados.

De acordo com Morettin e Tolo (1987) o método de regressão é um modelo paramétrico em que existe uma função do tempo $f(t)$ completamente determinada e a_t é uma sequência aleatória, independente de $f(t)$. Qualquer influência do tempo age apenas na parte determinística $f(t)$ e modelos, onde Z_t depende funcionalmente de Z_{t-1}, Z_{t-2}, \dots , não estão incluídos em

$$Z_t = f(t) + a_t, t = 1, \dots, N, \quad (2.29)$$

em que $f(t)$ é chamada sinal e a_t o ruído.

A série a_t satisfaz

$$E(a_t) = 0, E(a_t^2) = \sigma_t^2, E(a_t a_s) = 0 \text{ } t \neq s \quad (2.30)$$

é chamada de ruído branco.

Sendo que neste modelo as observações são não correlacionadas, isto é, não há medição dos graus de associação entre duas séries temporais distintas.

O método de estatística clássica da análise de regressão pode ser normalmente usado para estimar os parâmetros de modelos comuns de tendência não constante (Cryer e Chan, 2008).

2.6.2.3 Modelo ARIMA

O modelo ARIMA é um modelo auto-regressivo, aplicável a série não estacionária. Tal metodologia consiste em ajustar modelos auto-regressivos integrados de médias móveis - ARIMA (CAMPOS, 2008).

Tendo em vista, duas classes de processos descritos pelo modelo ARIMA

i) Processo Linear estacionário

$$Z_t - \mu = a_t + \varphi_1 a_{t-1} + \varphi_2 a_{t-2} + \dots = \sum_{k=0}^{\infty} \varphi_k a_{t-k}, \varphi_0 = 1 \quad (2.31)$$

Em que

a_t é o ruído branco e $\mu = E(Z_t)$

$\varphi_1, \varphi_2, \varphi_3, \dots$ é uma seqüência de parâmetros tal que

$$\sum_{k=0}^{\infty} \varphi_k^2 < \infty \quad (2.32)$$

ii) Processo Linear não estacionário homogêneo

Esses modelos são descritos adequadamente pelos modelos auto-regressivos de médias móveis de ordem p, d e q : ARIMA (p, d, q) , que podem ser generalizados pela inclusão de um operador sazonal.

Morettin e Toloi(1987) enunciam tal processo como uma generalização dos processos lineares estacionários. Como foi descrito anteriormente supõe-se que a série Z_t seja não

estacionária em nível e/ou em inclinação, tornando-se estacionária a partir de finitas diferenças, d .

Caso

$$W_t = \Delta^d Z_t \quad (2.33)$$

seja estacionária, W_t é um modelo ARMA(p, q), isto é,

$$\phi(B)W_t = \theta(B)a_t \quad (2.34)$$

Considere W_t uma diferença de Z_t , logo Z_t é uma integral de W_t , por consequência Z_t segue um modelo auto-regressivo-integrado-médias móveis, ou modelo ARIMA,

$$\phi(B)\Delta^d Z_t = \theta(B)a_t \quad (2.35)$$

de ordem (p, d, q) , em que p e q são as ordens de $\phi(B)$ e $\theta(B)$, respectivamente.

2.6.2.3.1 Estruturação do Modelo ARIMA A estrutura do modelo ARIMA (Autoregressive Integrated Moving Average) é baseada nos estágios do ciclo iterativo do método de Box e Jenkins.

Os modelos de Box-Jenkins partem da idéia de que os valores de uma série temporal são altamente dependentes, ou seja, cada valor pode ser explicado por valores prévios da série. Os modelos ARIMA representam a classe mais geral de modelos para a análise de séries temporais (PELLEGRINI, 2000).

Bortolini et. al. (2011) descrevem os estágios de ciclos iterativos como: *i*) especificação de uma classe geral de modelos; *ii*) identificação de modelos com base na análise de autocorrelações, autocorrelações parciais e outros critérios; *iii*) estimação dos parâmetros do modelo identificado; *iv*) verificação da estrutura residual (ruído branco). Se o modelo não for adequado refaz-se o ciclo, retornando a fase de identificação.

A fase que requer mais atenção é a identificação do particular modelo ARIMA a ser ajustado aos dados. Como essa fase é baseada principalmente nas autocorrelações e autocorrelações parciais, seguem-se suas definições.

Autocorrelação: é uma medida que informa o quanto o valor de uma realização de uma variável aleatória é capaz de influenciar seus vizinhos, definida pela distância, ou atraso com que se deseja medi-la.

A autocorelação para Pellegrini (2000) pode ser usada para descrever a correlação entre dois valores da mesma série temporal, em diferentes períodos de tempo. Assim, um

coeficiente de autocorrelação ρ_1 mede a correlação entre dois valores adjacentes na série, e a autocorrelação, neste caso, é dita autocorrelação de lag (ou defasagem).

A função de autocorrelação (fac) ρ_j é estimada por:

$$r_j = \frac{c_j}{c_0}, j = 0, 1, \dots, N - 1 \quad (2.36)$$

em que c_j é a estimativa da cov Y_j .

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})], j = 0, 1, \dots, N - 1 \quad (2.37)$$

sendo $\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t$ a média amostral e $r_{-j} = r_j$.

Auto Correlação Parcial: a autocorrelação parcial é bem similar a autocorrelação, a respeito de permitir a análise de uma série temporal a partir do relacionamento entre os seus valores, por outro lado a autocorrelação parcial mede o grau de associação entre z_t e z_{t-k} , isso acontece quando os efeitos de outros lags são removidos.

Denardin (2009) destaca que a análise da Função Auto-Correlação Parcial (FACP) visa medir o grau de relacionamento que existe entre Z_t e Z_{t-k} , mantidos constantes o efeito dos demais Z 's intermediários do modelo Auto Regressivo. Isto é, procura-se mostrar o grau de relacionamento que existe entre as componentes do par Z_t e Z_{t-1} , mantendo-se fixos os efeitos de $Z_{t-2}, Z_{t-3}, \dots, Z_{t-k}$, ou mostrar o grau de relacionamento entre Z_t e Z_{t-2} , mantendo-se fixo o efeito de $Z_{t-1}, Z_{t-3}, \dots, Z_{t-k}$, e assim sucessivamente.

Prado (2011) destaca que Box, Jenkins e Reinsel (1994) propuseram a utilização da função de autocorrelação parcial (facp) como instrumento de facilitação do procedimento de identificação de um modelo, sendo útil também na estimação do modelo escolhido.

O objetivo da fase de identificação é determinar os valores de p , d e q do modelo ARIMA (p, d, q) , além de estimativas preliminares dos parâmetros a serem usados no estágio de estimação. Tal procedimento consiste em duas partes, a primeira é fazer diferenças quantas vezes necessárias das séries para se obter uma série estacionária, simplificando o processo em um ARMA (p, q) ; o segundo identifica o processo ARMA (p, q) resultante por meio das autocorrelações e autocorrelações parciais estimadas.

2.6.2.4 Modelos Sazonais

Bezerra (2006) classifica os modelos sazonais como determinístico e estocástico este último é conhecido como modelo de Box e Jenkins, para séries sazonais, denominado de SARIMA, este é uma generalização do modelo ARIMA para lidar com a sazonalidade, também chamado de ARIMA sazonal multiplicativo. Apesar da eliminação da sazonalidade determinística, ainda é possível que reste correlação significativa em:

i) "lags" de baixa ordem, indicando que os resíduos ainda são correlacionados, podendo-se ajustá-los através de modelo ARIMA, por exemplo;

ii) "lags" sazonais, isto é, múltiplos de período s . isto significa que há necessidade de se considerar uma sazonalidade estocástica, ou seja, ajustar a série original um modelo ARIMA sazonal (SARIMA).

Em grande parte das séries temporais existe a presença da componente periódica sazonal, esta se repete a cada s ($s > 1$) observações. O exemplo disso são as séries compostas por observações mensais e sazonalidade anual, s é igual a 12 (PRADO, 2011).

Define-se $\nabla_s = (1 - B^s)$ como sendo o operador de diferença sazonal. Assim,

$$\nabla_s Z_t = (1 - B^s)Z_t = Z_t - Z_{t-s} \quad (2.38)$$

é a primeira diferenciação sazonal. Em geral, D diferenciações sazonais podem ser requeridas para produzir uma série estacionária. Neste caso, o operador de diferenciação sazonal de ordem D é $\nabla_s^D = (1 - B^s)^D$. Assim, a forma geral do modelo sazonal autoregressivo integrado a média móvel de ordem (P, D, Q) (BOX et al., 1994 citado por PELLEGRINI, 2000).

$$\Phi(B^s)\nabla_s^D z = \Theta(B_s)\alpha_t \quad (2.39)$$

Em que $\Phi(B_s)$ e $\Theta(B_s)$ são polinômios em B^s de graus P e Q , respectivamente, satisfazem as condições de estacionariedade e invertibilidade. A representação da ordem (P, D, Q) é feita em letras maiúsculas, para diferenciá-la da representação feita nos modelos não sazonais. Logo o modelo Arima sazonal é simbolizado por SARIMA (p, d, q) (P, D, Q) .

O modelo SARIMA $(p, d, q)(P, D, Q)$, conforme Bortolini et. al. (2011) considera a tendência e a sazonalidade da série, e pode ser representado da seguinte forma:

$$\phi_p(B)\Phi_p(B^s)^d(1 - D)^D Z_t = \theta_q(B)\Theta_Q(B^s)a_t \quad (2.40)$$

Sendo que:

B é o operador Translação para o passado, definido por $B^j Z_t = Z_{t-j}$;

$\Phi_p(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_p B^p$ é o polinômio auto-regressivo de ordem p ;

$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ é o polinômio de médias móveis de ordem q ;

$\phi_P(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_P B^{Ps}$ é o polinômio sazonal autoregressivo de ordem P ;

$\Theta_Q(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_Q B^{Qs}$ é o polinômio sazonal de médias móveis de ordem Q ;

d é o número de diferenças necessárias para retirar a tendência da série;

D é o número de diferenças de "lag's" para retirar a sazonalidade da série.

Para identificar um modelo SARIMA $(p, d, q)(P, D, Q)_S$, inicialmente deve-se determinar os valores de d e D através da análise gráfica das funções de autocorrelação e de autocorrelação parcial. Se for observado um comportamento não-estacionário, com persistência de elevados valores para os coeficientes de autocorrelação e de autocorrelação parciais, pode-se inferir a necessidade de diferenciação da série (LAMOUNIER, 2007).

Da mesma forma, se valores elevados e persistentes nos lags com defasagens próximas aos múltiplos do período sazonal (próximas a 12, 24 e 36, por exemplo) pode-se inferir também a necessidade de diferenciação sazonal da série.

2.6.2.4.1 Sazonalidade Determinística A sazonalidade é determinística quando o padrão de repetição anual é exato, podendo ser predita perfeitamente a partir de meses anteriores.

No momento em que Z_t apresenta um comportamento sazonal determinístico com período 12, um modelo que pode ser útil segundo Morettin (1987) é:

$$Z_t = \mu_t + N_t \quad (2.41)$$

2.6.2.4.2 Sazonalidade Estocástica A sazonalidade estocástica pode ser encontrada em séries temporais cuja componente sazonal varia com o tempo (MORETTIN e TOLOI, 1987). Considere μ_t em (2.29) como um processo estocástico satisfazendo:

$$(1 - B^{12})\mu_t = Y_t \quad (2.42)$$

em que Y_t é o processo estacionário.

Empregando o operador $(1 - B^{12})$ na equação (2.29) obtém-se

$$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B)N_t \quad (2.43)$$

Substituindo a equação (2.30) em (2.31), tem-se

$$(1 - B^{12})Z_t = Y_t + (1 - B^{12})N_t$$

$$(1 - B^{12})Z_t = (1 - B^{12})\mu_t + (1 - B)N_t \quad (2.44)$$

Capítulo 3

Resultados

Neste capítulo são apresentados os resultados da aplicação dos modelos de regressão linear múltipla e $SARIMA(p, d, q)(P, D, Q)_s$. O primeiro parte da análise gráfica para identificar as componentes tendência e sazonalidade na série. No segundo foram ajustados três modelos para previsões.

3.1 Aplicação do Modelo de Regressão Múltipla

As séries de consumo mensal de energia para os quatro setores foram representadas graficamente na figura 3.1.

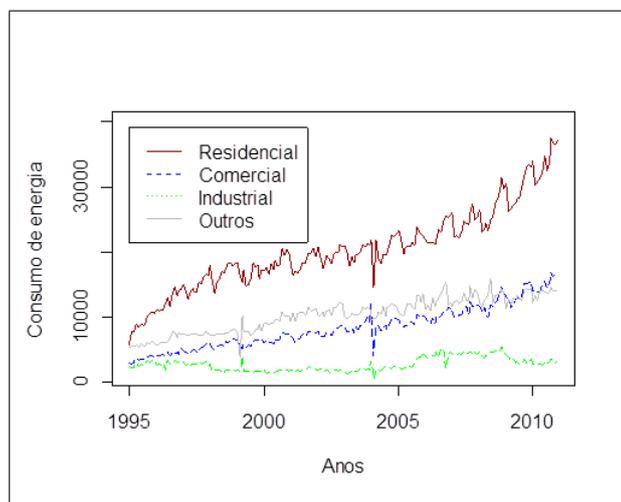


Figura 3.1: Gráfico de séries temporais para o consumo de energia nos diferentes setores.

De acordo com a figura 3.1, pode-se interpretar o comportamento do consumo mensal de energia no Estado do Amapá para cada um dos setores pesquisados. No setor residencial, é possível verificar indícios de que dois componentes determinem o comportamento da série, sendo eles a sazonalidade e a tendência. A tendência foi verificada graficamente devido ao crescimento constante no valor do consumo de energia com o passar dos anos,

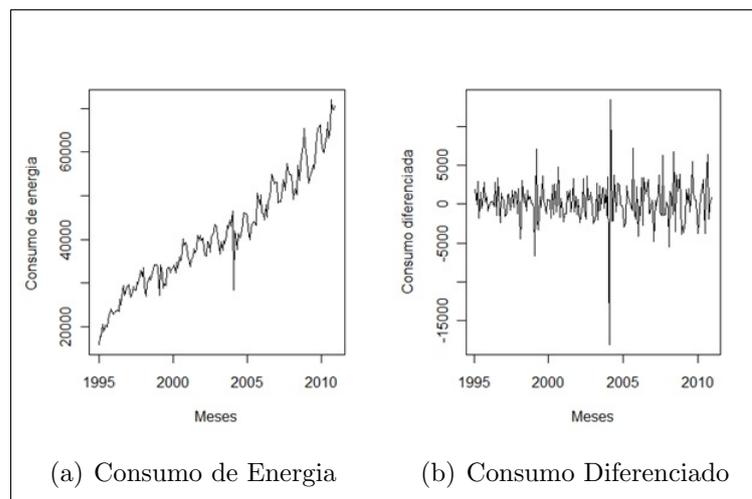


Figura 3.2: Gráfico de linhas da série de consumo total e da série diferenciada.

e a sazonalidade devido ao comportamento repetido de tempos em tempos ao longo de toda a série. No setor comercial houve um crescimento tendencial, exceto em 2004, que se verificou um decréscimo acentuado.

No setor industrial não se verificou graficamente indícios de tendência crescente ou decrescente no período considerado. Este setor apresentou o menor consumo em todo o período analisado, mantendo-se aproximadamente constante até o ano de 2005, quando apresentou um aumento de consumo e em seguida uma queda em 2010. Nos outros setores a série se manteve aproximadamente constante até 1999 quando apresentou uma baixa, seguida de um aumento com sazonalidade presente principalmente a partir de 2000.

A série de consumo total mensal foi representada na figura 3.2 (a), na qual verifica-se uma forte indicação de tendência crescente nos dados, e ainda uma evidência também da presença de um componente sazonal. Na figura 3.2(b) tem-se a série de consumo total com uma diferença para eliminar o componente de tendência. Verifica-se no gráfico que a média da nova série se manteve em torno de um valor constante ao longo de todo o período analisado, ou seja, a tendência foi completamente eliminada. No entanto, ainda é possível verificar a presença do componente sazonal.

A análise de regressão linear múltipla foi utilizada para verificar a significância estatística desses componentes já verificados graficamente.

Na tabela 3.1 tem-se que a maioria dos coeficientes da regressão, referentes aos meses (α_j), foram significativos, confirmando a significância de sazonalidade determinística nos dados. O coeficiente β_1 referente aos anos também se verificou significativo, concluindo-se assim, que realmente existe tendência na série.

De acordo com Fonseca, Martins e Toledo (1985) a tendência é identificada por movimentos persistentes em alguma direção, sendo na maioria dos casos possível descrever

Tabela 3.1: Parâmetros da análise de regressão múltipla da variável consumo de energia como função dos meses e dos anos.

Coefficiente	Estimativa	Erro	Padrão t	Prob > t
intercepto	23058,69	2040,266	11,30181	0,000000
β_1	0,952995	0,015939	59,78979	0,000000
α_1	-0,095817	0,045710	-2,09617	0,037481
α_2	-0,169537	0,045710	-3,70896	0,000278
α_3	-0,159599	0,045710	-3,49153	0,000606
α_4	-0,131978	0,045710	-2,88727	0,004368
α_5	-0,139024	0,045710	-3,04142	0,002711
α_6	-0,090500	0,045710	-1,97986	0,049260
α_7	-0,113280	0,045710	-2,47822	0,014136
α_8	-0,077822	0,045710	-1,70251	0,090406
α_9	-0,011162	0,045710	-0,24418	0,807370
α_{10}	-0,011826	0,045710	-0,25871	0,796156
α_{11}	-0,001317	0,045710	-0,02880	0,977054
α_{12}	0,007821	0,043408	0,18018	0,857213

tais variações por meio do ajustamento de uma reta ou função matemática. Assim na figura 3.3, a tendência da série foi representada graficamente, e quantificada na equação de regressão para os dados de consumo total anual de energia.

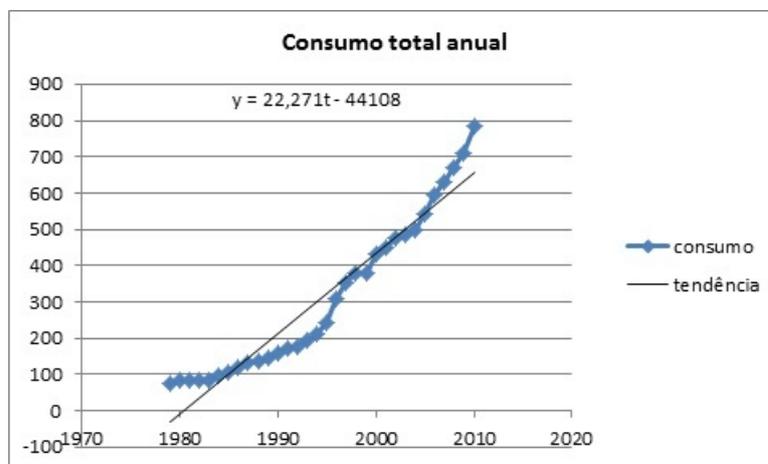


Figura 3.3: Gráfico da série de consumo anual (em 1000 MVH) com reta de tendência ajustada.

A equação de tendência que melhor se ajusta aos dados de consumo de energia é $y = 22,271t - 44108$ em que 22,271 é o coeficiente angular, t é a variável preditora e 44108 é o intercepto. Verificou-se um aumento significativo no consumo anual de energia elétrica ao longo do tempo, provavelmente em decorrência do aumento do número de consumidores.

Observa-se ainda uma dependência da sazonalidade em relação à tendência, pois a reta ajustada apresenta uma inclinação diferente de zero em relação ao eixo das abscissas

o que induz ao modelo multiplicativo. Daí decorre a utilização do modelo ARIMA sazonal multiplicativo (SARIMA) que será apresentado na próxima seção.

3.2 Ajuste dos Modelos SARIMA

Após a identificação dos componentes tendência e sazonalidade na série de consumo de energia, foram tomadas diferenças na série, para torná-la estacionária, e assim, ajustar um modelo para fins de previsão.

Considerando a presença de sazonalidade na série, utilizou-se o operador diferença para torná-la estacionária. Além disso, verificou-se na função de autocorrelação que ainda existiam alguns "lags" múltiplos de 12 significativos, sugerindo a presença da sazonalidade estocástica. Assim, o modelo a ser ajustado deve também incluir um componente sazonal. O modelo de Box e Jenkins da classe SARIMA(p, d, q)(P, D, Q), ou seja, um modelo auto-regressivo integrado de médias móveis sazonal, pode ser adequado para o ajuste. As ordens dos modelos identificados para essa série foram determinadas com base na análise das funções de autocorrelação e de autocorrelação parcial.

A verificação dos resíduos dos modelos foi feita com base na função de autocorrelação. O teste de Box-Pierce permite aceitar ou rejeitar a hipótese de que os resíduos do modelo são ruído branco, concluindo-se assim que o modelo ajusta-se bem aos dados. A estatística Q da função de autocorrelação é comparada com o valor crítico da distribuição qui-quadrado com graus de liberdade igual ao número de lags menos o número de parâmetros do modelo ajustado. Se os resíduos são um ruído branco, a maioria dos lags se encontrará dentro dos limites de confiança.

Os valores do teste estão apresentados na tabela a seguir, considerando a hipótese nula de que o resíduo é ruído branco.

Tabela 3.2: Parâmetros da análise de regressão múltipla da variável consumo de energia como função dos meses e dos anos.

Modelo	Estatística de teste Q	tabelado
SARIMA(1, 1, 1)(1, 1, 0) ₁₂	46,48	$\chi_{37}^2 = 52, 19$
SARIMA (0, 1, 2)(0, 1, 1) ₁₂	41,09	$\chi_{37}^2 = 52, 19$
SARIMA(0, 1, 1)(0, 1, 1) ₁₂	41,89	$\chi_{38}^2 = 53, 38$

De acordo com os valores da estatística de teste Q, a hipótese nula não é rejeitada para os três modelos, logo, conclui-se que os resíduos dos três são ruídos brancos. Assim, tem-se um bom ajuste dos três modelos.

Dessa forma, foram identificados três modelos para os dados, cujas funções de autocorrelação estão apresentadas a seguir.

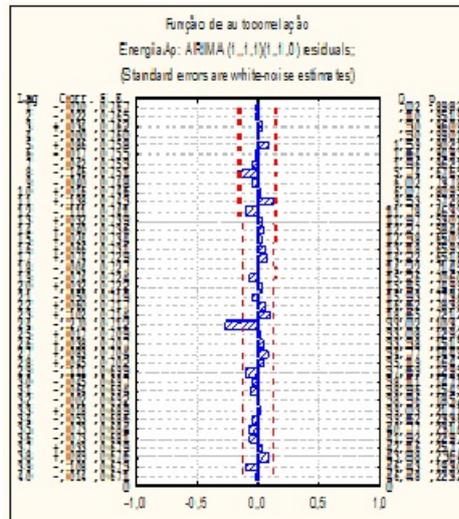


Figura 3.4: Função de autocorrelação para os resíduos dos modelos ajustados.

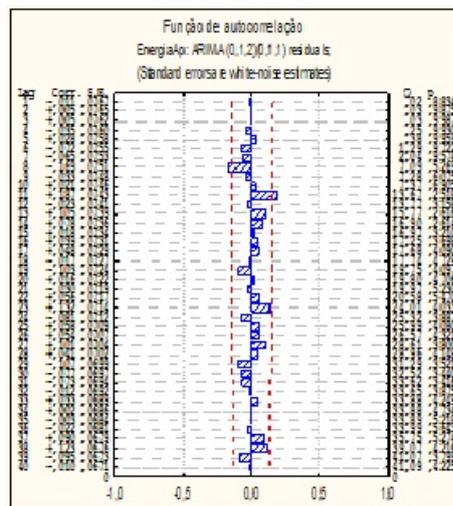


Figura 3.5: Função de autocorrelação para os resíduos dos modelos ajustados.

Pode-se verificar pelos gráficos que a maioria dos lags estão entre os limites de confiança, logo os resíduos dos modelos são ruído branco. As estimativas dos parâmetros dos modelos estão apresentadas na tabela 3.3.

Os modelos ajustados foram então utilizados para fazer previsões de consumo mensal para o ano de 2010, cujos valores observados foram reservados para comparações com essas previsões. As previsões estão apresentadas na tabela 3.4.

Considerando que o principal interesse no ajuste é a previsão de valores futuros, o modelo mais adequado é aquele que apresente as melhores previsões. O critério utilizado para a escolha do melhor modelo foi o do menor erro quadrático médio de previsão, com resultados apresentados na tabela 3.5.

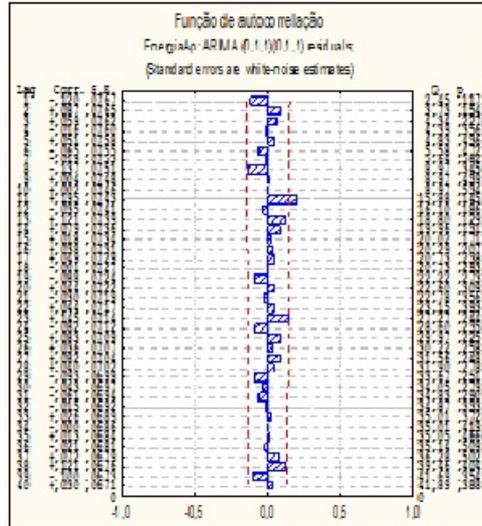


Figura 3.6: Função de autocorrelação para os resíduos dos modelos ajustados.

Tabela 3.3: Parâmetros da análise de regressão múltipla da variável consumo de energia como função dos meses e dos anos.

Estimativas dos parâmetros						
Modelo	Φ_1	θ_1	θ_2	Φ_2	Θ_1	Θ_2
SARIMA(1,1,1)(1,1,0)12	-0,25	0,67		-0,44		
SARIMA (0,1,2)(0,1,1)12		0,81	-0,17		0,68	
SARIMA(0,1,1)(0,1,1)12		0,70		0,69		

Tabela 3.4: Consumo total mensal, observado e previsto para o ano de 2010, com os modelos M1: SARIMA(1,1,1)(1,1,0)12; M2: SARIMA(0,1,2)(0,1,1)12; M3: SARIMA(0,1,1)(0,1,1)12.

Ano 2010		Valores previstos		
Mês	Valor observado	M1	M2	M3
Jan	62562	62025,60	63243,32	63190,55
Fev	60124	58214,50	59311,74	58844,02
Mar	59796	57588,51	59305,50	58896,25
Abr	61726	59095,44	60636,80	60217,95
Mai	63699	58586,34	60266,53	59836,36
Jun	66904	62696,52	63626,36	63152,86
Jul	63101	60630,56	62407,42	61961,16
Ago	65519	63459,87	64040,33	63564,42
Set	72000	67301,64	68256,26	67733,93
Out	70065	69251,82	68986,98	68419,28
Nov	69858	71226,98	69672,57	69077,79
Dez	70672	69832,82	69168,62	68586,48

De acordo com os valores do erro quadrático médio, o modelo mais adequado para fins de previsão é o modelo SARIMA (0,1,2)(0,1,1)12 . A expressão para esse modelo

Tabela 3.5: Valores do erro quadrático médio de previsão (MSE) para cada modelo.

Modelo	MSE
SARIMA(1,1,1)(1,1,0) ₁₂	7935378,85
SARIMA (0,1,2)(0,1,1) ₁₂	3767752,32
SARIMA(0,1,1)(0,1,1) ₁₂	5424157,47

considerando as estimativas dos parâmetros, é dada pela expressão:

$$(1 - B_{12})(1 - B)Z_t = (1 - 0,81B + 0,17B_2)(1 - 0,68B_2)a_t \quad (3.1)$$

As previsões feitas a partir desse modelo são visualizadas na figura 3.7 a seguir.

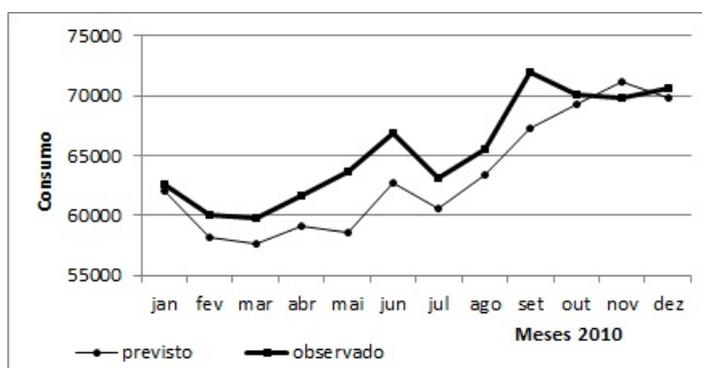


Figura 3.7: Comparações dos valores observados com os valores previstos para o ano de 2010 com o modelo SARIMA (0, 1, 2)(0, 1, 1)₁₂.

De acordo com o gráfico, pode-se verificar que as previsões para o consumo de energia em 2010 acompanharam as variações dos valores reais. No geral, as previsões foram subestimadas em relação aos valores reais.

Conclusão

O método de regressão múltipla, realizado para verificar a presença de tendência e sazonalidade foi significativo principalmente no setor residencial.

A série de consumo de energia elétrica do Estado do Amapá apresentou as componentes tendência e sazonalidade, constatando-se a presença de sazonalidade determinística.

O consumo de energia elétrica no Estado do Amapá apresenta um comportamento tendencioso crescente.

Os modelos da classe SARIMA ajustados permitiram obter previsões de consumo de energia com valores próximos dos valores reais observados.

Referências Bibliográficas

- [1] BEZERRA, Manoel Ivanildo Silvestre. **Apostila de Análise de Séries Temporais**, 2006. Disponível em <http://pt.scribd.com/doc/5515941/Apostila-Series-Temporais>; Acesso em: 22 set.2011.
- [2] CAMPOS, Roger Júnior. **Previsão de Séries Temporais com Aplicações a Séries de Consumo de Energia Elétrica**.2008. Disponível em: [http : //www.cpdee.ufmg.br/ roger/search.html](http://www.cpdee.ufmg.br/roger/search.html) > Acesso em: 19/08/ 2011.
- [3] FONSECA, Jairo Simon da; MARTINS, Gilberto de Andrade; TOLEDO, Geraldo Luciano **Estatística Aplicada**.2.ed.:Atlas,1985.
- [4] LAMOUNIER,Wagner Moura.**Tendência, ciclos e sazonalidade nos preços spot do café brasileiro na NYBOT**.Gest. Prod., São Carlos, v. 14, n. 1, p. 13-23, jan.-abr. 2007 .
- [5] MILONE, Giuseppe.**Estatística**:geral e aplicada. 2.ed. São Paulo: Thomson Learning, 2006.
- [6] MORETTIN, Pedro A.; TOLOI, Clélia M.**Séries Temporais**:Métodos Quantitativos. 2. ed. São Paulo: Atual,1987.
- [7] MORETTIN, Pedro Alberto; TOLOI, Clélia Maria**Previsão de Séries Temporais**. 2.ed. São Paulo: Atual,1987.
- [8] PELLEGRINI, Fernando Rezende.**Metodologia para Implementação de Sistemas de Previsão de Demanda**.2000. Disponível em:< [http : //www.producao.ufrgs.br/arquivos/publicacoes/Fernando%20R%20Pellegrini.pdf](http://www.producao.ufrgs.br/arquivos/publicacoes/Fernando%20R%20Pellegrini.pdf) > Acesso em: 06/09/ 2011.
- [9] RIBEIRO,Claudio Vasconcelos; GOLDSCHMIDT, Ronaldo Ribeiro; CHOREN, Ricardo. **Métodos para Previsão de Séries Temporais e suas Tendências de Desenvolvimento**. Disponível em < [http : //www.comp.ime.eb.br/techreports/repositorio/2009_3.pdf](http://www.comp.ime.eb.br/techreports/repositorio/2009_3.pdf) > Acesso em: 06/09/2011.

- [10] REIS, Allan Braune. **Previsão de Carga a Curto Prazo Via Modelagem Local**. 2005.
- [11] IQUIAPAZA, Roberto Aldo; BARBOSA, Francisco Vidal; BRESA, Aureliano Angel. **Teste de não Estacionariedade em Séries Financeiras com Dados em Painel: Uma Síntese Aplicada**. V Encontro Brasileiro de Finanças, Universidade Presbiteriana Mackenzie, julho 18, 2005 - julho 19, 2005.
- [12] RODRIGUES, Luciene Cleia; SILVA, Paula Pinho Cabral da; LINDEN Ricardo. **Séries Temporais no Consumo de Energia Elétrica no Estado do Rio de Janeiro**. Disponível em < [http : //pt.wikipedia.org/wiki/Autocorrela%C3%A7%C3%A3o](http://pt.wikipedia.org/wiki/Autocorrela%C3%A7%C3%A3o) > acessado em: 20/10/2011.