

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Thiago Gomes de Araujo

**Ajuste das Séries de Carga de Energia
Elétrica Influenciadas pelos Ofensores
Calendário e Temperatura**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio.

Orientador: Prof. Reinaldo Castro Souza

Rio de Janeiro
Abril de 2013



Thiago Gomes de Araujo

**Ajuste das Séries de Carga de Energia Elétrica
Influenciadas pelos Ofensores Calendário e
Temperatura**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Reinaldo Castro Souza

Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Nivalde José de Castro

UFRJ

Prof. José Francisco Moreira Pessanha

UERJ

Prof. Plutarcho Maravilha Lourenço

Centro de Pesquisas de Energia Elétrica

Prof. José Eugenio Leal

Coordenador Setorial do Centro

Técnico Científico

Rio de Janeiro, 03 de abril de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Thiago Gomes de Araujo

Graduou-se em Estatística na ENCE (Escola Nacional de Ciências Estatísticas) em 2010 com a monografia intitulada “Ajuste Sazonal do PIB Trimestral: X-12- Arima e Modelo Estrutural – Análise Comparativa dos Resultados”.

Ficha Catalográfica

Araujo, Thiago.

Ajuste das Séries de Carga de Energia Elétrica Influenciadas pelos Ofensores Calendário e Temperatura / Thiago Gomes de Araujo; orientador: Reinaldo Castro Souza. – 2013.

81 f.: il. ; 29,7 cm

1. Dissertação (Mestrado em Engenharia Elétrica) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.

Inclui referências bibliográficas.

1. Engenharia Elétrica – Teses. 2. Energia Elétrica. 3. Séries Temporais. 4. Regressão Linear. 5. Regressão Dinâmica. I. Souza, Reinaldo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Este trabalho é dedicado aos meus pais, Marco Antonio e Alzenir Angelo, por toda dedicação, apoio, compreensão e confiança.

Agradecimentos

A Deus por ter abençoado a minha vida e guiado meu caminho.

A CAPES e a PUC-Rio pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Ao meu orientador Prof. Dr. Reinaldo Castro Souza pela ajuda e confiança depositada.

Ao Pedro Ferreira e Fernando Cyrino pela ajuda, sugestão de melhorias e novas ideias em toda a elaboração deste trabalho.

Aos meus pais, Marco e Zeni, e irmão, Igor Araujo, pela ajuda nas horas de dificuldade.

À minha família e meus amigos, em especial Vinícius Pereira, Thiago Ismério, Paulo Ricardo e Carolina Portilho, pela torcida e pelas orações.

Aos professores pelos ensinamentos.

Ao meu amigo Julio Siqueira pela amizade e ajuda na realização deste trabalho.

Resumo

Araujo, Thiago Gomes de; Souza, Reinaldo Castro (Orientador). **Ajuste das Séries de Carga de Energia Elétrica Influenciadas pelos Ofensores Calendário e Temperatura**. Rio de Janeiro, 2013. 81p. Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

O objetivo do presente trabalho é a geração de uma série mensal de carga elétrica livre das variações de calendário e de temperatura. Para tal, foram comparadas duas abordagens, uma totalmente empírica e outra híbrida com métodos empíricos e modelagens de regressão dinâmica, para identificar a mais adequada para a retirada desses ofensores. Os dados utilizados são provenientes de observações diárias de cada um dos quatro subsistemas que integram o Sistema Interligado Nacional (SIN), porém a ideia é produzir séries mensais do SIN e não apenas de cada um dos subsistemas. A série trimestral do PIB foi utilizada para decidir qual abordagem melhor ajustou os dados de Carga. A série mensal de carga ajustada do SIN será utilizada para subsidiar decisões, de compra e venda de energia nos leilões, das empresas distribuidoras de energia elétrica.

Palavras-chave

Energia Elétrica, Séries Temporais; Regressão Linear; Regressão Dinâmica.

Abstract

Araujo, Thiago Gomes de; Souza, Reinaldo Castro (Advisor). **Adjusting Load Series by the Calendar and Temperature Effects**. Rio de Janeiro, 2013. 81p. MSc. Dissertation – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis proposes a method to generate monthly load series free of variations coming from two sources: calendar and temperature. Two approaches were considered, one totally empirical and another one called hybrid, as it use empirical procedure to remove the calendar effect and a dynamic regression type of model to remove the temperature effects. The data set used comes found to daily observations from each one of the four subsystems that form the SIN (Brazilian Integrated Grid). However the final task is to obtain a unique monthly series for the SIN and not only the four subsystems monthly series. The quarterly PIB series was used to check the performance of the two proposed methods. Such adjusted series are quite important tools to hold on the decision of acquisitions and dailes of energy in the energy audits.

Keywords

Electric Energy, Time Series; Linear Regression; Dynamic Regression.

Sumário

1 INTRODUÇÃO	13
2 REFERENCIAL TEÓRICO	21
2.1. Séries Temporais	21
2.2. Processos Estocásticos	21
2.3. Previsão de séries temporais	23
2.4. Regressão Dinâmica	24
2.5. Testes Estatísticos	28
3 METODOLOGIA	34
3.1. Método Empírico (Multiplicativo)	34
3.2. Regressão Dinâmica (Método Aditivo)	44
4 RESULTADOS	45
4.1. Análise Exploratória dos dados	46
4.2. Retirada do efeito Calendário	52
4.3. Retirada do efeito Temperatura	55
4.4. Retirada do efeito Semana Morta	59
4.5. Obtenção das séries Mensais	64
4.6. Obtenção da Série Mensal do Sistema Interligado Nacional (SIN)	66
4.7. Obtenção da Série Trimestral do SIN	66
4.8. Comparação dos Resultados	67
5 CONCLUSÃO E TRABALHOS FUTUROS	70
REFERÊNCIAS BIBLIOGRÁFICAS	72
APÊNDICE	74
ANEXO A MODELAGEM DA SÉRIE DIÁRIA DE CARGA VIA REGRESSÃO DINÂMICA	75

Lista de Gráficos

Gráfico 2.1: Fluxograma para construção de um modelo de Regressão Dinâmica	27
Gráfico 3.1: Fluxograma da metodologia empírica de limpeza da série diária de Carga até a série trimestral de Carga	43
Gráfico 4.1: Série de carga do subsistema Norte	47
Gráfico 4.2: Série de carga do subsistema Nordeste	47
Gráfico 4.3: Série de carga do subsistema Sul	47
Gráfico 4.4: Série de carga do subsistema Sudeste / Centro-Oeste	48
Gráfico 4.5: Monthplot do subsistema Norte	49
Gráfico 4.6: Monthplot do subsistema Nordeste	49
Gráfico 4.7: Monthplot do subsistema Sul	49
Gráfico 4.8: Monthplot do subsistema Sudeste / Centro-Oeste	50
Gráfico 4.9: Histograma do subsistema Norte	50
Gráfico 4.10: Histograma do subsistema Nordeste	51
Gráfico 4.11: Histograma do subsistema Sul	51
Gráfico 4.12: Histograma do subsistema Sudeste / Centro-Oeste	52
Gráfico 4.13: Série de Carga com e sem ajuste calendário do subsistema Norte	54
Gráfico 4.14: Série de Carga com e sem ajuste calendário do subsistema Nordeste	54
Gráfico 4.15: Série de Carga com e sem ajuste calendário do subsistema Sul	54
Gráfico 4.16: Série de Carga com e sem ajuste calendário do subsistema Sudeste / Centro-Oeste	55
Gráfico 4.17: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Nordeste	56
Gráfico 4.18: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sul	56
Gráfico 4.19: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sudeste / Centro-Oeste	57
Gráfico 4.20: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Nordeste	58

Gráfico 4.21: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sul	58
Gráfico 4.22: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sudeste / Centro-Oeste	59
Gráfico 4.23: Série de Carga com ajuste Calendário, Temperatura utilizando o Método Empírico e com e sem ajuste Semana Morta para o subsistema Nordeste	61
Gráfico 4.24: Série de Carga com ajuste Calendário, Temperatura utilizando o Método Empírico e com e sem ajuste Semana Morta para o subsistema Sul	61
Gráfico 4.25: Série de Carga com ajuste Calendário, Temperatura utilizando o Método Empírico e com e sem ajuste Semana Morta para o subsistema Sudeste / Centro-Oeste	62
Gráfico 4.26: Série de Carga com ajuste Calendário, Temperatura utilizando Regressão Dinâmica e com e sem ajuste Semana Morta para o subsistema Nordeste	62
Gráfico 4.27: Série de Carga com ajuste Calendário, Temperatura utilizando Regressão Dinâmica e com e sem ajuste Semana Morta para o subsistema Sul	63
Gráfico 4.28: Série de Carga com ajuste Calendário, Temperatura utilizando Regressão Dinâmica e com e sem ajuste Semana Morta para o subsistema Sudeste / Centro-Oeste	63
Gráfico 4.29: Série Mensal de Carga ajustada do subsistema Norte	64
Gráfico 4.30: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do subsistema Nordeste	64
Gráfico 4.31: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do subsistema Sul	65
Gráfico 4.32: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do subsistema Sudeste / Centro-Oeste	65
Gráfico 4.33: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do SIN	66
Gráfico 4.34: Série Trimensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do SIN	67
Gráfico 4.35: Séries Trimestrais de variação de Carga ajustada em relação ao mesmo trimestre do ano anterior	67
Gráfico 4.36: Séries Trimestrais de variação de Carga ajustada e PIB em relação ao mesmo trimestre do ano anterior	68

Gráfico A.1: Carga x Carga ajustada do subsistema Nordeste	76
Gráfico A.2: Carga x Carga ajustada do subsistema Sul	77
Gráfico A.3: Carga x Carga ajustada do subsistema Sudeste / Centro-Oeste	77
Gráfico A.4: Dispersão entre as variáveis carga ajustada e resíduos do subsistema Nordeste	78
Gráfico A.5: Dispersão entre as variáveis carga ajustada e resíduos do subsistema Sul	79
Gráfico A.6: Dispersão entre as variáveis carga ajustada e resíduos do subsistema Sudeste / Centro-Oeste	79
Gráfico A.7: FAC dos resíduos para o subsistema Nordeste	80
Gráfico A.8: FAC dos resíduos para o subsistema Sul	80
Gráfico A.9: FAC dos resíduos para o subsistema Sudeste / Centro- Oeste	81

Lista de Tabelas

Tabela 4.1: Estatísticas das Séries de carga dos subsistemas	48
Tabela 4.2: Teste de Jarque-Bera para os quatro subsistemas	52
Tabela 4.3: Média de Carga com e sem ajuste calendário por dia da semana	53
Tabela 4.4: Coeficientes de regressão de cada período de cada subsistema por mês	56
Tabela 4.5: Média de Carga com ajuste Calendário e Temperatura utilizando o Método Empírico, para uma semana comum e uma semana de festas	60
Tabela 4.6: Média de Carga com ajuste Calendário e Temperatura utilizando Regressão Dinâmica, para uma semana comum e uma semana de festas	60
Tabela 4.7: Correlação entre as séries trimestrais de variação de Carga com a série trimestral de variação do PIB	69
Tabela Ap.1: Valores de R^2 para diferentes tamanhos de janela	74
Tabela A.1: MAPE e R^2 dos modelos	76
Tabela A.2: Teste para verificar média zero nas séries	77
Tabela A.3: Teste para verificar estacionariedade das séries	78
Tabela A.4: Teste de Ljung-Box para os resíduos dos modelos.	81

1 Introdução

A energia elétrica que alimenta as indústrias, comércio e lares no Brasil é gerada (ou produzida), principalmente, em usinas hidrelétricas (cerca de 80%), onde a passagem da água por turbinas geradoras transformam a energia mecânica, originada pela queda d'água, em energia elétrica com um baixo custo. Esta forma de geração de energia é considerada limpa, pois não há queima de combustível, além disso a água é um recurso renovável. A geração de energia através de usinas hidrelétricas é útil em países com grande vazão hidrográfica (rios e bacias) e acidentes geográficos (quedas d'água, por exemplo). Além das vantagens citadas, o aproveitamento hidrelétrico proporciona outros usos tais como irrigação, navegação e amortecimentos de cheias.

Para que a energia elétrica chegue até os consumidores finais, além da geração, é necessário ainda que faça um percurso que envolva os sistemas de transmissão e distribuição, que serão explicadas mais adiante.

Segundo Reis (2003), antes de chegar aos locais de consumo, a energia elétrica passa pelo estágio da distribuição – a não ser em casos muito especiais de grande consumidores em alta tensão, alimentados diretamente pela transmissão.

Em todos os segmentos da sociedade moderna, a energia elétrica é um insumo de grande importância, pois ela viabiliza, dentre outros, atividades industriais de grande porte (complexos siderúrgicos, por exemplo) e hábitos cotidianos mais simples (iluminação residencial, por exemplo).

As tecnologias e processos da distribuição são muito diferentes dos realizados na transmissão, essa diferença caracteriza empresas e processos de gestão diferentes, associados a cadeias de negócios totalmente diversas.

O setor elétrico brasileiro, ao contrário do que ocorre em outros países, ainda não possui um mercado spot eficiente, que cumpre algumas funções importantes numa indústria desregulamentada, tais como: aumentar a flexibilidade das transações; permitir ajustes entre a energia contratada e a energia gerada e ser

referência para contratos de longo prazo. Isto é, um mercado spot é um importante mecanismo de ajuste entre demanda e oferta. (Newbery, 1998).

Iniciada nos anos 90, a reforma do setor elétrico brasileiro, seguindo uma tendência verificada em inúmeros países, teve como objetivo central introduzir um ambiente de competição e aumentar a participação privada nos investimentos do setor, na medida em que se constatava o fim do modelo de financiamento público. Visando estimular novos investimentos, o propósito inicial da reforma do setor elétrico brasileiro era a criação de um sistema no qual a competição e os preços definidos no mercado spot fossem estimulados – da geração à distribuição.

A separação dos três segmentos que compõem a cadeia de produção: geração, transmissão e distribuição, chamado de processo de desverticalização das empresas, é uma característica importante da reforma da indústria de eletricidade. A prevenção do comportamento predatório e o aumento de competidores no segmento da geração, dado o livre acesso à rede, era o objetivo da desverticalização. (Leite & Santana, 2006).

Porém, a reforma proposta apresentou falhas significativas, tanto de planejamento como de execução. O cronograma das reformas e do processo de privatização sofreu atrasos. No que diz respeito às privatizações, elas não tiveram a aprovação da sociedade sobre sua real necessidade, como Peci (2007) relata, e foram acompanhadas de inconsistências significativas nas novas regras contratuais. Potenciais fontes de conflitos de interesses entre as geradoras e distribuidoras foram encerrados com aspectos referentes às cláusulas de contratos de concessões e aos contratos iniciais entre os agentes (Correia *et al.*, 2006).

É importante observar que no período pós-acionamento, as empresas concessionárias passaram a enfrentar séria crise de liquidez e prejuízos operacionais (Pires *et al.*, 2002). Com isso, foi iniciado no Brasil, a partir de 2003, um processo de ajuste e reestruturação do setor elétrico, resultante de uma nova estrutura de poder, decidido no processo eleitoral. Dessa forma, no âmbito do Ministério das Minas e Energia (MME), foi iniciado uma série de estudos para formular e implementar um novo modelo para o setor elétrico brasileiro, cujas bases institucionais e legais foram aprovadas pelo Congresso Nacional através das Leis 10.847 e 10.848 de 2004. Este novo modelo tem dois objetivos: a garantia de suprimento de eletricidade com modicidade tarifária.

Com relação à segurança do suprimento, o novo e atual modelo brasileiro incorporou:

a) A inversão do foco dos contratos de energia elétrica do curto para o longo prazo, com o objetivo de reduzir a volatilidade do preço e criar um mercado de contratos de longo prazo (Power Purchase Agreements – PPAs) que possa ser utilizado como garantia firme para os financiamentos;

b) A obrigatoriedade de cobertura contratual, pelas distribuidoras e consumidores livres, de 100% de seu consumo de energia elétrica;

c) Criação do Comitê de Monitoramento de Setor Elétrico (CMSE) com a função de acompanhar para os horizontes temporais de curto, médio e longo prazo o equilíbrio entre oferta e demanda;

d) A exigência prévia de licenças ambientais para se permitir a participação de um novo empreendimento no processo de licitação;

e) A retomada do planejamento setorial integrado e centralizado pelo Estado, na figura da Empresa de Pesquisa Energética (EPE), criada pelo novo decreto.

O mercado brasileiro de energia elétrica foi dividido pelo modelo em dois ambientes de comercialização, com lógicas e estruturas nitidamente distintas. O primeiro, denominado Ambiente de Contratação Regulada (ACR), atende as demandas dos consumidores cativos, onde prevalecem os consumidores residenciais, de serviço e indústria com níveis de consumo menores. O segundo, voltado exclusivamente para empresas com maior volume de consumo e necessidades estratégicas de maior volume de energia em curto prazo, chamados de consumidores livres, é denominado de Ambiente de Contratação Livre (ACL) e busca assegurar a concorrência e a liberdade efetiva dos seus participantes. No ACL, com um mínimo de intervenção do governo, os agentes podem realizar, livremente, contratos bilaterais, definindo-se preços, volumes, prazos e cláusulas de hedge.

Uma importante característica do modelo é que a integração vertical das empresas não é mais permitida. As distribuidoras não podem ser proprietárias de ativos de geração, nem comercializar diretamente com consumidores livres. Isto é, elas podem atuar apenas no ACR, comprando por meio de leilões e vendendo para consumidores cativos. Essa característica do modelo é uma vantagem, porque reduz os prováveis e indesejáveis subsídios cruzados, em que as tarifas dos

consumidores cativos poderiam subsidiar menores preços para os consumidores livres, como relata Castro & Leite (2008).

No contexto descrito até aqui, vale dizer que a função de comprar grandes blocos de energia das supridoras é das empresas de distribuição de energia elétrica. Cabe a elas também o controle do nível de tensão em patamares próprios para o consumo de sua clientela, geralmente formada por milhares de consumidores, além de manter a rede de distribuição e as instalações técnicas operando adequadamente.

Segundo Araújo (2007), a Câmara de Comercialização de Energia Elétrica (CCEE) iniciou seu trabalho em 2004, sucedendo ao Mercado Atacadista de Energia (MAE). A CCEE possui papel estratégico para viabilizar as operações de compra e venda de energia elétrica, registrando e administrando contratos firmados entre geradores, comercializadores, distribuidores e consumidores livres.

A CCEE viabiliza a comercialização de energia elétrica no Sistema Interligado Nacional nos Ambientes de Contratação Regulada e Contratação Livre, além de efetuar a contabilização e a liquidação financeira das operações realizadas no mercado de curto prazo. Os procedimentos de comercialização utilizados pela CCEE provém da ANEEL. As principais atribuições do CCEE são:

- Manter o registro de todos os contratos fechados nos Ambientes de Contratação Regulada (ACR) e de Contratação Livre (ACL);
- Promover a medição e registro dos dados de geração e consumo de todos os Agentes da CCEE;
- Apurar o Preço de Liquidação de Diferenças (PLD) do Mercado de Curto Prazo por submercado;
- Efetuar a Contabilização dos montantes de energia elétrica comercializados no Mercado de Curto Prazo e a Liquidação Financeira;
- Apurar o descumprimento de limites de contratação de energia elétrica e outras infrações e, quando for o caso, por delegação da ANEEL, nos termos da Convenção de Comercialização, aplicar as respectivas penalidades;
- Apurar os montantes e promover as ações necessárias para a realização do depósito, da custódia e da execução de Garantias Financeiras, relativas às Liquidações Financeiras do Mercado de Curto Prazo, nos termos da Convenção de Comercialização;

- Promover Leilões de Compra e Venda de energia elétrica, conforme delegação da ANEEL;
- Promover o monitoramento das ações empreendidas pelos Agentes, no âmbito da CCEE, visando à verificação de sua conformidade com as Regras e Procedimentos de Comercialização, e com outras disposições regulatórias, conforme definido pela ANEEL.

Ainda segundo Araujo (2007), de acordo com o estabelecido pelo artigo 11 do Decreto nº 5.163/2004 e artigo 2º da Lei nº 10.848/2004, é por meio de licitação na modalidade de leilões que as concessionárias, as permissionárias e as autorizadas de serviço público de Distribuição de energia elétrica do Sistema Interligado Nacional (SIN) devem garantir o atendimento a todo o seu mercado no Ambiente de Contratação Regulada (ACR).

A regulação das licitações para contratação regulada de energia elétrica e a realização do leilão diretamente ou por intermédio da Câmara de Comercialização de Energia Elétrica (CCEE) é atribuição da ANEEL, conforme determinado no parágrafo 11 do artigo 2º da Lei nº 10.848/2004.

Os vencedores de um leilão serão aqueles que ofertarem energia elétrica pelo menor preço por Mega-Watt (MWh) hora para atendimento da demanda prevista pelas Distribuidoras, isto é, o critério de menor tarifa (inciso VII, do art. 20, do Decreto nº 5.163/2004) será utilizado para definir os vencedores do leilão. Dessa forma, os contratos de Comercialização de Energia Elétrica em Ambiente Regulado (CCEAR) serão celebrados entre os vencedores e as Distribuidoras que declararam necessidade de compra para o início de suprimento da energia contratada no leilão (Reis, 2003).

Entretanto para que as empresas Distribuidoras façam boas previsões de Carga de Energia Elétrica a ser ofertada, de forma que o abastecimento do país seja garantido, é necessário que as séries de Carga estejam bem tratadas e livres de ofensores não econômicos, tais como Calendário e Temperatura. Além disto, será demonstrado aqui neste trabalho a alta correlação entre a variação de Carga ajustada de um trimestre para o mesmo trimestre do ano anterior com a variação do Produto Interno Bruto (PIB) trimestral com o mesmo trimestre do ano anterior. Dessa forma, uma boa previsão da série de Carga trimestral livre dos ofensores Calendário e Temperatura pode ser utilizada para a previsão do PIB trimestral.

Dada à importância da geração de séries de Carga livre dos ofensores não econômicos, o principal objetivo deste trabalho é produzir séries mensais e trimestrais de Carga livres dos ofensores Calendário e Temperatura.

Este ajuste para retirada dos ofensores não econômicos de Temperatura e Calendário das séries de Carga é feito em alguns países do mundo.

Na Espanha, a Rede Eléctrica de España (2012) divulga em seu *Boletín Mensual* os dados das medições físicas de carga e os dados corrigidos para compensar efeitos de temperaturas atípicas e diferenças no número de dias úteis. As séries de dados são disponíveis em uma frequência mensal e separam a variação da carga que foi ocasionada pela temperatura, pelas diferenças de calendário e o que decorre da atividade econômica.

Em Portugal é feito um trabalho parecido, onde a Rede Eléctrica Nacional (2012) divulga através da *Informação Mensal do Sistema Eletroprodutor* a medição física da carga e os ajustes para os efeitos da temperatura e do número de dias úteis.

No Brasil, Castro *et al.* (2012) desenvolveram uma metodologia adequada às especificidades do Brasil para produção de dados ajustados de carga, capaz de isolar a influência de fatores não econômicos na carga e listaram os fatores não econômicos que influenciam a carga no Brasil e que podem ser estimados com base em dados disponíveis. São eles:

- Descontinuidades de medição: ocorrem quando, por exemplo, há a incorporação de um sistema isolado ao Sistema Interligado Nacional (SIN). Em outubro de 2009, o antigo sistema isolado de Acre-Rondônia foi conectado ao SIN. A partir de meados desse mês a carga destes dois estados, que à época somava cerca de 400 MWmed, passou a ser computada no SIN. Claro está que este acréscimo de carga, da ordem de 0,75% da carga do SIN de um mês para o outro, não se deveu a nenhum crescimento econômico, mas é fruto do simples fato de que o ONS passou a medir a carga de uma parcela maior do país.

- Perdas da rede de transmissão: podem oscilar bastante ao longo do ano por razões estritamente ligadas à operação do sistema. Embora as perdas na transmissão tenham como fundamento razões elétricas e não econômicas, na maior parte dos países as perdas na transmissão tendem a ser pequenas e a variar pouco – e esta é provavelmente a razão pela qual o ajuste das perdas não é feito no tratamento de dados de carga em outros países. No Brasil é diferente: o sistema de

transmissão brasileiro é muito extenso e transporta grandes volumes de energia, pois foi construído para aproveitar as diferenças nos regimes hidrológicos entre as diversas bacias com aproveitamentos hidroelétricos. O ONS se vale do sistema de transmissão para transportar energia de onde ela é abundante no momento para onde ela é escassa. O uso mais ou menos intenso das grandes interligações afetam consideravelmente o nível de perdas. As perdas aumentam, por exemplo, com uso mais intenso do transporte de energia a longa distância e diminuem quando o consumo é abastecido basicamente por geração mais próxima.

- Ocorrência de temperaturas atípicas: também podem influenciar fortemente a carga. O consumo industrial tende a não ser sensível às variações de temperatura, mas o consumo comercial e, sobretudo, o consumo residencial, respondem fortemente a variações de temperaturas. Parte do efeito da temperatura sobre a carga é sazonal, sendo normal um grande aumento do consumo residencial durante o verão. Contudo, a ocorrência de temperaturas atípicas também pode fazer com que a carga fuja do padrão sazonal normal. Por exemplo, a ocorrência de um verão com temperaturas relativamente amenas pode resultar uma carga anormalmente baixa. Essa redução da carga tem pouco ou nenhum significado econômico e ela pode ser estimada com base em dados disponíveis.

- Variações no calendário: a literatura internacional sobre o tema normalmente se atém à influência do calendário no número de dias úteis de um mês. Para comparar corretamente um indicador de um determinado mês com a leitura do mesmo indicador para o mesmo mês do ano anterior, é preciso fazer um ajuste para refletir a diferença entre o número de dias úteis.

Essa dissertação propõe uma abordagem original para o tratamento do efeito de temperatura, para isto será utilizado um modelo de Regressão Dinâmica que permitirá utilizar os parâmetros e estimar os efeitos na variável Carga. O diferencial deste trabalho em relação ao que já foi feito é o embasamento estatístico dado na retirada deste ofensor.

Vale ressaltar também que foram utilizados os Softwares R, FPW e Excel para compor os resultados deste trabalho.

O texto restante desta dissertação está organizado da seguinte forma: No capítulo 2 está todo o referencial teórico e os principais conceitos metodológicos utilizados na dissertação. O capítulo 3 trata da metodologia empírica e da documentação do que foi feito na dissertação. No capítulo 4 é feita uma análise de

exploratória das séries diárias de Carga, além da aplicação das metodologias e comparação entre os métodos. A dissertação é finalizada no capítulo 5, onde são apresentadas as conclusões sobre o desenvolvimento e as sugestões de futuros trabalhos.

2 Referencial Teórico

2.1. Séries Temporais

Segundo Souza & Camargo (2004) define-se por Série Temporal, também chamada de série cronológica ou histórica, um conjunto de observações de uma dada variável, ordenada sequencialmente no tempo, em geral em intervalos equidistantes (a cada minuto, a cada 15 minutos, horário, diário, semanal, mensal, trimestral, anual). Seja Z_t a representação do valor da variável aleatória Z no instante t , a série temporal pode ser descrita por Z_1, Z_2, \dots, Z_T , onde T representa o tamanho da série ou o número de observações seriais da variável. Segundo Morettin & Toloí (2004), uma série temporal é um conjunto de observações compreendidas no tempo. Uma série temporal pode ser analisada no domínio do tempo ou no domínio da frequência, este trabalho utilizará apenas a abordagem no domínio do tempo.

2.2. Processos Estocásticos

Uma família de variáveis aleatórias indexadas no tempo, ou seja, uma família $\mathbf{Z} = \{Z(t), t \in T\}$ tal que, para cada $T \in R$, $Z(t)$ é uma variável aleatória, é denominado processo estocástico.

Seja Z_t um experimento repetido sob condições idênticas, onde a cada experimento obtém-se um registro dos valores Z_t no tempo, a denominação “ensemble” é dada ao conjunto de todos os possíveis registros. Cada registro particular é uma realização do processo. Deste modo, uma série temporal é uma realização ou função amostral de um determinado processo estocástico. (Souza & Camargo, 2004).

Um processo estocástico é dito estatisticamente determinado quando suas funções de distribuição de probabilidade conjuntas são conhecidas, entretanto isso ocorre raramente e tem-se apenas uma realização do processo estocástico (a série

temporal observada), a partir da qual se deseja inferir todas as características do mecanismo gerador da série. Deste modo, são consideradas duas restrições: Estacionariedade e Ergodicidade.

- Estacionariedade: Um processo estocástico é dito estacionário caso suas características sejam invariantes no tempo; caso contrário, o processo é dito não estacionário. A estacionariedade pode ser do tipo fraca (de segunda ordem) ou forte (estritamente estacionário).

A estacionariedade fraca ocorre quando a média, a variância e as covariâncias são constantes ao longo do tempo, deste modo, a mesma ocorre se, e somente se:

$$\text{i) } E(Z_t) = \mu \quad \forall t$$

$$\text{ii) } \text{Var}(Z_t) = \sigma^2 \quad \forall t$$

iii) $\text{Cov}(Z_t, Z_{t+k}) = \text{Cov}(Z_{t-k}, Z_t) = \gamma_k$, não depende de t , mas apenas do intervalo entre os instantes de tempo (lag k).

Se o processo estocástico for gaussiano (Z_t segue uma distribuição normal) e estacionário de segunda ordem, ele será estritamente estacionário, pois a distribuição normal é determinada unicamente em termos do primeiro e do segundo momento.

Na estacionariedade forte, a forma da distribuição conjunta do processo permanece sem variação mediante uma translação no tempo (Medeiros, 2005). Diferentemente da estacionariedade fraca, esta versão é difícil de ser encontrada, tendo em vista que, em geral, não se consegue especificar a distribuição conjunta de um processo estocástico.

- Ergodicidade: Um processo é dito ergódico quando apenas uma realização do mesmo é suficiente para se obter todas as estatísticas (Teorema Ergódico). Assim, todo processo ergódico também será estacionário, haja vista que uma realização de um processo não estacionário não poderá conter todas as informações necessárias para a especificação do processo.

2.3. Previsão de séries temporais

Um dos objetivos mais importantes na análise de séries temporais é a previsão de valores futuros a partir das informações passadas e atuais. A previsão é essencial para o planejamento e controle de operação em uma variedade de áreas.

Trata-se de uma análise que visa obter as características comportamentais sistemáticas da série e busca apresentar uma descrição dos mecanismos da natureza aleatória do processo estocástico que gerou a série de valores passados e presente. O estabelecimento da estrutura probabilística da realização amostral da variável proporcionará os meios para inferir conclusões sobre as probabilidades associadas a cada possível valor futuro da variável (Souza & Camargo, 2004).

Para que seja realizada a previsão é necessário definir um horizonte de “k” períodos de tempo futuro e o período de origem da previsão “t”.

Deste modo a previsão denotada por $\hat{Z}_t(k)$ é determinada a partir do valor esperado condicional de Z_{t+k} a partir dos valores passados:

$$\hat{Z}_t(k) = E(Z_{t+k} | Z_t, Z_{t-1}, \dots) \quad (2.1)$$

onde Z_{t+k} representa os valores a serem previstos para $k = 1, 2, \dots$

Uma previsão dos valores futuros de uma série pode ser elaborada a partir de alguns métodos estatísticos, e será caracterizada por sua origem, pelo número de passos-à-frente, por seu valor pontual e por uma medida da incerteza associada (em geral a variância). Independentemente da metodologia estatística que for aplicada é necessário formular um modelo matemático que seja capaz de representar o comportamento e as características da série e há várias técnicas que podem ser empregadas na previsão de séries temporais (Miranda & Souza, 2007).

Os modelos univariados utilizam somente os valores passados da própria série para explicar os valores futuros. Podem-se citar os métodos de decomposição, métodos de amortecimento exponencial e os modelos de Box & Jenkins.

Já os modelos de função de transferência e regressão dinâmica (modelos causais), usam os valores passados da própria série a ser prevista e de outra(s) série(s) que seja(m) correlacionada(s), ou seja, uma série de saída está relacionada a uma ou mais séries. Por exemplo, para previsão da carga de energia elétrica pode-se usar o histórico e uma série de alguma variável meteorológica.

Existem também os modelos multivariados, utilizados para prever várias séries simultaneamente, e os modelos de inteligência artificial, que consistem em desenvolver algoritmos que sejam capazes de reproduzir tarefas que são realizadas por seres humanos e necessitam de cognição, como o raciocínio, a aprendizagem e o auto-aperfeiçoamento.

Apesar de toda essa gama de modelos, no presente estudo será abordado apenas o modelo de regressão dinâmica.

2.4. Regressão Dinâmica

Em uma análise de regressão que envolva séries de tempo, a inclusão, não somente de valores correntes, mas também de valores passados (defasados) no modelo matemático denomina esta análise de regressão dinâmica. Tal modelo encontra-se representado genericamente na equação (2.2).

$$y = f \left(\begin{array}{l} y_{t-1}, \dots, y_{t-k}, x_{1,t}, x_{1,t-1}, \dots, x_{1,t-k}, \\ x_{2,t}, x_{2,t-1}, \dots, x_{2,t-k}, x_{m,t}, x_{m,t-1}, \dots, x_{m,t-k} \end{array} \right) \quad (2.2)$$

Para transformar o modelo matemático da equação (2.2) em um modelo estatístico é preciso:

- Escolher uma forma funcional linear;
- Acrescentar um termo estocástico (ϵ_t), que representa as variáveis não incluídas no modelo, assim como algumas flutuações normalmente distribuídas e insignificantes para o modelo;
- Fazer suposições básicas do modelo clássico de regressão linear sobre propriedades desse termo de erro.

Dessa forma e considerando ainda os valores defasados das variáveis, utiliza-se a seguinte equação dinâmica:

$$Y_t = \beta_0 + \alpha \cdot Y_{t-1} + \alpha \cdot Y_{t-2} + \dots + \alpha \cdot Y_{t-k} + \beta_1 \cdot X_{1,t} + \beta_1 \cdot X_{1,t-1} + \dots + \beta_1 \cdot X_{1,t-k} + \beta_2 \cdot X_{2,t} + \beta_2 \cdot X_{2,t-1} + \dots + \beta_2 \cdot X_{2,t-k} + \dots + \varepsilon_t \quad (2.3)$$

onde:

- Y_{t-k} é a variável dependente (endógena) no instante $t-k$, $k \geq 0$ e $t-k \geq 1$;
- α é o coeficiente da variável dependente;
- $X_{i,t-k}$ é a variável causal (exógenas) i no instante $t-k$, $i \geq 1$, $k \geq 0$ e $t-k \geq 1$;
- β_i é o coeficiente da variável causal i , $i \geq 1$;
- ε_t , ruído aleatório associado ao modelo, $\varepsilon_t \sim \text{NID}(0, \sigma^2)$;

Porém, segundo Barros & Souza (1995), a hipótese de independência dos ruídos não é realista na tentativa de modelagem de uma série temporal através de um modelo de regressão, além disso, os resultados e testes usados nos modelos de regressão não são válidos. Ainda segundo os autores, algumas das consequências da autocorrelação dos resíduos são:

1) Os estimadores usuais por mínimos quadrados são ainda não tendenciosos, mas não têm variância mínima.

2) Os estimadores da variância e dos erros padrões dos coeficientes da regressão são subestimados, o que levaria à conclusão de que os estimadores são mais precisos do que na realidade.

3) Os intervalos de confiança para os parâmetros da regressão e os testes de hipóteses relacionados a estes intervalos perdem a validade, como uma consequência direta de 2).

Desta forma, estes três pontos mostram a necessidade de tratamento para o problema de autocorrelação dos erros, pois se sabe que ignorá-los leva, em geral, a inúmeras conclusões falsas. Os modelos de regressão dinâmica estendem os modelos usuais de regressão ao levantarem que a hipótese de independência dos erros não é realista no contexto de séries temporais.

Os modelos de regressão dinâmica combinam a dinâmica de séries temporais e o efeito de variáveis explicativas. Eles devem ser usados quando existe uma estrutura de dependência entre a variável de interesse e possíveis variáveis causais e, ao mesmo tempo, quando a estrutura de correlação da série dependente indicar que não se pode supor a independência dos erros. Nestes

modelos, a variável dependente é explicada por seus valores defasados e pelos valores atuais e passados de variáveis causais ou exógenas. (Zanini, 2000).

O modelo de regressão dinâmica pode ser considerado um caso particular do que é conhecido na literatura como modelo de regressão generalizado de Cochrane & Orcutt (1949), que é dado por:

$$\varphi(B)y_t = \beta x_t + w_t \quad (2.4)$$

$$R(B)w_t = \varepsilon_t \quad (2.5)$$

onde:

- $\varphi(B)$ é o polinômio auto-regressivo de ordem p , isto é: $\varphi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$;
- B é o operador de retardo;
- $R(B)$ é um polinômio auto-regressivo.

A equação (2.4) ainda pode ser reescrita como:

$$w_t = \varphi(B)y_t - \beta x_t$$

Substituindo esta última expressão na equação (2.5), tem-se que:

$$R(B) \cdot [\varphi(B)y_t - \beta x_t] = \varepsilon_t \quad (2.6)$$

Nota-se que na equação (2.6), o modelo de regressão generalizado de Cochrane & Orcutt (1949) introduz defasagens tanto na variável dependente (Y_t) quanto nas causais. Esta equação também indica que a introdução do polinômio autoregressivo $R(B)$ não afeta a relação de causalidade entre y_t e x_t . A equação (2.6) pode ainda ser escrita em termos de novas variáveis y_t^* e x_t^* dadas por:

- $Y_t^* = R(B)y_t$
- $X_t^* = R(B)x_t$

onde $R(B)$ é chamado de “fator comum” e representa a estrutura de correlação presente no erro w_t .

Utilizando-se estas novas variáveis, a equação obtida é:

$$\varphi(B)y_t^* = \beta x_t^* + \varepsilon_t \quad (2.7)$$

Dessa forma, o modelo de regressão generalizado de Cochrane & Orcutt (1949) se reduz ao modelo de regressão dinâmica usual ao se considerar as novas variáveis y_t^* e x_t^* .

O modelo original de Cochrane e Orcutt tem como fator comum $[R(B) = 1 - \alpha B]$ e então a equação (2.7) se reduz a:

$$\varphi(B) \cdot [y_t - \alpha y_{t-1}] = \beta [x_t - \alpha x_{t-1}] + \varepsilon_t \quad (2.8)$$

Neste modelo, o procedimento de estimação é sequencial, isto é, a estimativa inicial de α é igual a 0, e a partir dela estima-se β e $\varphi(B)$ por mínimos quadrados ordinários. A partir destas estimativas encontra-se um estimador de $R(B)$ através da equação $R(B)w_t = \varepsilon_t$. O polinômio estimado $R(B)$ é então usado para transformar y_t e x_t e reestimar β e $\varphi(B)$. O processo é repetido até que se alcance a convergência dos parâmetros. Para maiores detalhes do modelo e do procedimento de estimação veja Barros & Souza (1995).

Na construção de um modelo de regressão dinâmica, normalmente, se emprega a estratégia *bottom-up*, ou seja, inicialmente se considera um modelo simples e com a inclusão de novas variáveis, este modelo é melhorado até encontrar um modelo apropriado. Este procedimento, em grande parte dos casos, não é de fácil elaboração, pois é preciso escolher também as defasagens das variáveis a serem incluídas no modelo.

Para definir um modelo adequado, é importante verificar se os coeficientes estimados são coerentes, além de observar a significância dos parâmetros.

O Gráfico 2.1 representa um fluxograma que mostra um passo-a-passo a ser seguido para a construção de um modelo de regressão dinâmica.

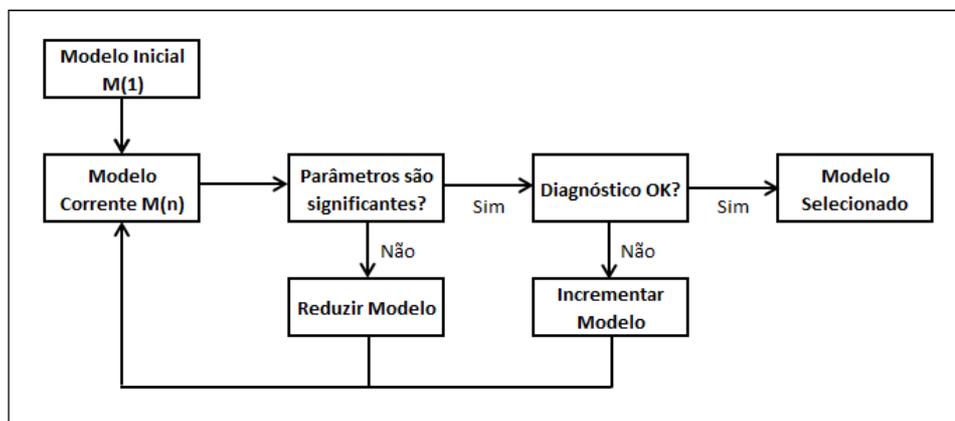


Gráfico 2.1: Fluxograma para construção de um modelo de Regressão Dinâmica
Fonte: Zanini, 2000.

2.5. Testes Estatísticos

Nesta seção serão apresentados os conceitos dos testes que serão utilizados para a série original e para verificar a adequação dos modelos após uma análise gráfica dos resíduos dos mesmos.

2.5.1. Teste de Dickey-Fuller Aumentado (ADF)

A diferença entre um processo autoregressivo e um passeio aleatório é a magnitude do parâmetro da defasagem de primeira ordem. Caso este parâmetro seja igual à unidade, a tendência estocástica (raiz unitária) é detectada.

Para testar a existência de apenas uma raiz unitária (não estacionariedade) pode-se utilizar o teste ADF (Said & Dickey, 1984), cujo propósito é a verificação da existência ou não de tendência estocástica nas séries e se baseia na significância do coeficiente ρ na seguinte equação estimada por mínimos quadrados ordinários (que contém a tendência determinística, na sua versão mais geral):

$$\Delta Y_t = a + \beta t + \rho Y_{t-1} + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \varepsilon_t$$

Hipóteses do Teste:

$H_0: \rho = 0$ ou “Existe raiz unitária” (Não estacionariedade).

$H_1: \rho < 0$ ou “Não existe raiz unitária” (Estacionariedade).

Regra de decisão:

A hipótese nula será rejeitada quando o nível de significância fixado em α é maior que o P-valor.

2.5.2. Testes de Normalidade

2.5.2.1. Shapiro-Wilk

O teste de Shapiro-Wilk (Shapiro & Wilk, 1965), calcula uma estatística W que testa se uma amostra aleatória, X_1, X_2, \dots, X_n , é proveniente de uma distribuição normal. Valores pequenos de W são evidências do pressuposto de

normalidade. Esse método foi obtido através de simulações de Monte Carlo e reproduzidas por Pearson & Hartley (1972). Além disso, dadas quaisquer variáveis aleatórias X_1, X_2, \dots, X_n , as estatísticas de ordem são, também, variáveis aleatórias $X_{(1)}, X_{(2)}, \dots, X_{(n)}$, definidas pela escolha dos valores de X_1, X_2, \dots, X_n em ordem crescente.

A Estatística de Teste W é calculada da seguinte forma:

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

onde:

- $(a_1, a_2, \dots, a_n) = \frac{\mathbf{m}^T \cdot \mathbf{V}^{-1}}{(\mathbf{m}^T \cdot \mathbf{V}^{-1} \cdot \mathbf{V}^{-1} \cdot \mathbf{m})^{1/2}}$;
- $\mathbf{m} = (m_1, m_2, \dots, m_n)^T$;
- m_i é o valor esperado da estatística de ordem de variáveis aleatórias independentes e estatisticamente distribuídas (iid) amostradas de uma distribuição normal padrão, $1 \leq i \leq n$;
- \mathbf{V} é a matriz de covariância da estatística de ordem;
- \bar{x} é a média da amostra.

Hipóteses do Teste:

H_0 : “A série segue uma distribuição normal”.

H_1 : “A série não segue uma distribuição normal”.

Regra de decisão:

Fixado um nível de significância α , a hipótese nula (H_0) será rejeitada caso o P-valor do teste seja menor que esse nível, onde o P-valor é calculado a partir da estatística W . Quanto maior W , menor será o P-valor e conseqüentemente haverá mais evidência para se rejeitar H_0 .

2.5.2.2. Jarque-Bera (JB)

O teste de normalidade Jarque-Bera (1980) é baseado nas diferenças entre os coeficientes de assimetria (3º momento central) e curtose (4º momento central), servindo para testar a hipótese nula de que a amostra foi extraída de uma distribuição normal. Para a realização deste teste, calculam-se, primeiramente, a assimetria S e a curtose C dos resíduos e utiliza-se o seguinte teste estatístico:

$$JB = n \left[\frac{S^2}{6} + \frac{(C - 3)^2}{24} \right]$$

onde:

- n = número de observações;

$$S = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}};$$

$$C = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}.$$

A estatística de teste JB segue uma distribuição qui-quadrado com dois graus de liberdade se n dados vem de uma distribuição normal. Se o valor de JB for muito baixo, a hipótese nula de normalidade da distribuição dos erros aleatórios não pode ser rejeitada. Se o valor de JB for muito alto, rejeita-se a hipótese de que os resíduos ou erros aleatórios se comportam como uma distribuição normal.

2.5.3. Testes para autocorrelação serial

2.5.3.1. Durbin-Watson

Desenvolvido pelos estatísticos Durbin & Watson (1950), é o teste mais utilizado na literatura para detecção da correlação serial dos resíduos.

A estatística de teste d é calculada da seguinte forma:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

onde:

- n representa o número de observações da série.
- e_t é o resíduo no instante t .

Hipóteses do Teste:

H_0 : “Ausência de Autocorrelação serial de primeira ordem”.

H_1 : “Existe Autocorrelação serial de primeira ordem”.

Regra de decisão:

O uso da estatística d depende de valores críticos d_L e d_U obtidos por Durbin e Watson, estes valores dependem do número de observações n e do número de variáveis explicativas. A tabela D5 do Gujarati (2006) contém estes limites.

Para testar a existência de autocorrelação positiva com nível de significância α , a estatística de teste d é comparada d_L e d_U da seguinte forma:

- $d < d_L$: Existe evidência de autocorrelação positiva nos erros, rejeita-se H_0 ;
- $d > d_U$: Existe evidência de que os erros não são positivamente correlacionados;
- $d_L < d < d_U$: Teste inconclusivo.

Para testar a existência de autocorrelação negativa com nível de significância α , a estatística de teste $(4-d)$ é comparada d_L e d_U da seguinte forma:

- $(4-d) < d_L$: Existe evidência de autocorrelação negativa nos erros, rejeita-se H_0 ;
- $(4-d) > d_U$: Existe evidência de que os erros não são negativamente correlacionados;
- $d_L < (4-d) < d_U$: Teste inconclusivo.

2.5.3.2. Teste Ljung-Box & Box-Pierce

Box & Pierce (1970) sugeriram um teste para as autocorrelações dos resíduos estimados, o qual indica se estes valores estão muito altos.

Inicialmente, é calculada uma estimativa para as autocorrelações:

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{e}_t \hat{e}_{t-k}}{\sum_{t=1}^n \hat{e}_t^2}$$

onde:

- \hat{e}_t é o resíduo padronizado no instante t ;
- n é o tamanho da amostra.

A estatística de teste é:

$$Q = n \sum_{k=1}^K \hat{r}_k^2$$

onde K é o número de defasagens.

Se o modelo for apropriado, a estatística de teste Q terá aproximadamente uma distribuição Qui-Quadrado com K graus de liberdade.

Hipóteses do Teste:

H_0 : “Os resíduos não são correlacionados”.

H_1 : “Os resíduos são correlacionados”.

Regra de decisão:

Fixado um nível de significância α , rejeita-se a hipótese nula se $Q > \chi_{1-\alpha, K}^2$.

Esta estatística de teste Q é usada com frequência para testar se a série temporal é de ruído branco.

Existe na literatura uma variante da estatística Q definida como estatística de Ljung-Box (Ljung & Box, 1978):

$$LB = n(n + 2) \sum_{k=1}^K \left(\frac{\hat{r}_k^2}{n - k} \right) \sim \chi_{K-p-q}^2$$

onde p e q são as ordens do modelo ajustado.

Hipóteses do Teste:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0.$$

$$H_1: \text{Pelo menos um } \rho_k \neq 0.$$

Regra de decisão:

Se o LB calculado excede o valor crítico de LB na distribuição de Qui-Quadrado ao nível de significância escolhido, há indícios para se rejeitar a hipótese nula de que todos os (verdadeiros) ρ_k são iguais a zero.

3 Metodologia

A seguir será descrita uma metodologia empírica, a partir de um método multiplicativo, para a retirada dos efeitos Calendário e Temperatura da série de Carga. Em seguida será proposta uma abordagem alternativa utilizando Regressão Dinâmica (método aditivo). É importante salientar que esta Regressão será aplicada apenas para a retirada do efeito de Temperatura da série de Carga.

Para efeito de notação e melhor entendimento, as variáveis Carga e Temperatura serão escritas da seguinte forma:

$$\text{CARGA}_{tjs}^{r,d,m,a} \quad \text{TEMP}_{tjs}^{r,d,m,a} \quad (3.1)$$

onde:

- t é o índice que referencia o dia em análise. Na base ordenada pelo tempo, isto é, do dia mais antigo até o dia mais recente, este índice pode variar de 1 até o número total de dias considerado. Com exceção do índice s , todos os outros são dependentes do índice t ;
- s é o índice que indica a qual subsistema a observação pertence;
- a indica o ano da observação;
- m indica o mês da observação, $1 \leq m \leq 12$;
- d indica o dia da observação, $1 \leq d \leq 31$;
- r indica o dia da semana, varia de domingo à sábado, além dos feriados e semi-feriados, isto é $1 \leq r \leq 9$. Domingo será codificado com o valor de $r = 1$, Segunda-feira com o valor de $r = 2$, assim sucessivamente até o Sábado que receberá o valor de $r = 7$. Os feriados receberão a codificação de $r = 9$ e os semi-feriados, $r = 8$.

3.1. Método Empírico (Multiplicativo)

Segundo Castro *et al.* (2012), as séries ajustadas de carga são construídas de forma a excluir a influência de fatores não econômicos e/ou fortuitos sobre a carga. Elas apresentam duas vantagens sobre as séries tradicionais. Em primeiro lugar, elas são uma boa base para a previsão de carga no médio e longo prazos, na medida em que espelham não a carga efetivamente verificada, mas aquela que teria sido observada se não ocorressem fatores fortuitos. Em segundo lugar, as séries ajustadas de carga permitem uma análise mais fina da relação entre

economia e carga, constituindo um indicador de atividade econômica de boa qualidade.

3.1.1. Ajuste da Série diária de Carga pelo fator Calendário

O ajuste calendário para a carga dos subsistemas consiste em reestimar a série diária de carga para obter uma série que não reflete as especificidades do calendário verificado.

O ajuste calendário consiste em recalcular a carga de cada dia multiplicando-a pelo inverso da razão normalmente observada entre a carga do dia da semana em questão e a carga semanal.

Mais detalhes desta metodologia podem ser encontradas em Castro *et al.* (2012).

3.1.1.1. Pesos para cada dia

Num primeiro momento, será preciso calcular o peso de cada observação dentro da semana em que esta observação pertence, isto é, peso do dia r de uma semana específica ($\text{Peso}_{t|s}^{r,d,m,a}$). Uma vez feito o referido cálculo para todos os dias da série, é possível calcular o peso típico de um dia da semana em uma semana qualquer (Peso_Típico^r).

Para realizar o cálculo do $\text{Peso}_{t|s}^{r,d,m,a}$, deve-se observar ainda se a semana contém feriado ou semi-feriado. Sabendo que a semana se inicia num domingo e termina em um sábado, a semana que não possui um feriado ou semi-feriado neste intervalo, será considerada semana do tipo 1. Caso a semana contenha feriado ou semi-feriado, será considerada semana do tipo 2.

A fórmula para calcular o peso do dia r dentro de uma semana do tipo 1 é:

$$\text{Peso}_{t|s}^{r,d,m,a} = \frac{\text{CARGA}_{t|s}^{r,d,m,a}}{\sum_{i=1-r}^{7-r} \text{CARGA}_{t+i|s}^{r+i,d,m,a}} \quad (3.2)$$

Como foi visto na fórmula (3.2), o peso do dia t é calculado dividindo-se a carga do dia t pela média das cargas dos dias da semana que o dia t pertence. Entende-se que para uma semana com feriado/semi-feriado esse peso é distorcido. Desta forma, compara-se a carga do dia com a média entre a carga média da semana anterior com a carga média da semana posterior se ambas não possuem feriados. Se uma das duas semanas em questão tiver um feriado ou semi-feriado, calcula-se a média do par de semanas sem feriados ou semi-feriados equidistantes da semana a que t pertence e o mais próximo possível dela.

O peso típico de cada dia da semana será a média dos $\text{Peso}_{t|s}^{r,d,m,a}$ para todos os dias r :

$$\text{Peso_Típico}^r = \frac{\sum_{i=1}^s \text{Peso}_{m|s}^{r,d,m,a}}{s} \quad (3.3)$$

onde s é o número de ocorrências do dia r na base.

3.1.1.2. Carga diária ajustada pelo fator calendário

O valor da Carga diária ajustada pelo fator Calendário será calculado pela fórmula:

$$\text{CARGA_adj}_{t|s}^{r,d,m,a} = \frac{\text{CARGA}_{t|s}^{r,d,m,a}}{\text{Peso_Típico}^r} \quad (3.4)$$

3.1.2. Ajuste da Série diária de Carga pelo fator Temperatura

O ajuste de temperatura consiste em compensar o efeito da ocorrência de temperaturas atípicas sobre a carga. Isto implica em fazer uma reestimativa das séries de carga diária utilizando as temperaturas diárias típicas em lugar das temperaturas verificadas. Se a temperatura verificada é maior do que a normal para a época do ano, a carga, que reage positivamente ao calor, também deve ter sido maior do que aquela que normalmente se verificaria. O ajuste de temperatura consiste, nesse caso, em calcular uma carga ajustada menor que a verificada,

sendo o ajuste proporcional à diferença entre a temperatura verificada e a temperatura típica. (Castro *et al.*, 2012)

3.1.2.1.

Verificação de correlação entre as séries diárias de Carga e Temperatura

Sabe-se que a influência da variável Temperatura na variável Carga é diferente em cada época do ano. Dessa forma a série de dados será dividida em doze bancos de dados, onde cada um desses 12 bancos reúne as informações de um mesmo mês, isto é, serão realizadas doze regressões lineares que levam em consideração todos os dias de um mês específico. Por exemplo: a 1ª regressão reúne apenas os dados referentes do mês de janeiro.

Será considerada também a padronização das séries, ou seja, serão realizadas sete diferenciações nas variáveis Carga_adj e Temperatura. Dessa forma:

$$DTemp_{t-i|s}^{r,d,m,a} = Temp_{t|s}^{r,d,m,a} - Temp_{t-i|s}^{r,d,m,a} \quad (3.5)$$

$$DCARGA_{t-i|s}^{r,d,m,a} = \frac{CARGA_adj_{t|s}^{r,d,m,a}}{CARGA_adj_{t-i|s}^{r,d,m,a}} - 1 \quad (3.6)$$

onde $1 \leq i \leq 7$.

Dessa forma, cada par de observações $(CARGA_adj_{t|s}^{r,d,m,a}; Temp_{t|s}^{r,d,m,a})$ será substituído por outros sete pares de observações $(DCARGA_{t-i|s}^{r,d,m,a}; DTemp_{t-i|s}^{r,d,m,a})$, onde i varia de 1 a 7.

A partir deste ponto, os novos bancos possuirão as variáveis $DCARGA_{k|s}^m$ e $DTemp_{k|s}^m$, onde k é igual a $(7t-i)$, i é a i -ésima defasagem de t e m o mês que caracteriza o banco.

3.1.2.2.

Análise do coeficiente angular das regressões

As regressões são desenvolvidas com intuito de se utilizar o coeficiente angular no ajuste do valor da Carga através do efeito de Temperatura. Caso este coeficiente seja significativo, será feito o ajuste da Carga pelo efeito da temperatura, caso contrário, não será feito o ajuste de temperatura. Dado o modelo geral de regressão linear simples:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \quad (3.7)$$

Para este trabalho, a variável $DCARGA_{k|s}^m$ será utilizada como a variável resposta Y e $DTemp_{k|s}^m$ será utilizada como a variável regressora X_1 .

O teste de hipóteses para o coeficiente angular pode ser descrito da seguinte maneira:

$$H_0: \beta_1 = 0 \quad x \quad H_1: \beta_1 \neq 0$$

Supondo H_0 verdadeira, pode ser realizado um teste t , onde a estatística t é calculada a partir da equação (3.7).

Essa estatística de teste segue uma distribuição t com $(n-2)$ graus de liberdade, onde n é o número de observações da amostra. Portanto, a regra de decisão para o teste de hipótese é dada por: Se $t_{calc} < -t_{tab}$ ou $t_{calc} > t_{tab}$, rejeita-se a hipótese nula H_0 , ou seja, há evidências para concluir que existe uma relação significativa entre as variáveis. O valor tabelado (t_{tab}) é dado pela tabela da distribuição t e depende do nível de significância α considerado.

Após ter sido calculado o coeficiente de regressão para as observações de cada sub-banco, é necessário retirar da análise os dias considerados atípicos. Para isto, serão calculados intervalos de confiança, considerando o erro padrão e o valor previsto da Carga, para avaliar se a variação da Carga de determinado dia não é um valor atípico. Caso o valor da variação da Carga se encontre dentro do intervalo de confiança calculado, isto significa que este valor não é atípico. Caso contrário, ou seja, o valor da variação da Carga não se encontre dentro do intervalo de confiança calculado, este valor será considerado um “outlier” e excluído da análise. Com a retirada desses “outliers”, o coeficiente de regressão é recalculado e utilizado no cálculo do Fator de Temperatura (subtópico 3.1.2.5).

A notação do coeficiente de regressão de um mês m específico, de um subsistema s é dada por: b_{1s}^m .

O valor previsto da Carga é:

$$\widehat{DCarga}_{k|s}^m = b_{0s}^m + b_{1s}^m \cdot DTemp_{k|s}^m \quad (3.8)$$

onde b_{0s}^m é o valor estimado para o coeficiente linear da equação de regressão, do mês m e subsistema s .

O intervalo de confiança será calculado da seguinte forma:

$$\left[\widehat{DCarga}_{k|s}^m \pm EP \cdot mg \right] \quad (3.9)$$

onde:

- EP é o erro padrão entre as variáveis DCarga e DTemp;
- mg é o número de erros-padrão considerado.

3.1.2.3. Temperatura Típica diária

Esta temperatura típica do dia, que leva em consideração todas as observações da série diária de Temperatura, será denotada por TTA e é a média de temperatura do mesmo dia e mesmo mês de todos os anos, isto é:

$$TTA_{L|s}^{d,m} = \frac{\sum_{i=w}^v Temp_{t|s}^{r,d,m,i}}{v - w + 1} \quad (3.10)$$

onde:

- L é o dia do ano, o dia 1º de janeiro, independente do ano, será referido como $L=1$ e o dia 31 de dezembro como $L=365$ ou seja, $1 \leq L \leq 365$;
- v é o último ano do banco de dados que contenha o dia/mês em análise;
- w é o primeiro ano do banco de dados que contenha o dia/mês em análise.

3.1.2.4. Temperatura Típica diária II

Conforme descrito em Castro *et al.* (2012), esta temperatura típica do dia é calculada a partir das TTA's e será denotada por TTM. A TTM é a média das 31 TTA's dos dias mais próximos, isto é:

$$TTM_{L|s}^{d,m} = \frac{\sum_{i=L-15}^{L+15} TTA_{i|s}^{d,m,a}}{31} \quad (3.11)$$

Foi utilizado o tamanho de janela 31 para o cálculo da TTM. Porém conforme pode ser visto no Apêndice, segundo o critério utilizado nesta dissertação, o tamanho de janela que otimiza a retirada do efeito temperatura da série de Carga é 3.

3.1.2.5. Fator de temperatura

Este fator será utilizado para ajustar o valor da Carga, observa-se que leva em consideração o parâmetro estimado pela regressão ($b_{l|s}^m$), desenvolvido no tópico 3.1.2.2.

$$\text{Fat_Temp}_{t|s}^{r,d,m,a} = 1 + \left[\frac{(\text{TTM}_{L|s}^{d,m} - \text{Temp}_{t|s}^{r,d,m,a}) \cdot b_{l|s}^m}{100} \right] \quad (3.12)$$

3.1.2.6. Carga Ajustada pela temperatura

O efeito da temperatura será corrigido através de um método multiplicativo, isto é:

$$\text{CARGA_adj}_{t|s}^{r,d,m,a} = \text{CARGA_adj}_{t|s}^{r,d,m,a} \cdot \text{Fat_Temp}_{t|s}^{r,d,m,a} \quad (3.13)$$

3.1.3. Ajuste da Série diária pelo fator Semana Morta

A semana entre o natal e o fim de ano é um período de baixíssimo consumo de energia elétrica, pois muitas indústrias aproveitam as festas para promover férias coletivas, parando ou reduzindo muito a produção. Observa-se que na semana de festas há redução de carga com relação à típica tanto em dias que não são formalmente feriados ou semi-feriados como nos próprios feriados. Notadamente, 25 de dezembro costuma ter uma carga baixíssima, muito menor que a de um feriado normal. Mais detalhes em Castro *et al.* (2012).

Os dias pertencentes às semanas de festas de fim de ano receberão um tratamento especial. Para cada observação da série de Carga será calculado um desvio do Peso Semanal (equação 3.2) para o Peso Típico do dia (equação 3.3):

$$\text{Desvio}_{t|s}^{r,d,m,a} = \frac{\text{Peso}_{t|s}^{r,d,m,a}}{\text{Peso_Típico}_{t|s}^{r,d,m,a}} - 1 \quad (3.14)$$

O fator para corrigir um dia que pertence à semana morta, será calculado da seguinte forma:

$$\text{Fat_Semana_Morta}_s = \frac{\sum_{t=1}^k \text{Desvio}_{t|s}^{r,d,m,a}}{k} \quad (3.15)$$

onde k é o número total de dias pertencentes à semana morta na série de Carga. Dessa forma, o somatório conterà apenas os dias de semana morta.

O fator de ajuste para a Semana Morta não será aplicada para dias comuns, com isso a série de Carga diária ajustada pelo fator Semana Morta será:

$$\text{CARGA_diaria}_{t|s}^{r,d,m,a} = \begin{cases} \text{CARGA_adjT}_{t|s}^{r,d,m,a} & , \text{ dia comum} \\ \frac{\text{CARGA_adjT}_{t|s}^{r,d,m,a}}{1 + \text{Fat_Semana_Morta}_s} & , \text{ dia da semana morta} \end{cases} \quad (3.16)$$

3.1.4. Ajuste da Carga mensal

A série ajustada mensal leva em conta tanto os dados disponíveis em frequência diária como os dados mensais de geração de pequenas centrais, descontinuidades de medição e perdas na rede básica. O ajuste neste caso é trivial e a Carga Mensal Líquida de Perdas está exemplificada nos 5 passos abaixo.

3.1.4.1. Carga Mensal

É a média dos valores da série de Carga diária.

$$\text{CARGA_Mensal}_s^{m,a} = \frac{\sum_{t=1}^v \text{Carga}_{t|s}^{r,d,m,a}}{v} \quad (3.17)$$

onde v é o número de dias do mês m em um ano a específico.

3.1.4.2. Carga Ajustada Mensal

É a média dos valores da Carga diária ajustada

$$\text{CARGA_Mensal_adj}_s^{m,a} = \frac{\sum_{t=1}^v \text{CARGA_diaria}_{t|s}^{r,d,m,a}}{v} \quad (3.18)$$

onde v é o número de dias do mês m em um ano a específico.

3.1.4.3.**Fator de ajuste mensal**

É a razão entre as séries Carga_Mensal e Carga_Mensal_adj.

$$\text{Ajuste_Mensal}_s^{m,a} = \frac{\text{Carga_Mensal_Adj}_s^{m,a}}{\text{Carga_Mensal}_s^{m,a}} \quad (3.19)$$

3.1.4.4.**Carga Mensal Ajustada**

É a Carga Mensal, obtida no subtópico (3.1.4.3), somada à Geração de Energia de Pequenas Centrais e aplicado o fator de correção mensal.

$$\text{CARGA_Ajustada}_s^{m,a} = \left(\frac{\text{CARGA_Mensal}_s^{m,a} + \text{Geração_PEQ_Centrais}_s^{m,a}}{\text{Geração_PEQ_Centrais}_s^{m,a}} \right) \cdot \text{Ajuste_Mensal}_s^{m,a} \quad (3.20)$$

3.1.4.5.**Carga Mensal Ajustada Líquida de Perdas**

Finalmente, para chegar à série Mensal de Carga Ajustada final, basta descontar as perdas na rede básica da Carga Mensal Ajustada.

$$\text{CARGA_Ajustada_Líquida}_s^{m,a} = \text{CARGA_Ajustada}_s^{m,a} \cdot (1 - \text{Perdas}_s^{m,a}) \quad (3.21)$$

onde $\text{Perdas}_s^{m,a}$ representa um valor, em porcentagem, de carga perdida na rede básica.

3.1.5.**Carga Trimestral Ajustada**

É a média da Carga Ajustada Líquida de Perdas entre três meses consecutivos, isto é, dentro do ciclo de um ano, serão obtidas 4 observações. A média entre os meses de janeiro, fevereiro e março formarão a primeira observação do ano, a segunda será formada com a média dos valores mensais de abril, maio e junho, assim por diante até o último mês do ano (dezembro).

$$CARGA_Tri_s^{q,a} = \frac{\sum_{i=3q-2}^{3q} CARGA_Ajustada_Líquida_s^{i,a}}{3} \quad (3.22)$$

onde $1 \leq q \leq 4$.

O gráfico 3.1 mostra um fluxograma para os passos da metodologia empírica de limpeza da série diária de Carga até a série trimestral de Carga.

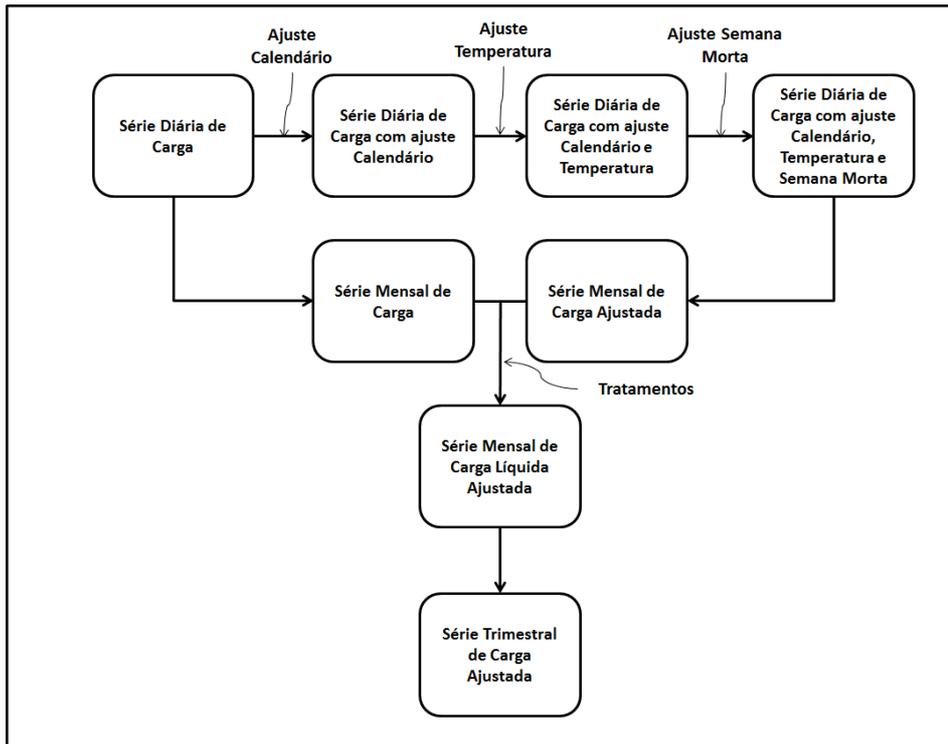


Gráfico 3.1: Fluxograma da metodologia empírica de limpeza da série diária de Carga até a série trimestral de Carga

3.1.6.

Varição do Trimestral em relação ao mesmo Trimestre do ano anterior

Este Indicador, que é adimensional, dirá quanto um trimestre variou percentualmente em comparação ao mesmo trimestre do ano anterior.

$$CARGA_Tri_Var_s^{q,a} = \frac{CARGA_Tri_s^{q,a}}{CARGA_Tri_s^{q,a-1}} - 1 \quad (3.23)$$

3.2. Regressão Dinâmica (Método Aditivo)

A proposta deste trabalho é retirar o efeito de Temperatura da variável Carga a partir do modelo de Regressão Dinâmica, proposto em 2.4. Será ajustado um modelo para a série diária de Carga_adj com a série diária de Temperatura como variável causal. A equação genérica de ajuste é:

$$CARGA_adj_{t|s}^{r,d,m,a} = \beta_0 + \sum_{i=1}^{k_1} \alpha_i \cdot CARGA_adj_{t-i|s}^{r,d,m,a} + \sum_{i=1}^{k_2} \beta_i \cdot Temp_{t-i+1|s}^{r,d,m,a} + \varepsilon_t \quad (3.24)$$

No tópico (2.4) é mostrado como se identifica as defasagens da variável Temperatura que explicam a variável Carga_adj e como são estimados os parâmetros α e β .

Como um dos objetivos deste método é retirar o efeito de Temperatura da variável de Carga, basta subtrair o terceiro termo do lado direito da equação (3.24) da Carga_adj. Desta forma, a Carga_adj ajustada pelo efeito temperatura é representada da seguinte maneira:

$$CARGA_adj_{t|s}^{r,d,m,a} = CARGA_adj_{t|s}^{r,d,m,a} - \sum_{i=1}^{k_2} \beta_i \cdot Temp_{t-i+1|s}^{r,d,m,a} \quad (3.25)$$

4 Resultados

As séries de carga foram obtidas a partir de dados disponibilizados pelo ONS, que são utilizados na rotina de previsão de carga e no acompanhamento dos resultados da operação do sistema. Os dados se referem a cada um dos quatro subsistemas do Sistema Interligado Nacional (Nordeste, Sul e Sudeste / Centro-Oeste). Não foi possível obter os dados de temperatura da região Norte. Esses dados, que serão utilizados para a criação das séries ajustadas, têm frequência diária. São eles:

- *Carga por subsistema.* O Centro Nacional de Operação do Sistema (CNOS) disponibiliza séries de carga por subsistema com várias periodicidades. A metodologia de ajuste escolhida utilizará dados em frequência diária.

- *Temperaturas equivalentes por subsistema.* O ONS utiliza na previsão de carga de curto prazo um programa desenvolvido pela COPPE/UFRJ que calcula temperaturas equivalentes por subsistema a cada hora. Os dados de temperatura gerados por este programa foram utilizados nos testes e foram incorporados à metodologia. As temperaturas equivalentes são médias ponderadas das temperaturas observadas em diversos locais de um subsistema (aerportos essencialmente). A ponderação das temperaturas locais foi calculada por um Algoritmo Genético cuja função objetivo é maximizar a correlação entre a própria temperatura equivalente e a carga horária. Do ponto de vista prático, o peso dado à medição de temperatura em uma localidade não espelha a representatividade da carga local no subsistema, mas sim o quão sensível é a carga do subsistema a variações na temperatura do local. Por exemplo, embora a carga do Rio de Janeiro seja bem menor do que a carga de São Paulo, no cálculo das temperaturas equivalentes do subsistema Sudeste / Centro-Oeste, o peso das medições de temperatura do Rio é maior que o peso das medições em São Paulo. Isso ocorre porque o consumo de energia para refrigeração no Rio é muito elevado e é bastante sensível a variações de temperatura.

- *Calendário por subsistema.* A rotina de previsão de carga diária do ONS inclui a manutenção de um calendário. Esse calendário inclui não apenas os dias de semana e feriados nacionais, mas também os chamados de semi feriados. Semi feriados são dias úteis com carga consideravelmente reduzida. Exemplos são os dias "enforcados" – uma segunda feira em uma semana em que a terça feira é feriado nacional – e os feriados em estados ou municípios muito representativos da carga do subsistema – por exemplo o dia 9 de julho, feriado em São Paulo para comemoração da Revolução Constitucionalista de 1932.

Outro conjunto de dados utilizados só está disponível em frequência mensal. São eles:

- *Geração de pequenas centrais.* As estatísticas de carga do CNOS, mencionadas anteriormente, contemplam apenas usinas que são despachadas ou programadas pelo ONS. Mas há diversas pequenas centrais que estão conectadas diretamente à rede de distribuição e que são despachadas pelas distribuidoras. Os dados de geração dessas usinas não estão disponíveis *on-line* para o ONS, sendo fornecidos pelas distribuidoras através de um relatório mensal. A geração de pequenas centrais representa em torno de 3,5% da carga medida pelo CNOS.

- *Perdas na rede básica.* As perdas no sistema de transmissão são calculadas pelo ONS desde junho de 2005. São dados mensais que estão disponíveis somente ao final do mês seguinte ao mês de referência;

Carga de sistemas isolados anterior à incorporação ao SIN. O ONS solicitou às distribuidoras do antigo Sistema Isolado Acre-Rondônia as medições de carga anteriores à interligação ao SIN. Estes dados foram utilizados para corrigir a série de carga no período anterior à incorporação dos sistemas isolados.

4.1. Análise Exploratória dos dados

Os gráficos 4.1, 4.2, 4.3 e 4.4 exibem as séries originais de Carga diária em um período de junho de 2003 até julho de 2012, dos Subsistemas Norte, Nordeste, Sul e Sudeste/Centro-Oeste, respectivamente, evoluindo no tempo. A tabela 4.1 mostra algumas estatísticas destas séries.

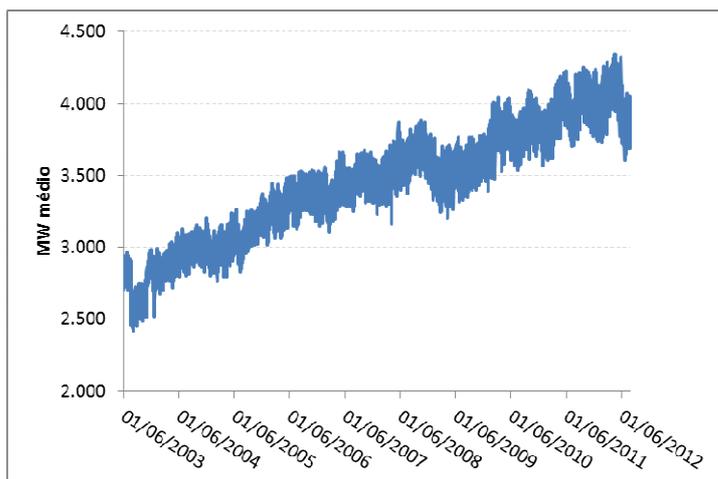


Gráfico 4.1: Série de carga do subsistema Norte
Fonte: ONS, elaboração própria.

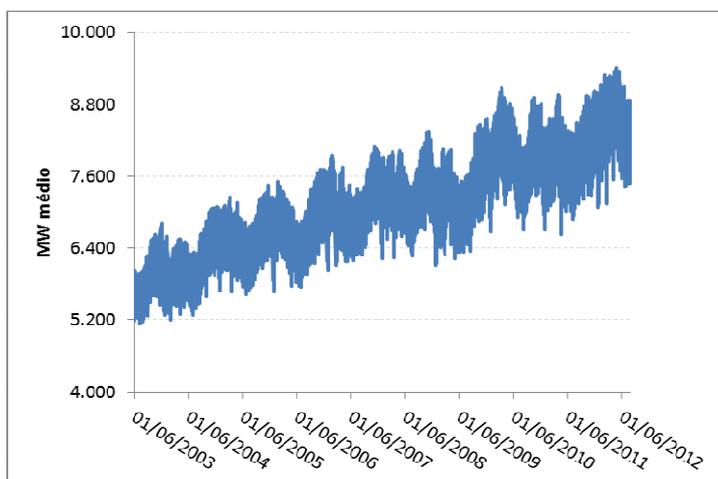


Gráfico 4.2: Série de carga do subsistema Nordeste
Fonte: ONS, elaboração própria.

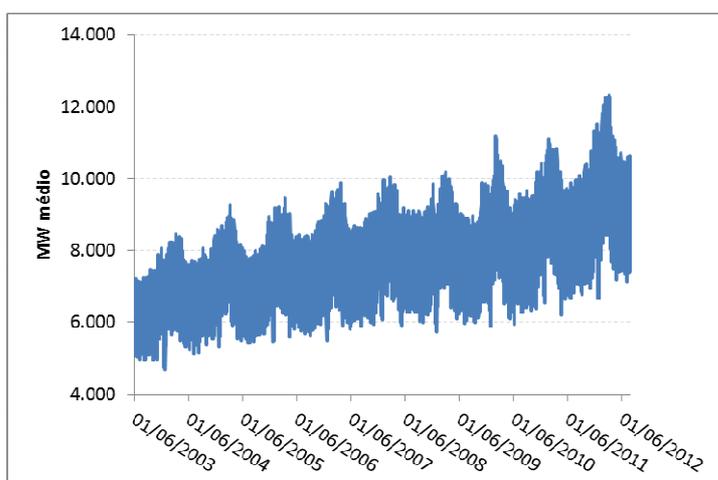


Gráfico 4.3: Série de carga do subsistema Sul
Fonte: ONS, elaboração própria.

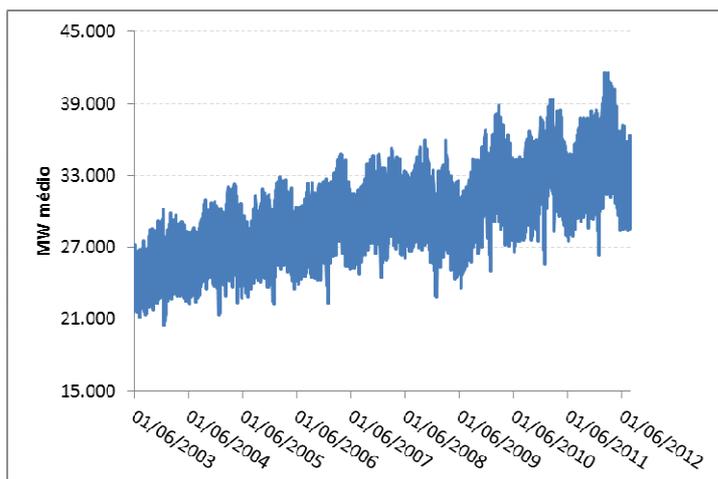


Gráfico 4.4: Série de carga do subsistema Sudeste / Centro-Oeste

Fonte: ONS, elaboração própria.

Tabela 4.1: Estatísticas das Séries de carga dos subsistemas

Subsistemas	Média	Mínimo	Máximo	Desvio-Padrão
Norte	3.485	2.415	4.344	407
Nordeste	7.313	5.146	9.413	880
Sul	8.247	4.677	12.330	1.302
Sudeste / Centro-Oeste	30.795	20.448	41.582	3.738

Fonte: ONS, elaboração própria.

A seguir serão feitos alguns gráficos e testes estatísticos para explorar as séries diárias de Carga. Os gráficos 4.5, 4.6, 4.7 e 4.8 mostram o monthplot dos 4 subsistemas. Este gráfico expressa no eixo das abcissas os dias do ciclo da série, variando de 1 a 365, o eixo da ordenadas mostra os valores de todos os anos para o dia referente ao indicado no eixo das abcissas, além disso também é visto a média de Carga de todas as ocorrências de um determinado dia do ano. Através destes gráficos, será confirmada a ideia de que há aumento da carga no início do ano e queda no meio do ano para os subsistemas Nordeste, Sul e Sudeste / Centro-Oeste, além de que, no meio do segundo semestre, há um novo aumento da carga. O subsistema Norte não apresenta mudança na Carga ao longo do ano, isto é, não há mudança de Carga com o aumento ou diminuição da temperatura.

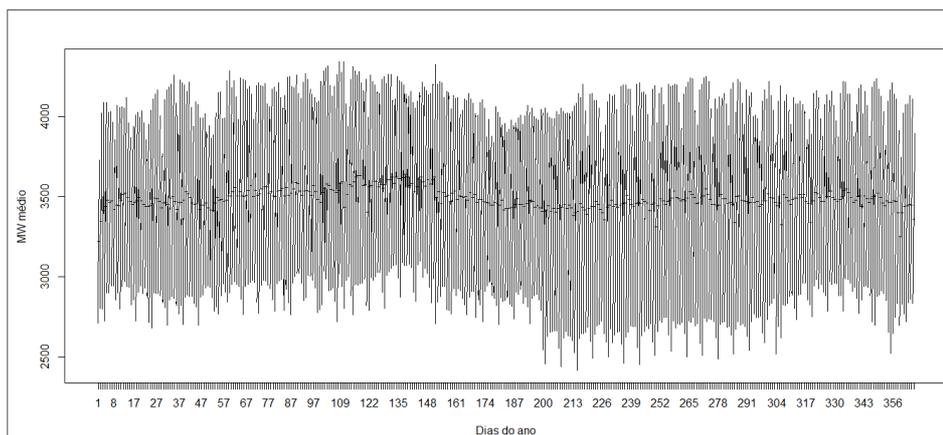


Gráfico 4.5: Monthplot do subsistema Norte
Fonte: ONS, elaboração própria.

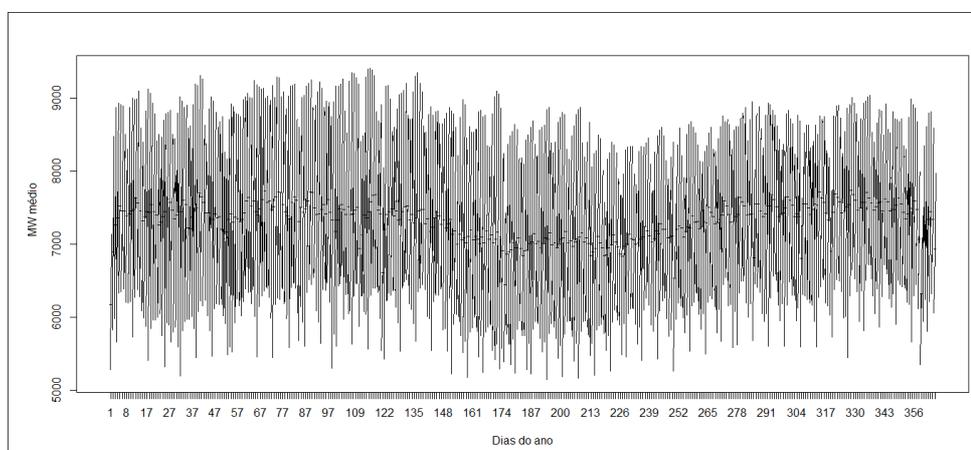


Gráfico 4.6: Monthplot do subsistema Nordeste
Fonte: ONS, elaboração própria.

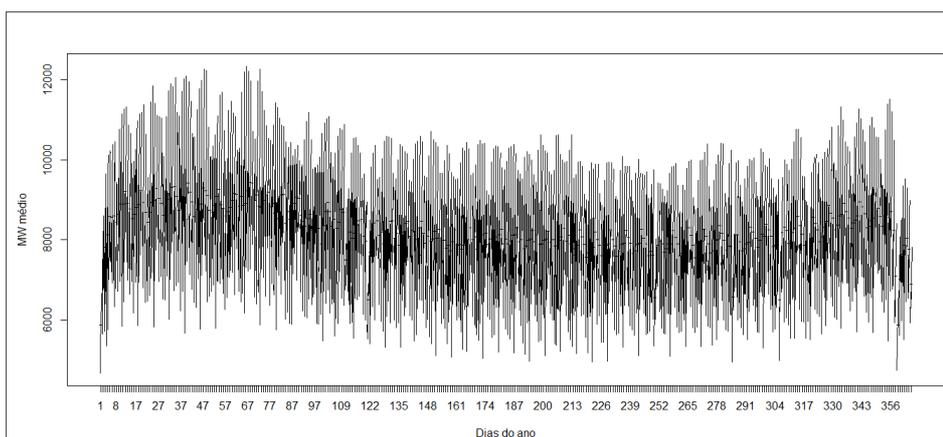


Gráfico 4.7: Monthplot do subsistema Sul
Fonte: ONS, elaboração própria.

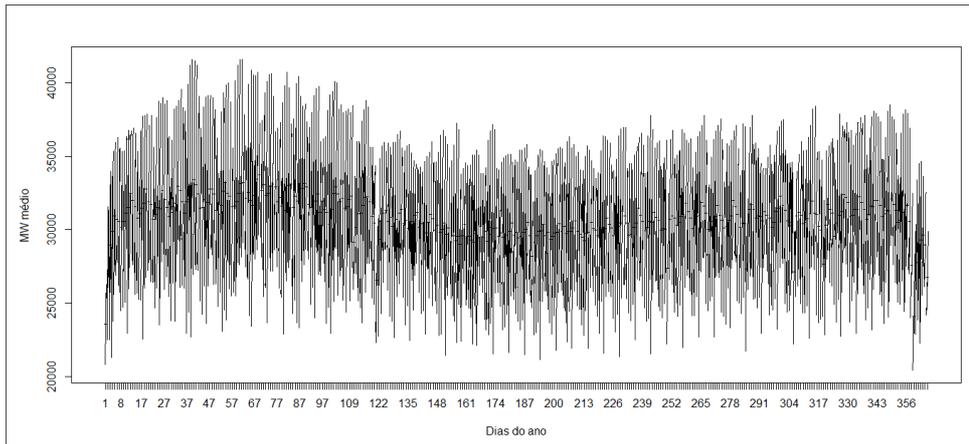


Gráfico 4.8: Monthplot do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

Os gráficos 4.9, 4.10, 4.11 e 4.12 mostram os Histogramas de cada subsistema.

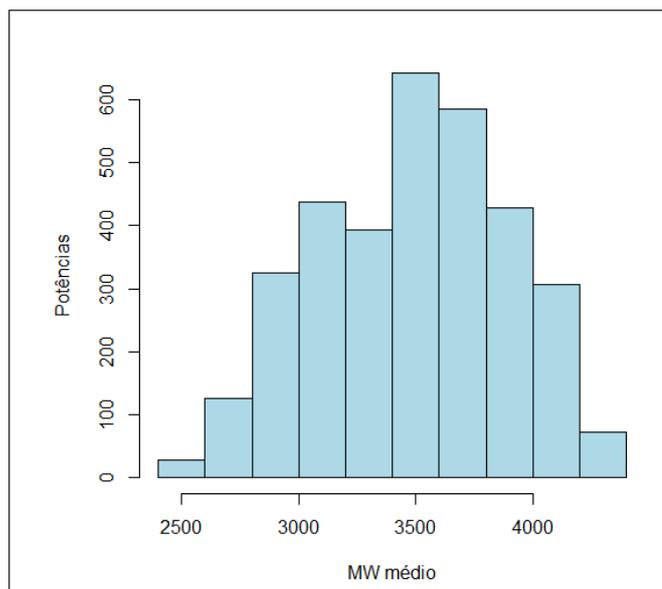


Gráfico 4.9: Histograma do subsistema Norte
Fonte: ONS, elaboração própria.

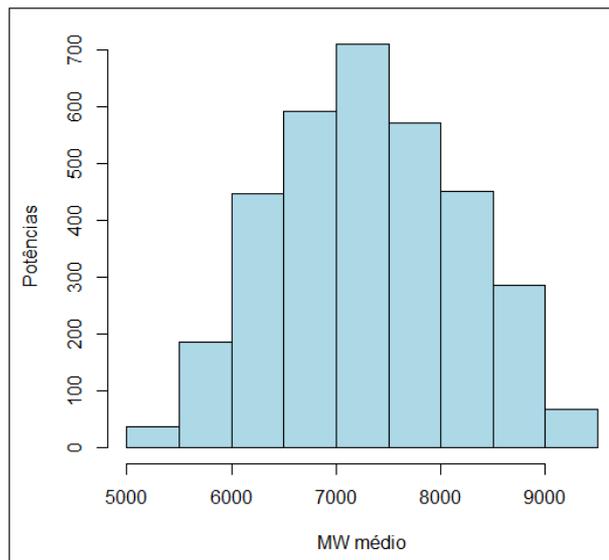


Gráfico 4.10: Histograma do subsistema Nordeste
Fonte: ONS, elaboração própria.

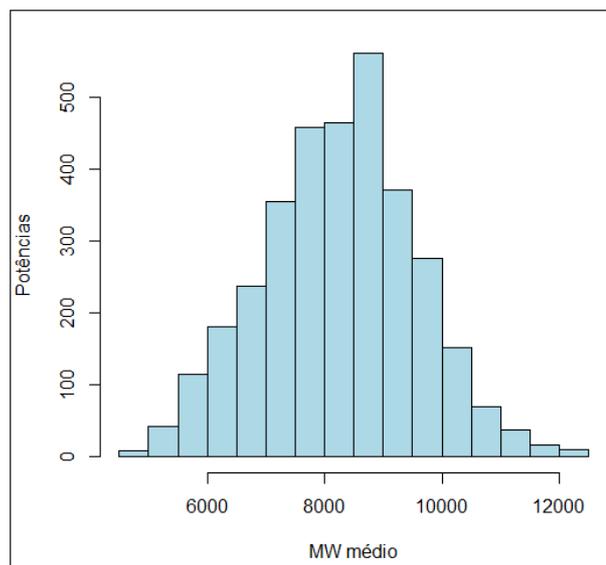


Gráfico 4.11: Histograma do subsistema Sul
Fonte: ONS, elaboração própria.

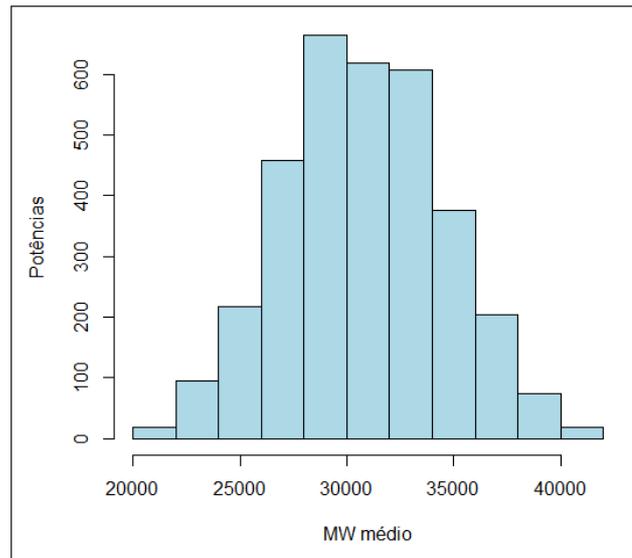


Gráfico 4.12: Histograma do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

Através da análise dos gráficos de Box-plot e Histograma das séries diárias de Carga, é notado que apenas o subsistema Norte não apresenta simetria nos dados. Outro ponto importante é a presença de outliers nas séries do subsistema Sul e Sudeste / Centro-Oeste, além de que apenas a série do subsistema Sul parece seguir uma distribuição normal.

A tabela 4.2 mostra o teste de normalidade de Jarque-Bera para as séries diárias de Carga para os quatro subsistemas. Este teste confirma a hipótese inicial de que apenas o subsistema Sul segue uma distribuição normal, pois foi o único que apresentou um p-valor alto, isto é, maior do que o nível de significância usual de 5%, não rejeitando a hipótese nula de normalidade.

Tabela 4.2: Teste de Jarque-Bera para os quatro subsistemas

Subsistemas	Est. Teste	P-valor	Decisão
Norte	95,05	0,00	Rejeita H_0
Nordeste	66,63	0,00	Rejeita H_0
Sul	4,72	0,09	Não Rejeita H_0
Sudeste / Centro-Oeste	18,23	0,00	Rejeita H_0

Fonte: ONS, elaboração própria.

4.2. Retirada do efeito Calendário

Para prever e modelar a carga de um dia qualquer é preciso saber, em primeiro lugar, qual dia da semana está sendo tratado. A curva de carga de cada

dia da semana é diferente. A madrugada de segunda-feira, por exemplo, costuma apresentar uma carga baixíssima, semelhante à carga da mesma hora de um domingo. Já a madrugada de terça costuma ter uma carga em um patamar maior do que a de segunda, sugerindo que há processos industriais que operam em três turnos, mas não durante as noites do fim de semana. Isso faz com que em geral a carga da terça seja maior que a carga da segunda, que, por sua vez é sempre bem maior que a carga de um domingo.

Aplicando a metodologia descrita no capítulo 3, será retirado o efeito de sazonalidade semanal das séries. A tabela 4.3 mostra a média da série diária de Carga e a média da série diária da Carga com ajuste calendário para cada dia da semana. A diferença entre a média de Carga de um domingo e a média de Carga de uma segunda-feira no subsistema Sudeste / Centro-Oeste, por exemplo, é de 19%. Após a aplicação da metodologia, as diferenças entre os dias da semana são bem próximas de 0%.

Tabela 4.3: Média de Carga com e sem ajuste calendário por dia da semana

Dia da Semana	Norte		Nordeste		Sul		Sudeste / Centro-Oeste	
	Carga sem ajuste	Carga com ajuste	Carga sem ajuste	Carga com ajuste	Carga sem ajuste	Carga com ajuste	Carga sem ajuste	Carga com ajuste
Domingo	3.302	3.490	6.565	7.346	6.430	8.327	26.435	31.003
Segunda-feira	3.524	3.490	7.410	7.346	8.536	8.324	31.413	30.995
Terça-feira	3.544	3.492	7.558	7.348	8.840	8.324	32.212	30.997
Quarta-feira	3.544	3.490	7.558	7.344	8.886	8.337	32.344	31.034
Quinta-feira	3.546	3.494	7.585	7.365	8.909	8.350	32.420	31.082
Sexta-feira	3.546	3.494	7.565	7.354	8.873	8.338	32.381	31.054
Sábado	3.432	3.491	7.163	7.342	7.722	8.323	29.692	31.012
Semi-feriado	3.459	3.504	7.166	7.452	7.573	8.415	28.928	30.710
Feriado	3.364	3.534	6.675	7.469	6.707	8.403	27.301	31.515

Fonte: ONS, elaboração própria.

Os gráficos 4.13, 4.14, 4.15 e 4.16 mostram as séries diárias de Carga com e sem ajuste calendário para cada subsistema. As séries geradas pela aplicação do ajuste calendário são, perceptivelmente, menos voláteis do que as séries originais, isto porque a maior causa da oscilação da carga diária é o fator calendário.

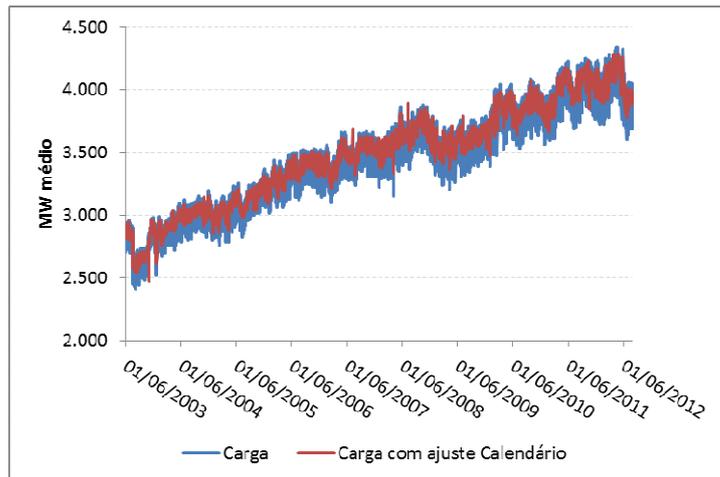


Gráfico 4.13: Série de Carga com e sem ajuste calendário do subsistema Norte

Fonte: ONS, elaboração própria.

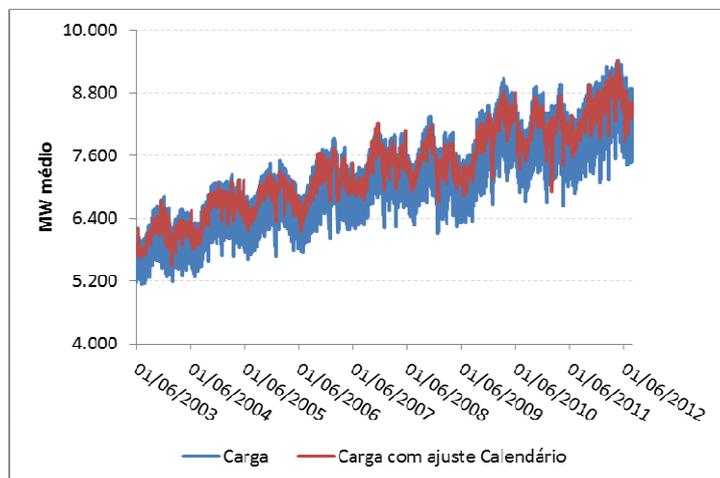


Gráfico 4.14: Série de Carga com e sem ajuste calendário do subsistema Nordeste

Fonte: ONS, elaboração própria.

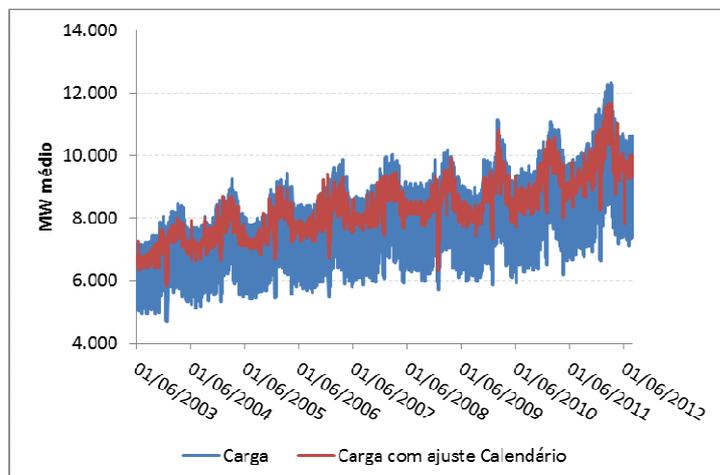


Gráfico 4.15: Série de Carga com e sem ajuste calendário do subsistema Sul

Fonte: ONS, elaboração própria.

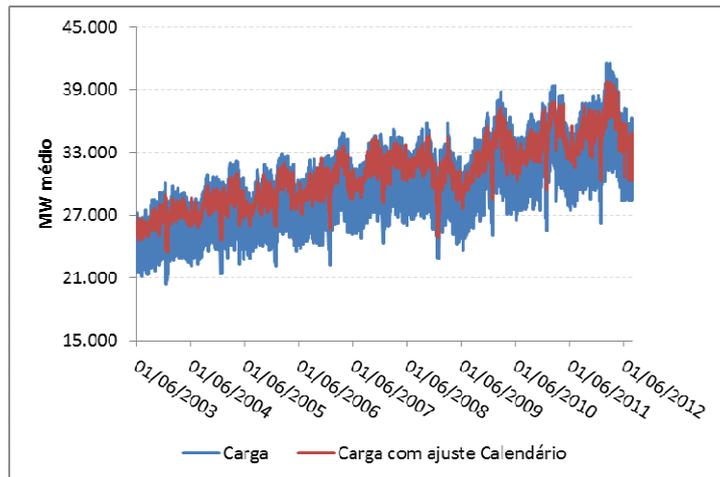


Gráfico 4.16: Série de Carga com e sem ajuste calendário do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

4.3.Retirada do efeito Temperatura

Para retirar o efeito de Temperatura dos dados diários de Carga com ajuste de calendário, serão utilizadas duas metodologias diferentes (propostas no capítulo 3) e ao final será feita uma comparação entre as mesmas.

Nesta etapa não será ajustada a série de Carga do subsistema Norte, isto porque estudos anteriores comprovaram não haver influencia da variação de temperatura na variável Carga.

4.3.1.Método Empírico (Multiplicativo)

Para este método, antes de ser aplicada a metodologia proposta, as séries de Carga com ajuste calendário, para os três subsistemas, foram divididas em duas séries menores, a primeira variando de 01/06/03 até 31/12/07 e a outra variando de 01/01/08 até 31/07/12. Isto é, para cada subsistema, serão feitas duas modelagens, uma para cada parte dos dados, e depois estas séries ajustadas com efeito temperatura serão juntadas novamente.

A tabela 4.4 mostra os coeficientes de regressão, a partir de variáveis derivadas da Carga com ajuste calendário e Temperatura, para todos os meses de cada período de cada subsistema.

Tabela 4.4: Coeficientes de regressão de cada período de cada subsistema por mês

Subsistema	Período	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Nordeste	2003-2007	1,14	2,11	1,53	0,89	1,53	1,36	0,60	1,20	0,85	1,04	2,39	1,95
	2008-2012	1,00	2,04	1,34	1,47	1,48	1,32	0,64	0,58	0,81	1,08	1,52	1,68
Sul	2003-2007	1,70	1,12	1,68	1,33	0,62	0,33	0,29	0,33	0,42	0,90	1,15	1,21
	2008-2012	1,77	1,69	1,43	1,24	0,60	0,18	0,00	0,10	0,58	0,65	1,16	1,41
Sudeste / Centro-Oeste	2003-2007	0,85	0,92	1,05	0,95	0,90	0,69	0,48	0,56	0,66	0,81	0,79	0,86
	2008-2012	0,93	0,95	1,18	1,23	0,91	0,73	0,52	0,56	0,62	0,85	1,05	0,96

Fonte: Castro *et al.* (2012).

Os gráficos 4.17, 4.18 e 4.19 mostram as séries diárias de Carga com ajuste calendário e as séries diárias de Carga com ajuste calendário e Temperatura, para cada um dos três subsistemas. Notam-se pequenas diferenças entre as séries diárias de Carga, com ajuste Calendário, com e sem ajuste Temperatura.

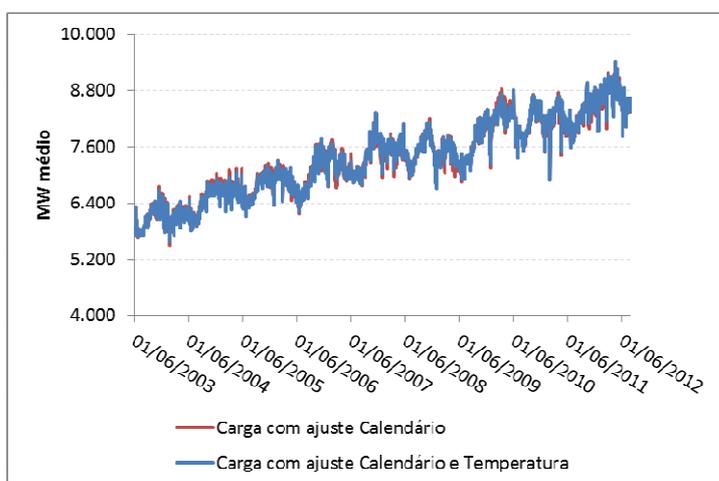


Gráfico 4.17: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Nordeste

Fonte: ONS, elaboração própria.

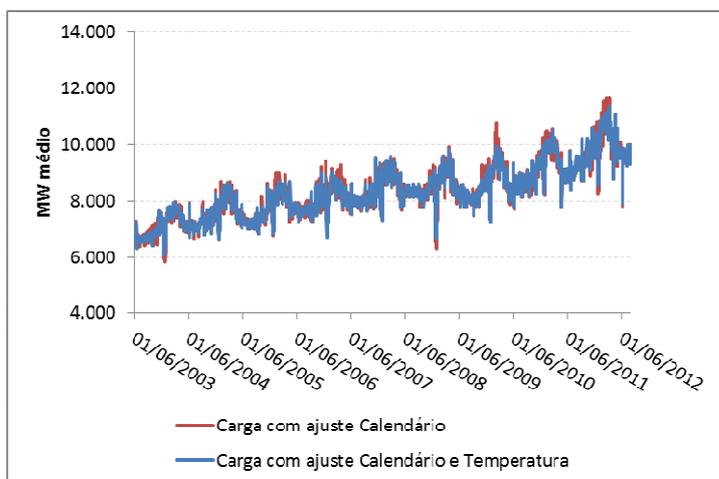


Gráfico 4.18: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sul

Fonte: ONS, elaboração própria.

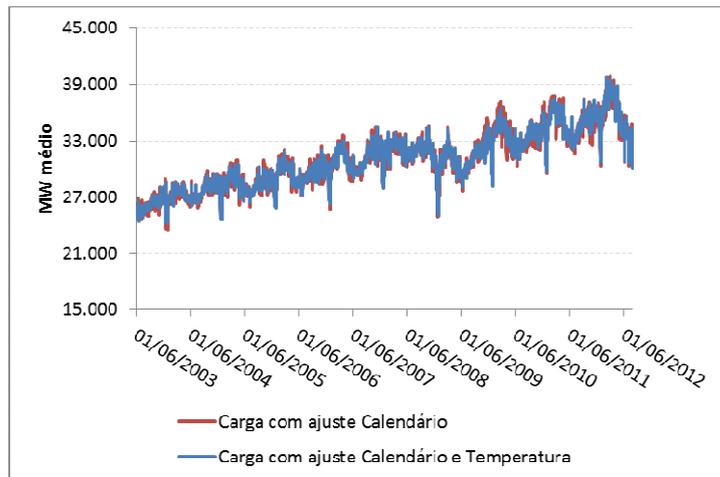


Gráfico 4.19: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

4.3.2. Método Regressão Dinâmica (Aditivo)

Mais detalhes e resultados das modelagens via Regressão Dinâmica se encontram no ANEXO A.

Para este método, será utilizada a série de Carga diária com ajuste Calendário, no período do dia 01/06/03 até o dia 31/07/12, para cada um dos três subsistemas.

As equações de ajustes para os subsistemas Nordeste, Sul e Sudeste se encontram abaixo, respectivamente:

$$\begin{aligned} \text{CARGA_adjT}_{t|\text{NE}}^{r,d,m,a} &= \text{CARGA_adj}_{t|\text{NE}}^{r,d,m,a} - 59,21 \cdot \text{Temp}_{t|\text{NE}}^{r,d,m,a} \\ &+ 12,49 \cdot \text{Temp}_{t-2|\text{NE}}^{r,d,m,a} + 22,62 \cdot \text{Temp}_{t-5|\text{NE}}^{r,d,m,a} \\ &+ 23,13 \cdot \text{Temp}_{t-8|\text{NE}}^{r,d,m,a} \end{aligned} \quad (4.1)$$

$$\begin{aligned} \text{CARGA_adjT}_{t|\text{S}}^{r,d,m,a} &= \text{CARGA_adj}_{t|\text{S}}^{r,d,m,a} - 45,53 \cdot \text{Temp}_{t|\text{S}}^{r,d,m,a} \\ &+ 25,56 \cdot \text{Temp}_{t-1|\text{S}}^{r,d,m,a} + 8,56 \cdot \text{Temp}_{t-3|\text{S}}^{r,d,m,a} \end{aligned} \quad (4.2)$$

$$\begin{aligned}
 \text{CARGA_adj}T_{t|SE}^{r,d,m,a} &= \text{CARGA_adj}j_{t|SE}^{r,d,m,a} - 159,97 \cdot \text{Temp}_{t|SE}^{r,d,m,a} \\
 &+ 47,86 \cdot \text{Temp}_{t-2|SE}^{r,d,m,a} + 45,01 \cdot \text{Temp}_{t-3|SE}^{r,d,m,a} \\
 &+ 31,27 \cdot \text{Temp}_{t-5|SE}^{r,d,m,a} + 26,85 \cdot \text{Temp}_{t-7|SE}^{r,d,m,a}
 \end{aligned}
 \tag{4.3}$$

Os gráficos 4.20, 4.21 e 4.22 mostram as séries diárias de Carga com ajuste calendário e as séries diárias de Carga com ajuste calendário e Temperatura, para cada um dos três subsistemas. Novamente, notam-se pequenas diferenças entre as séries diárias ajustadas de Carga.

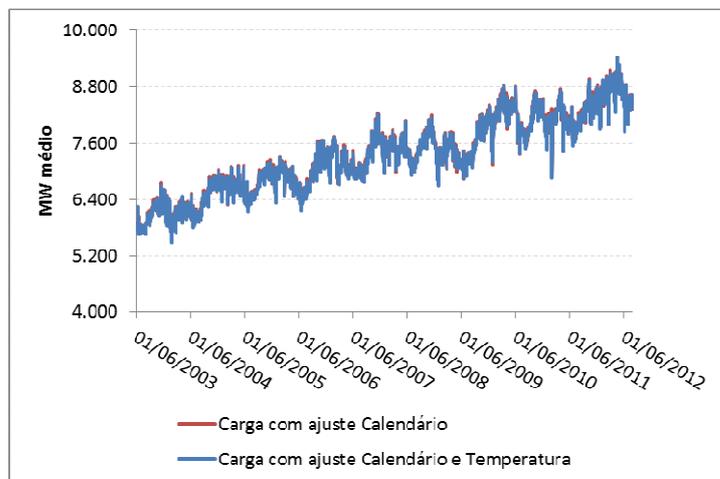


Gráfico 4.20: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Nordeste

Fonte: ONS, elaboração própria.

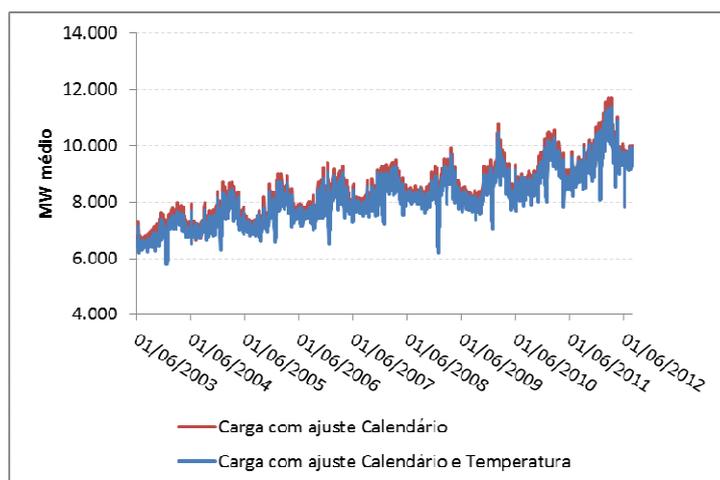


Gráfico 4.21: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sul

Fonte: ONS, elaboração própria.

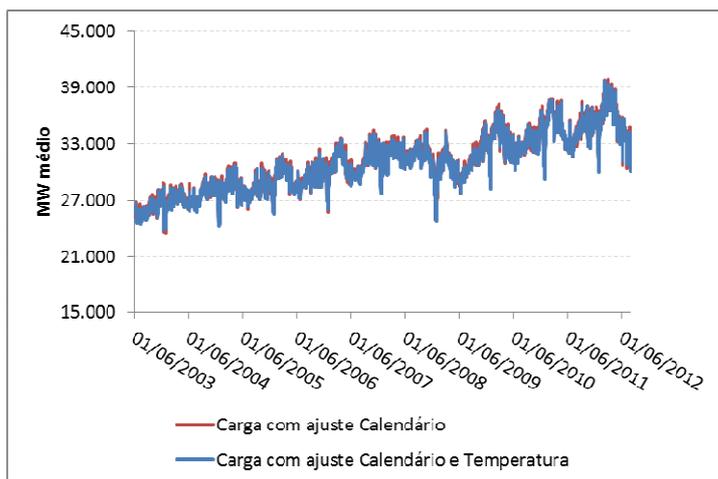


Gráfico 4.22: Série de Carga com ajuste Calendário e série de Carga com ajuste Calendário e Temperatura do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

4.4. Retirada do efeito Semana Morta

O ajuste da Carga diária tem ainda um fator a ser analisado, que é o ajuste para a semana de festas de fim de ano. O ajuste consiste em aplicar à carga dos dias da semana de festas de fim de ano um coeficiente de ajuste baseado na média dos desvios de Carga em dias de semana de festas com relação à Carga dos mesmos dias da semana em semanas normais. Novamente, este ajuste não será feito para o subsistema Norte.

A análise das tabelas 4.5 e 4.6, que informam a média de Carga com ajuste Calendário e Temperatura por dia da semana em uma semana comum e em uma semana morta, comprova o que foi dito anteriormente, isto é, a média de Carga de um feriado em uma semana morta tem valor bem menor do que a média de Carga de um feriado em uma semana comum. Este comportamento é notado tanto para a retirada do efeito Temperatura pela metodologia Empírica quanto para a metodologia de Regressão Dinâmica.

Tabela 4.5: Média de Carga com ajuste Calendário e Temperatura utilizando o Método Empírico, para uma semana comum e uma semana de festas

Dia da Semana	Nordeste		Sul		Sudeste / Centro-Oeste	
	Semana Comum	Semana Morta	Semana Comum	Semana Morta	Semana Comum	Semana Morta
Domingo	7.347	7.257	8.330	7.887	31.024	28.915
Segunda-feira	7.350	7.211	8.336	7.520	31.036	29.316
Terça-feira	7.353	7.129	8.337	7.604	31.041	28.959
Quarta-feira	7.347	7.155	8.332	7.982	31.052	30.080
Quinta-feira	7.362	7.398	8.353	7.894	31.091	29.629
Sexta-feira	7.353	7.345	8.342	7.643	31.080	28.703
Sábado	7.343	7.277	8.333	7.465	31.065	27.600
Semi-feriado	7.449	7.496	8.717	7.702	31.409	28.737
Feriado	7.547	7.064	8.620	7.486	32.221	27.707

Fonte: ONS, elaboração própria.

Tabela 4.6: Média de Carga com ajuste Calendário e Temperatura utilizando Regressão Dinâmica, para uma semana comum e uma semana de festas

Dia da Semana	Nordeste		Sul		Sudeste / Centro-Oeste	
	Semana Comum	Semana Morta	Semana Comum	Semana Morta	Semana Comum	Semana Morta
Domingo	7.328	7.218	8.125	7.705	30.847	28.754
Segunda-feira	7.328	7.196	8.134	7.234	30.858	28.966
Terça-feira	7.332	7.115	8.132	7.349	30.868	28.675
Quarta-feira	7.326	7.120	8.123	7.760	30.864	29.779
Quinta-feira	7.341	7.364	8.143	7.608	30.907	29.335
Sexta-feira	7.332	7.302	8.132	7.422	30.893	28.506
Sábado	7.321	7.240	8.126	7.185	30.879	27.423
Semi-feriado	7.418	7.454	8.465	7.417	31.116	28.536
Feriado	7.518	7.033	8.412	7.191	32.000	27.517

Fonte: ONS, elaboração própria.

Após a aplicação da metodologia para corrigir o efeito Semana Morta, a diferença entre a média de Carga de um feriado em uma semana comum e um feriado em uma semana morta diminuiu bastante. No subsistema Sul, por exemplo, considerando a média entre as duas metodologias, o desvio da Carga de um feriado em uma semana comum em relação a uma semana de festas foi reduzido de 16,1% para 3,2%.

Os gráficos 4.23, 4.24 e 4.25 introduzem as séries com ajuste Calendário, Temperatura pelo método Empírico e Semana Morta, para os três subsistemas. Já os gráficos 4.26, 4.27 e 4.28 mostram as séries com ajuste Calendário, Temperatura pelo método de Regressão Dinâmica e Semana Morta, para os três subsistemas. Elas são idênticas às séries anteriores em quase todo o ano,

discrepando apenas nos dias de semana de festas de fim de dezembro e início de janeiro.

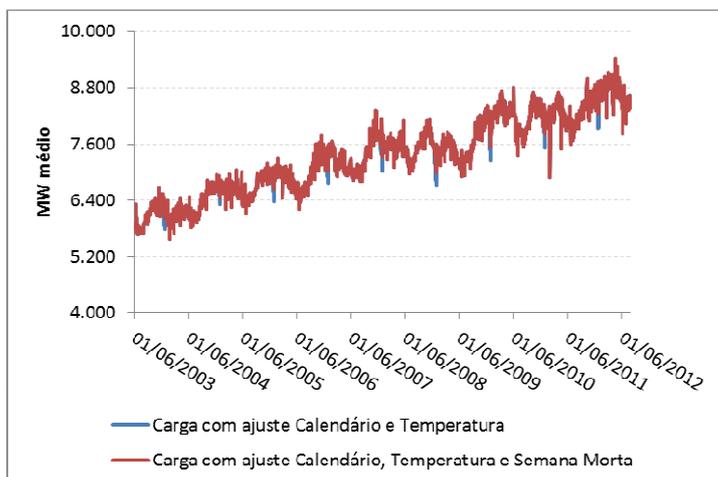


Gráfico 4.23: Série de Carga com ajuste Calendário, Temperatura utilizando o Método Empírico e com e sem ajuste Semana Morta para o subsistema Nordeste
Fonte: ONS, elaboração própria.

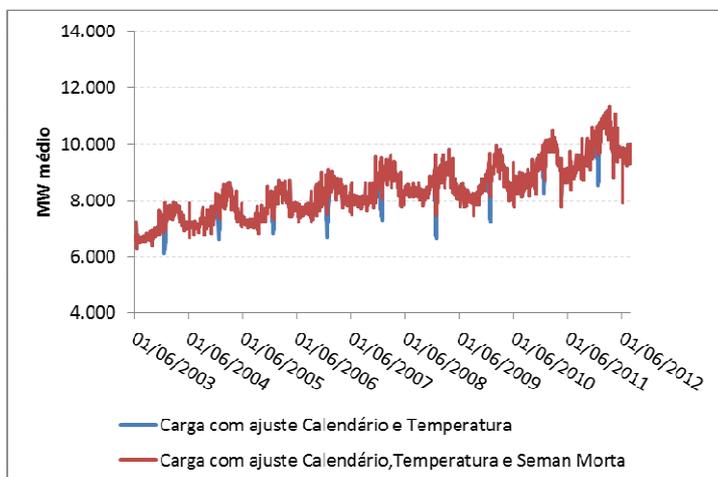


Gráfico 4.24: Série de Carga com ajuste Calendário, Temperatura utilizando o Método Empírico e com e sem ajuste Semana Morta para o subsistema Sul
Fonte: ONS, elaboração própria.

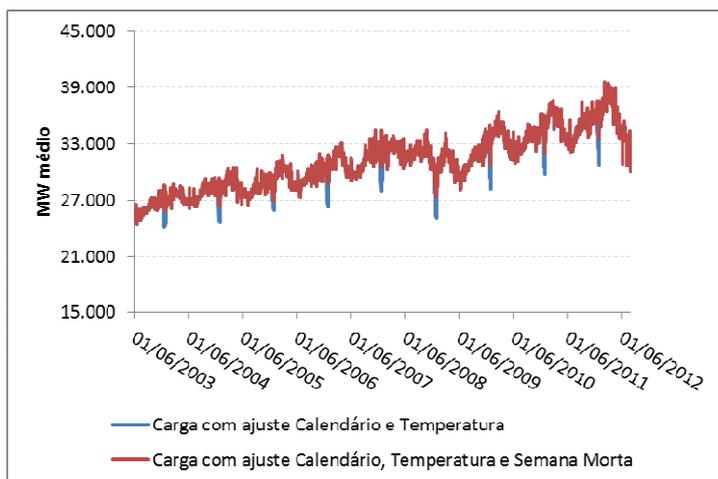


Gráfico 4.25: Série de Carga com ajuste Calendário, Temperatura utilizando o Método Empírico e com e sem ajuste Semana Morta para o subsistema Sudeste / Centro-Oeste

Fonte: ONS, elaboração própria.

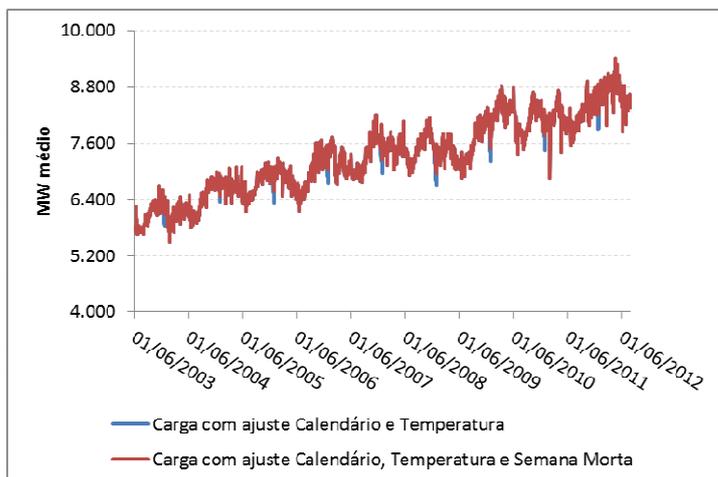


Gráfico 4.26: Série de Carga com ajuste Calendário, Temperatura utilizando Regressão Dinâmica e com e sem ajuste Semana Morta para o subsistema Nordeste

Fonte: ONS, elaboração própria.

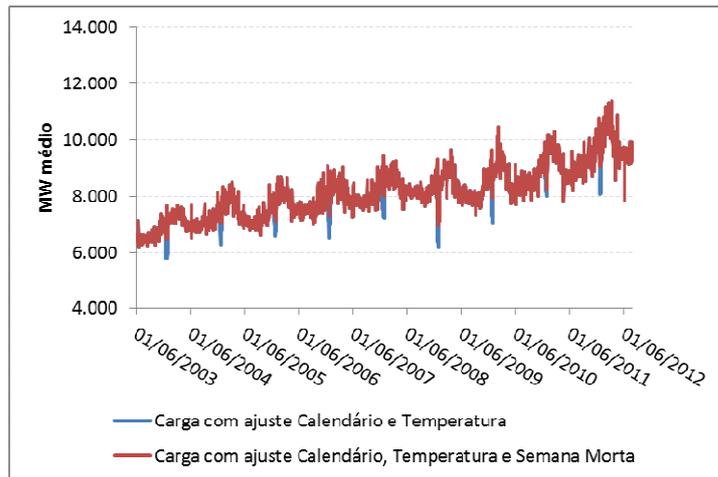


Gráfico 4.27: Série de Carga com ajuste Calendário, Temperatura utilizando Regressão Dinâmica e com e sem ajuste Semana Morta para o subsistema Sul
Fonte: ONS, elaboração própria.

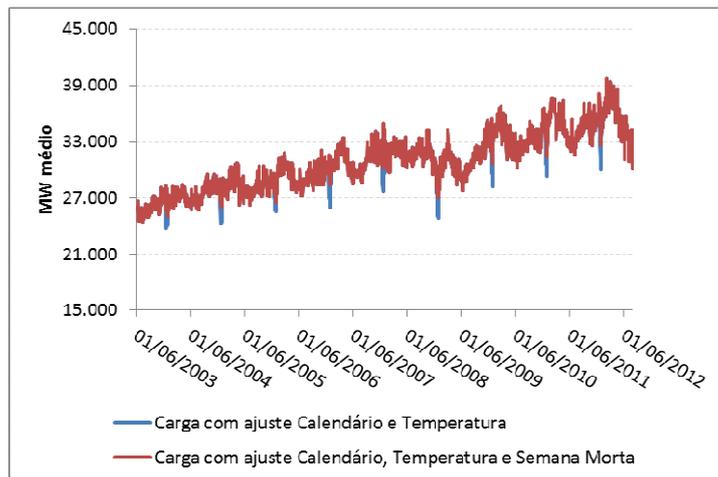


Gráfico 4.28: Série de Carga com ajuste Calendário, Temperatura utilizando Regressão Dinâmica e com e sem ajuste Semana Morta para o subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

É possível observar que a série diária ajustada tem um padrão sazonal no ano bastante regular para o subsistema Sul e Nordeste, e a partir do ano de 2009 também é notado este padrão sazonal para o subsistema Sudeste / Centro-Oeste. Há um aumento da carga no início do ano, associado ao calor do verão e à uma retomada da indústria depois do bimestre sazonalmente fraco de dezembro e janeiro. A carga no meio do ano é menor, em boa medida devido ao menor consumo com refrigeração. No meio do segundo semestre há um novo aumento da carga, fruto do pico da produção industrial.

4.5. Obtenção das séries Mensais

A série mensal ajustada é uma série que leva em conta tanto os dados disponíveis em frequência diária como os dados mensais de geração de pequenas centrais, descontinuidades de medição e perdas na rede básica.

As séries mensais serão obtidas utilizando-se das metodologias propostas no capítulo 3. Após aplicar os respectivos fatores e correções nas séries diárias, serão obtidas as séries mensais de cada subsistema, que estão representadas nos gráficos 4.29, 4.30, 4.31 e 4.32.

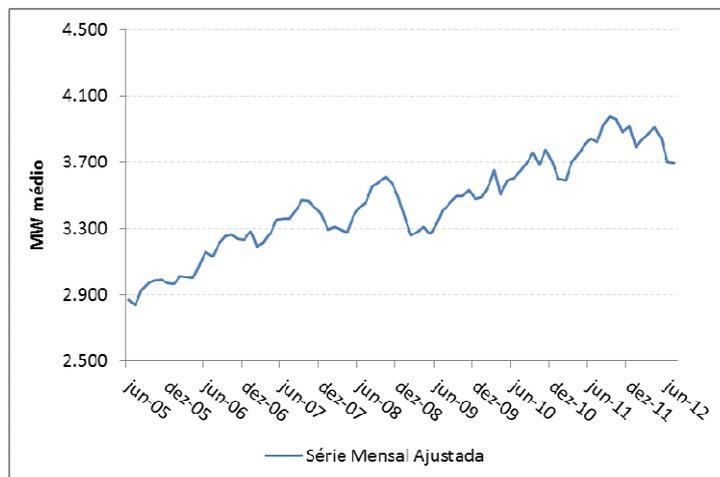


Gráfico 4.29: Série Mensal de Carga ajustada do subsistema Norte
Fonte: ONS, elaboração própria.

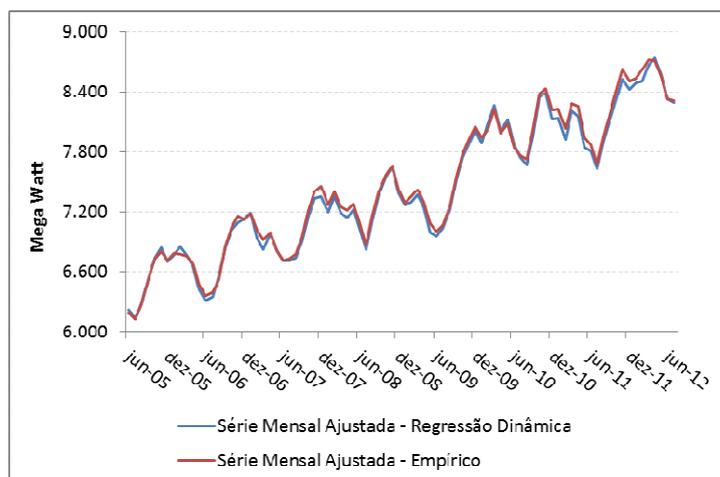


Gráfico 4.30: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do subsistema Nordeste
Fonte: ONS, elaboração própria.

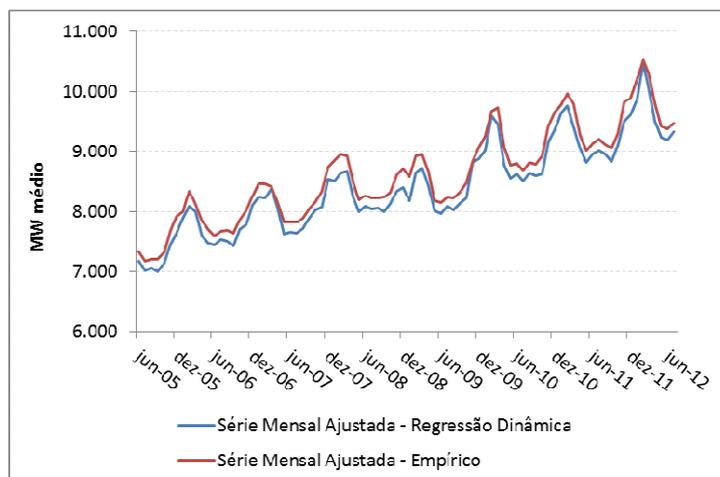


Gráfico 4.31: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do subsistema Sul
Fonte: ONS, elaboração própria.

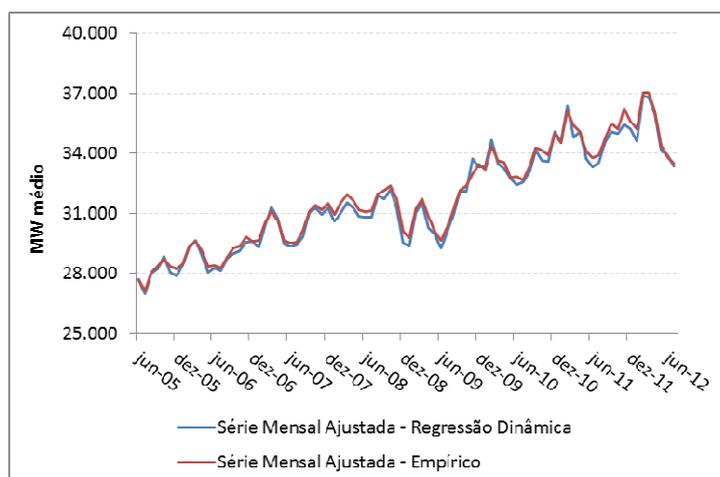


Gráfico 4.32: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

Para o subsistema Norte há apenas uma série Mensal de Carga ajustada, isto porque não foi feito ajuste Temperatura para este subsistema, apenas o ajuste Calendário.

Percebe-se, através dos gráficos anteriores, que as metodologias de Regressão Dinâmica e Método Empírico geram séries Mensais de Carga próximas para os subsistemas Nordeste e Sudeste / Centro-Oeste, enquanto que para o subsistema Sul, as séries Mensais de Carga ajustadas não são parecidas. O método Empírico, para retirada do efeito de Temperatura na série Diária, gerou

uma série Mensal de Carga estritamente maior do que o modelo de Regressão Dinâmica.

4.6. Obtenção da Série Mensal do Sistema Interligado Nacional (SIN)

Para obter a série do SIN, basta somar as séries Mensais de Carga Ajustada dos quatro subsistemas. Dessa forma, o gráfico 4.33 mostra a série Mensal do SIN das duas metodologias utilizadas na retirada do efeito de Temperatura na Série Diária.

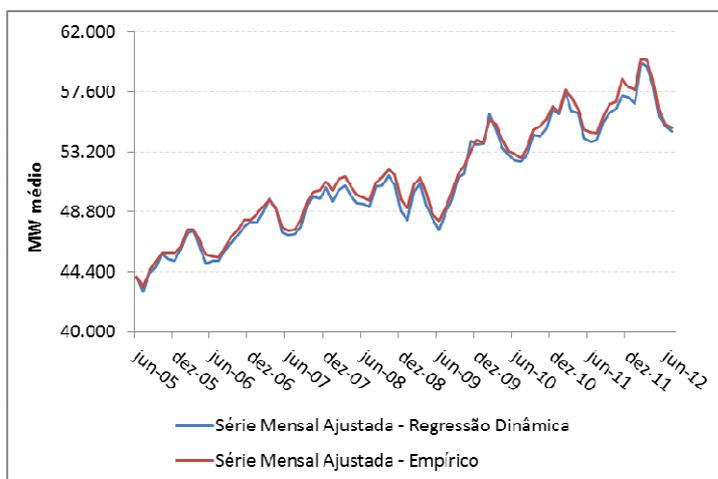


Gráfico 4.33: Série Mensal de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do SIN
Fonte: ONS, elaboração própria.

4.7. Obtenção da Série Trimestral do SIN

Finalmente, a obtenção da série Trimestral do SIN, objetivo deste trabalho para comparação com a série trimestral do PIB, é obtida com a média de Carga Mensal (conforme capítulo 3).

Os gráficos 4.34 e 4.35 mostram, respectivamente, a série Trimestral do SIN e a variação de Carga trimestral para a Carga do mesmo trimestre do ano anterior, para as duas metodologias utilizadas na retirada do efeito de Temperatura na Série Diária.

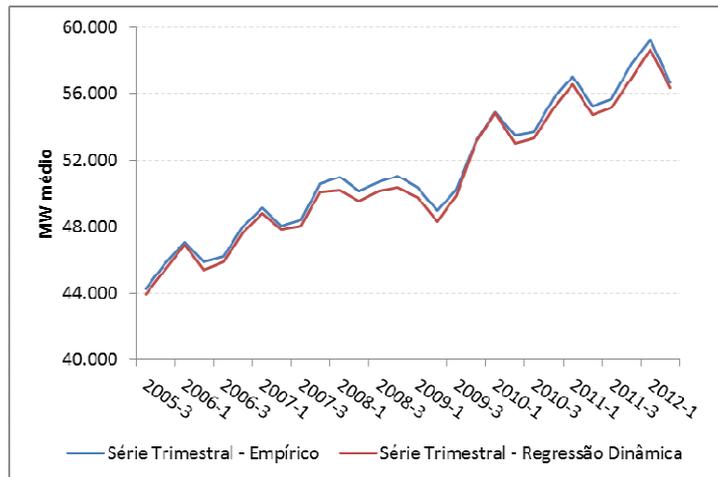


Gráfico 4.34: Série Trimestral de Carga ajustada, utilizando Regressão Dinâmica e Método Empírico para retirada do efeito Temperatura, do SIN
Fonte: ONS, elaboração própria.

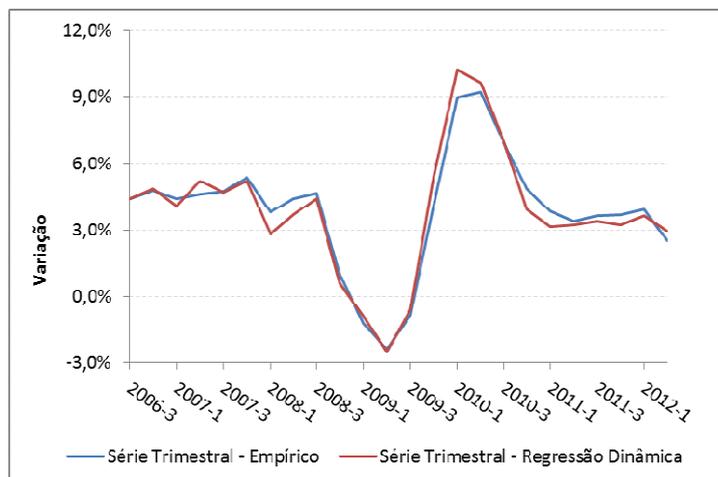


Gráfico 4.35: Séries Trimestrais de variação de Carga ajustada em relação ao mesmo trimestre do ano anterior
Fonte: ONS, elaboração própria.

4.8. Comparação dos Resultados

Neste ponto do trabalho, entende-se que, com exceção das variáveis econômicas, todas as variáveis que influenciam a Carga foram retirados. Desta forma, as séries Trimestrais geradas anteriormente são séries “limpas” em relação aos efeitos Calendário, Temperatura e Semana Morta. Com isso, para efeito de comparação entre a metodologia empírica feita inicialmente e a modelagem de Regressão Dinâmica, proposta para retirada do efeito de Temperatura, a série Trimestral do PIB será utilizada como indicador econômico.

A série de Carga Trimestral que mais acompanhar as variações do PIB e/ou melhor explicar o mesmo, será considerada melhor ajustada do que a outra.

A variação de um trimestre para o mesmo trimestre do ano anterior da série do PIB e da série de Carga sem ajuste será acrescentada no gráfico 4.35, gerando o gráfico 4.36.

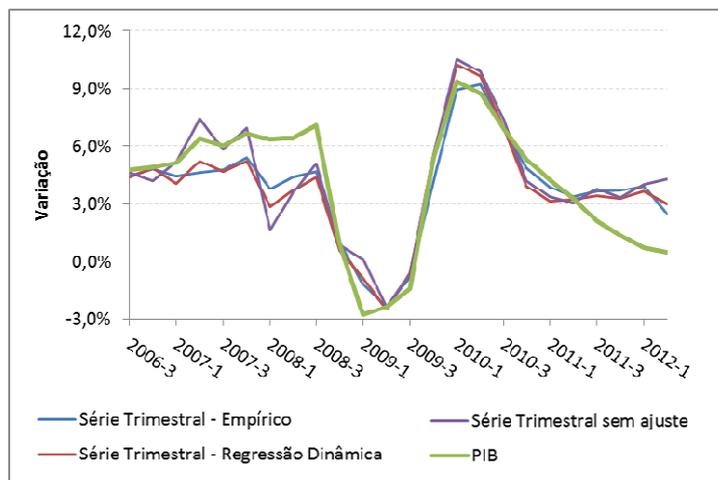


Gráfico 4.36: Séries Trimestrais de variação de Carga ajustada e PIB em relação ao mesmo trimestre do ano anterior

Fonte: ONS, elaboração própria.

Através do gráfico 4.36, pode-se perceber que as séries trimestrais de variação de Carga ajustada em relação ao mesmo trimestre do ano anterior seguem um comportamento parecido com a série trimestral de variação do PIB em relação ao mesmo trimestre do ano anterior. Os coeficientes de correlação entre as séries de variações de carga do trimestre com o mesmo trimestre do ano anterior e a série de variações de PIB do trimestral com o mesmo trimestre do ano anterior foram calculados e podem ser vistos na tabela 4.7. A série de Carga trimestral ajustada por Regressão Dinâmica apresentou um coeficiente de correlação de 92,0% enquanto que a série de Carga trimestral ajustada pelo Método Empírico teve um coeficiente de correlação de 90,7%. Enquanto que a série de Carga trimestral sem ajustes apresentou um coeficiente de correlação de 83,6% com a série trimestral do PIB. Desta forma, é mostrado que os ajustes Calendário, Temperatura e Semana Morta adicionaram um ganho de 10% em relação ao coeficiente de correlação com a série trimestral do PIB, evidenciando que estes

ajustes foram realizados com sucesso na tentativa de excluir fatores fortuitos que influenciam a série de Carga.

Tabela 4.7: Correlação entre as séries trimestrais de variação de Carga com a série trimestral de variação do PIB

Método	Correlação
Carga com Ajuste Empírico	90,7%
Carga com Ajuste Reg. Din.	92,0%
Carga sem Ajuste	83,6%

Outro artifício de comparação é explicar a série trimestral do PIB pelas séries trimestrais de Carga ajustadas. Isto é, a série do PIB será modelada, via Regressão Dinâmica, pela Carga ajustada e suas defasagens. Não serão incluídas no modelo as defasagens do PIB e do Erro. O critério de seleção escolhido será o valor do coeficiente de determinação (R^2). Ou seja, a série de Carga Trimestral ajustada que apresentar o maior valor de R^2 na modelagem, isto quer dizer que ajustou melhor os valores do PIB, será considerado a série com melhor ajuste.

O R^2 ajustado com a série de Carga trimestral ajustada por regressão dinâmica como regressora da variável PIB foi de 89,58%, enquanto que o R^2 ajustado com a série de Carga trimestral ajustada pelo método empírico como regressora da variável PIB foi de 89,55%. Em relação ao ajuste, é notado que ambas as metodologias “limparam” a série de Carga dos ofensores não econômicos e que não há diferença entre estas metodologias.

5 Conclusão e Trabalhos Futuros

Sabendo que as séries de Carga são influenciadas por diversos fatores, entre eles, efeitos de calendário, de temperatura e da economia, a série de Carga Mensal Ajustada e a série Trimestral de Variação de Carga Ajustada em relação ao mesmo trimestre do ano anterior apresentadas no capítulo 4 são muito importantes para dar embasamento e apoio à decisão das empresas que operam na compra e venda da energia elétrica nos leilões e para auxiliar Instituições que façam a previsão Trimestral do PIB, respectivamente. Para realizar os ajustes, foram utilizados dados diários de carga e temperatura, disponibilizados pelo ONS.

Para tais ajustes, foram utilizadas duas abordagens diferentes, um método totalmente empírico para retirada dos ofensores de Temperatura e Calendário nas séries diárias e mensais e uma outra abordagem com os mesmos ajustes para a série mensal e diária para o efeito Calendário, a exceção é no ajuste para retirada do efeito temperatura na série diária. O ajuste de temperatura na série diária na segunda abordagem é baseado em uma modelagem estatística de regressão dinâmica.

Após estas correções, entende-se que apenas o ofensor econômico influencia a série de Carga. Dessa forma, para comparação entre as abordagens, as séries mensais de carga foram consolidadas para frequência trimestral e comparadas com a série trimestral do PIB. Foi decidido que a série de carga trimestral que melhor ajustar a série do PIB e apresentar o maior coeficiente de correlação entre estas séries, será a série melhor ajustada em relação aos efeitos de calendário e temperatura. Para ajustar a série do PIB foi utilizado o modelo de regressão dinâmica apenas com a variável de Carga e suas defasagens como regressoras.

Foi notado que a diferença entre os ajustes é mínima e ambas as séries tiveram um ótimo poder de explicação para a série do PIB (valores de R^2 próximos de 90%). Vale observar que não foram incluídas defasagens do Erro e nem da série do PIB nas modelagens, mostrando que as séries de Carga Ajustada e

PIB são fortemente correlacionadas. O coeficiente de correlação entre a série de carga ajustada pela segunda abordagem e a série do PIB foi um pouco maior do que o coeficiente de correlação entre a série de carga ajustada pela primeira abordagem e a série do PIB. Os coeficientes foram 92,0% e 90,7%, respectivamente. Mais uma vez é mostrado a pouca diferença entre os ajustes. Além disso, o coeficiente de correlação entre a série de carga sem ajuste e a série do PIB também foi calculado, 83,6%. Isso mostra que ambas as abordagens geraram séries “limpas” das variações de calendário e temperatura.

No ajuste do fator de Temperatura, uma pequena vantagem da abordagem deste trabalho em relação à metodologia empírica foi a não divisão da série de Carga em duas séries menores, de forma que a metodologia proposta nesta dissertação realizou apenas uma modelagem nos dados. Outra vantagem é o embasamento estatístico proveniente dos modelos de Regressão Dinâmica.

Sugere-se como trabalho futuro, uma formalização através de modelos estatísticos do efeito calendário, pois nesta dissertação utilizou-se integralmente a proposta descrita em Castro *et al.* (2012). Merece também investigação outras formas de avaliar o desempenho das abordagens, pois neste trabalho utilizou-se a série trimestral do PIB, isso implicou na agregação trimestral das séries de carga livre dos ofensores, o ideal seria a comparação com a série de carga mensal filtrada dos efeitos calendário e temperatura.

Referências Bibliográficas

ARAÚJO, V. S.; **Sistemática dos leilões de energia**. Monografia, DEE, UFRJ, janeiro 2007.

BARROS, M.; SOUZA, R.C. **Regressão Dinâmica**. Núcleo de Estatística Computacional. PUCRio, 1995.

BERA, A.; JARQUE, C. **Efficient test for normality, heterocedasticity and serial independence of regression residuals**. *Econometrics Letters*, 6, pp. 255–259, 1980.

BOX, G. E. P. and PIERCE, D. A., **Distribution of residual correlations in autoregressive-integrated moving average time series models**. *Journal of the American Statistical Association*, 65, pp. 1509-1526, 1970.

CASTRO, N. J; LEITE, A.L.S. **Estrutura de Governança e a Formação de Holdings no Setor Elétrico Brasileiro**. *Estratégia e Negócios*, Florianópolis, v.1, n.2, pp. 8, 2008.

CASTRO, N. J. *et al*; **Pesquisa sobre Previsão de Carga com Base em Cenários Macro Econômicos**. Grupo de Estudos do Setor Elétrico (GESEL IE-UFRJ), Dezembro 2012.

COCHRANE, D.; ORCUTT, G.H. **Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms**. *Journal of the American Statistical Association*, 44, p. 32-41, 1949.

CORREIA, T.; MELO, E.; COSTA, A.; SILVA, A.; **Trajetória das reformas institucionais da indústria elétrica brasileira e novas perspectivas de mercado**. *Revista de Economia* 7 (3): pp. 607-627, 2006.

DURBIN, J.; WATSON G. S. **Testing for serial correlation in least squares regression I**. *Biometrika*, London, v.37, n.3/4, p. 409-428, 1950.

GUJARATI, D. **Econometria Básica**. Campus, Rio de Janeiro, 4th edition, pp. 786, 2006.

LEITE, A.L.S. & SANTANA, E.A.; **Mercado de capacidade: uma alternativa para o setor elétrico brasileiro**. *Revista de Desenvolvimento Econômico RDE*, Ano VIII, n. 14, p.23-33, 2006.

LJUNG, G. M. and BOX, G. E. P., **On a measure of lack of fit in time series models**. *Biometrika* 65, pp. 297-303, 1978.

MEDEIROS, M.C. **Técnicas de Pesquisa em Economia**, Notas de Aula, 2005.

MIRANDA, C.V.; SOUZA, R.C. **Previsão de dados de alta frequência para carga elétrica usando holt-winters com dois ciclos**. Dissertação de mestrado, DEE, PUC-Rio, março 2007.

MORETTIN, P.A.; TOLOI, C.M.C. **Análise de séries temporais**. São Paulo: E.Blucher, 2004.

NEWBERY, D.; **Competition, Contracts and Entry in the Electricity Spot Market**. *Rand Journal of Economics* 29(4): pp. 726-749, 1998.

PEARSON, A. V.; HARTLEY, H. O. **Biometrika Tables for Statisticians**, Vol 2, Cambridge, England, Cambridge University Press, 1972.

PECI, A.; **Reforma regulatória brasileira dos anos 90 à luz do modelo de Kleber Nascimento**. *Revista de Administração Contemporânea* 11 (1): pp. 11-30, 2007.

PIRES, J.C.L.; GIAMBIAGI, F.; SALES, A.F. **As perspectivas do setor elétrico após o racionamento**. *Revista do BNDES* 9(18), pp. 163-204, 2002.

REDE ELÉTRICA DE ESPAÑA, Boletín Mensual. Disponível em: <<http://www.ree.es>>. Acesso em: janeiro de 2012.

REDE ELÉTRICA NACIONAL, Informação Mensal do Sistema Eletroprodutor. Disponível em: <<http://www.ren.pt>>. Acesso em: janeiro de 2012.

REIS, L. B.; **Geração de energia elétrica: tecnologia, inserção ambiental, planejamento, operação e análise de viabilidade**, 3ª Edição – Editora Manole, 2003.

SAID, S.E.; DICKEY, D.A. **Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order**. *Biometrika*, 71: pp. 599-607, 1984.

SHAPIRO, S.S.; WILK, M.B. **An analysis of variance test for normality**. *Biometrika*, 52: pp. 591–611, 1965.

SOUZA, R.C.; CAMARGO, M.E. **Análise e Previsão de Séries Temporais: Os Modelos ARIMA**. 2ª Ed., 2004.

ZANINI, A.; SOUZA, R.C. **Redes Neurais e Regressão Dinâmica: Um modelo híbrido para previsão de curto prazo da demanda de gasolina automotiva no Brasil**. Dissertação de mestrado, DEE, PUC-Rio, abril 2000.

Apêndice

Foram simulados alguns tamanhos da janela para o cálculo da TTM e criadas as séries trimestrais de carga do SIN com estes diferentes valores para o tamanho da janela. Mais uma vez foi utilizada a série trimestral do PIB para comparar as diferentes séries de carga produzidas. Isto é, a série de carga que apresentar o maior R^2 na regressão dinâmica com a série do PIB como variável resposta será considerada a série melhor ajustada pelo efeito temperatura. A tabela Ap.1 mostra os diferentes valores de R^2 para cada uma das séries de carga geradas.

Tabela Ap.1: Valores de R^2 para diferentes tamanhos de janela

Tamanho	R^2
31	89,55%
17	89,59%
11	89,60%
3	89,61%

Fonte: ONS, elaboração própria.

De acordo com a Tabela Ap.1, é visto que o tamanho de janela 3 no cálculo da TTM teve uma pequena melhora em relação ao tamanho 31, aumentando o R^2 de 89,55% para 89,61%.

ANEXO A

Modelagem da série diária de Carga via Regressão Dinâmica

Para cada subsistema será considerado o modelo inicial, que conta apenas com o intercepto, para explicar o valor da Carga, e a partir deste modelo serão adicionadas variáveis defasadas da própria Carga, a variável Temperatura e suas defasagens e a variável de erro defasada. As variáveis serão incluídas no modelo uma de cada vez e por ordem de contribuição para o modelo, isto é, a variável incluída será aquela que mais ganho, em termos de explicação, trará para o modelo. Este procedimento será feito até incluir todas as variáveis que influenciam o valor da Carga.

Os modelos que explicam as séries de Carga para os subsistemas Nordeste, Sul e Sudeste / Centro-Oeste estão representados nas equações A.1, A.2 e A.3, respectivamente.

$$\begin{aligned} \text{Carga}_{t|\text{NE}} = & 0,62 \cdot \text{Carga}_{t-1|\text{NE}} + 0,09 \cdot \text{Carga}_{t-2|\text{NE}} + 0,07 \cdot \text{Carga}_{t-4|\text{NE}} + \\ & 0,05 \cdot \text{Carga}_{t-5|\text{NE}} + 0,12 \cdot \text{Carga}_{t-7|\text{NE}} + 0,04 \cdot \text{Carga}_{t-14|\text{NE}} + \\ & 59,21 \cdot \text{Temp}_{t|\text{NE}} - 12,49 \cdot \text{Temp}_{t-2|\text{NE}} - 22,62 \cdot \text{Temp}_{t-5|\text{NE}} \\ & - 23,13 \cdot \text{Temp}_{t-8|\text{NE}} + 0,05 \cdot \text{Erro}_{t-3|\text{NE}} \end{aligned} \quad (\text{A.1})$$

$$\begin{aligned} \text{Carga}_{t|\text{Sul}} = & 85,33 + 0,70 \cdot \text{Carga}_{t-1|\text{Sul}} + 0,09 \cdot \text{Carga}_{t-3|\text{Sul}} + \\ & 0,04 \cdot \text{Carga}_{t-7|\text{Sul}} + 0,04 \cdot \text{Carga}_{t-14|\text{Sul}} + 45,53 \cdot \text{Temp}_{t|\text{Sul}} - \\ & 25,56 \cdot \text{Temp}_{t-1|\text{Sul}} - 8,56 \cdot \text{Temp}_{t-3|\text{Sul}} + 0,16 \cdot \text{Erro}_{t-7|\text{Sul}} \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} \text{Carga}_{t|\text{SE}} = & 0,75 \cdot \text{Carga}_{t-1|\text{SE}} + 0,07 \cdot \text{Carga}_{t-3|\text{SE}} + 0,01 \cdot \text{Carga}_{t-4|\text{SE}} + \\ & 0,05 \cdot \text{Carga}_{t-7|\text{SE}} + 0,03 \cdot \text{Carga}_{t-14|\text{SE}} + 159,97 \cdot \text{Temp}_{t|\text{SE}} - \\ & 47,86 \cdot \text{Temp}_{t-2|\text{SE}} - 45,01 \cdot \text{Temp}_{t-3|\text{SE}} - 31,27 \cdot \text{Temp}_{t-5|\text{SE}} - \\ & 26,85 \cdot \text{Temp}_{t-7|\text{SE}} - 0,08 \cdot \text{Erro}_{t-1|\text{SE}} + 0,10 \cdot \text{Erro}_{t-7|\text{SE}} \end{aligned} \quad (\text{A.3})$$

A tabela A.1 mostra dois indicadores para avaliar o ajuste do modelo aos dados, o MAPE e o R². Percebe-se que o modelo foi bem ajustado, tendo em vista que o valor do MAPE foi próximo de 1,3% e os valores de R² foram altos, maiores do que 95%, para cada um dos três subsistemas.

Tabela A.1: MAPE e R² dos modelos

Subsistema	R ²	MAPE
Nordeste	98,0%	1,0%
Sul	95,2%	1,6%
Sudeste / Centro-Oeste	97,2%	1,1%

Fonte: ONS, elaboração própria.

Foi desenvolvido um teste t para detectar se as variáveis regressoras são significativas nos modelos. Estes testes foram desenvolvidos a cada variável nova que era incluída no modelo. Notou-se que todos os parâmetros, de cada um dos três modelos, apresentaram um p-valor inferior ao nível de significância usual de 5%, rejeitando, assim, a hipótese nula, ou seja, o parâmetro β_0 das variáveis regressoras é diferente de 0 (zero).

Foi desenvolvido um teste t para detectar se as variáveis regressoras são significativas nos modelos. Estes testes foram desenvolvidos a cada variável nova que era incluída no modelo. Notou-se que todos os parâmetros, de cada um dos três modelos, apresentaram um p-valor inferior ao nível de significância usual de 5%, rejeitando, assim, a hipótese nula, ou seja, o parâmetro β_0 das variáveis regressoras é diferente de 0 (zero).

Os gráficos A.1, A.2 A.3 ilustram as séries originais e ajustadas de cada um dos subsistemas. Os gráficos, mais uma vez, mostram o bom ajuste dos modelos.

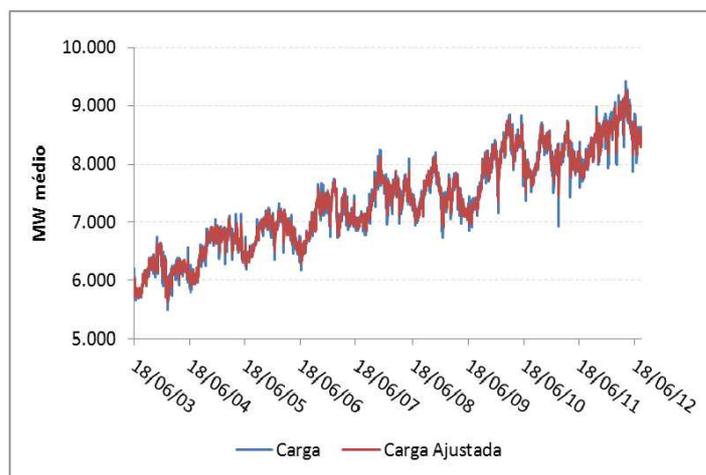


Gráfico A.1: Carga x Carga ajustada do subsistema Nordeste

Fonte: ONS, elaboração própria.

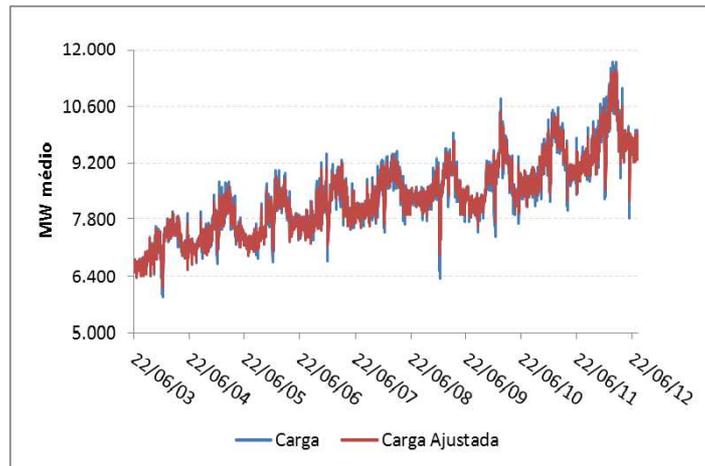


Gráfico A.2: Carga x Carga ajustada do subsistema Sul
Fonte: ONS, elaboração própria.

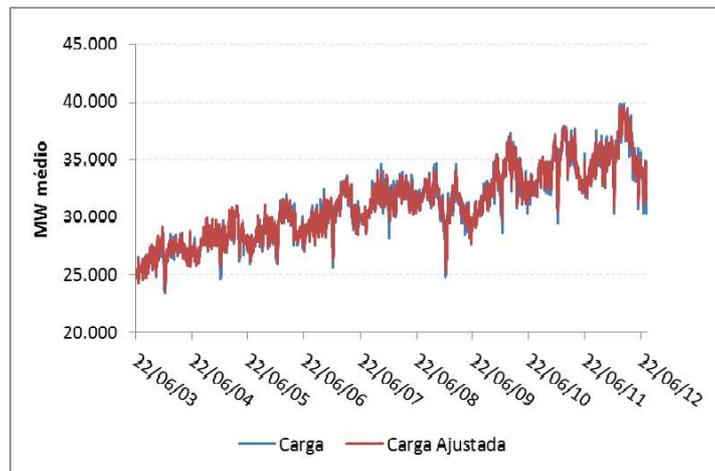


Gráfico A.3: Carga x Carga ajustada do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

Para analisar os resíduos dos modelos, serão feitos gráficos e testes estatísticos.

Um teste estatístico para verificar se a média da série é 0 (zero) será aplicado nos resíduos. A tabela A.2 ilustra os resultados deste teste para os três subsistemas.

Tabela A.2: Teste para verificar média zero nas séries

Subsistema	Est. Teste	P-valor	Decisão
Nordeste	0,06	0,95	Não Rejeitar H_0
Sul	0,00	1,00	Não Rejeitar H_0
Sudeste / Centro-Oeste	0,08	0,93	Não Rejeitar H_0

Fonte: ONS, elaboração própria.

Para os três subsistemas, o p-valor do teste é maior do que o nível de significância usual de 5%, isto indica a não rejeição da hipótese nula e as séries de resíduos, para os três subsistemas, apresentam média igual a 0 (zero).

Outro pressuposto dos resíduos é que eles sejam estacionários, para isto será feito um teste de Dickey-Fuller, onde a hipótese nula é não-estacionariedade da série. A tabela A.3 traz os resultados para os três subsistemas.

Tabela A.3: Teste para verificar estacionariedade das séries

Subsistema	Est. Teste	P-valor	Decisão
Nordeste	-16,20	0,01	Rejeitar H_0
Sul	-15,71	0,01	Rejeitar H_0
Sudeste / Centro-Oeste	-16,74	0,01	Rejeitar H_0

Fonte: ONS, elaboração própria.

Como pode ser visto na tabela A.3, todos os 3 subsistemas apresentaram p-valores baixos, rejeitando, assim, a hipótese nula de não-estacionariedade, ou seja, as séries de resíduos são estacionárias.

Para verificar homocedasticidade e auto-correlação serial, serão analisados os gráficos de dispersão entre as variáveis carga ajustada e resíduos (gráficos A.4, A.5 e A.6).

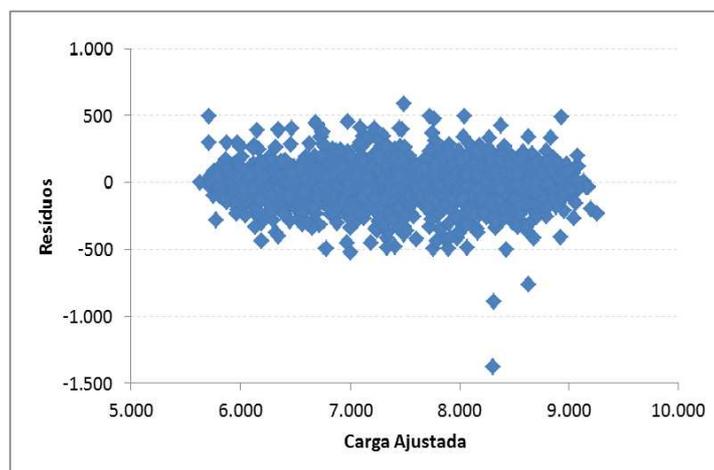


Gráfico A.4: Dispersão entre as variáveis carga ajustada e resíduos do subsistema Nordeste

Fonte: ONS, elaboração própria.

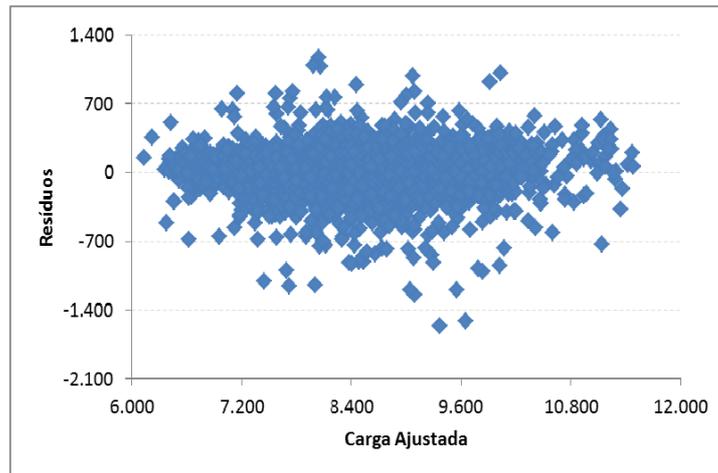


Gráfico A.5: Dispersão entre as variáveis carga ajustada e resíduos do subsistema Sul
Fonte: ONS, elaboração própria.

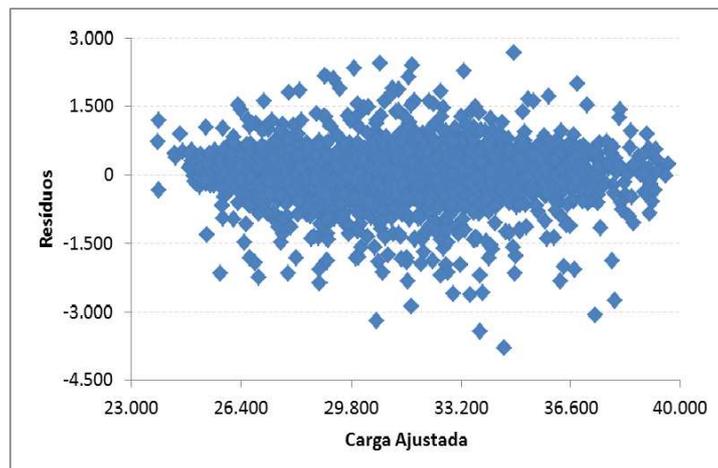


Gráfico A.6: Dispersão entre as variáveis carga ajustada e resíduos do subsistema Sudeste / Centro-Oeste
Fonte: ONS, elaboração própria.

Através dos gráficos A.4, A.5 e A.6, conclui-se que os resíduos dos modelos dos três subsistemas são homocedásticos, isto é, apresentam variância constante e não apresentam auto-correlação serial. Para ter mais certeza de que os resíduos seguem um ruído branco, serão apresentados os gráficos de Função de Auto-Correlação, representados nos gráficos A.7, A.8 e A.9 para os subsistemas Nordeste, Sul e Sudeste / Centro-Oeste, respectivamente.

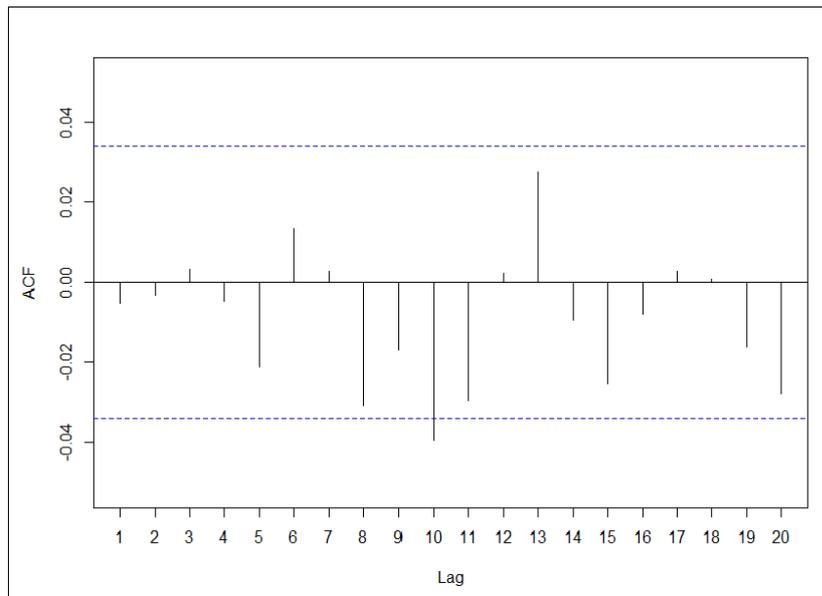


Gráfico A.7: FAC dos resíduos para o subsistema Nordeste
Fonte: ONS, elaboração própria.

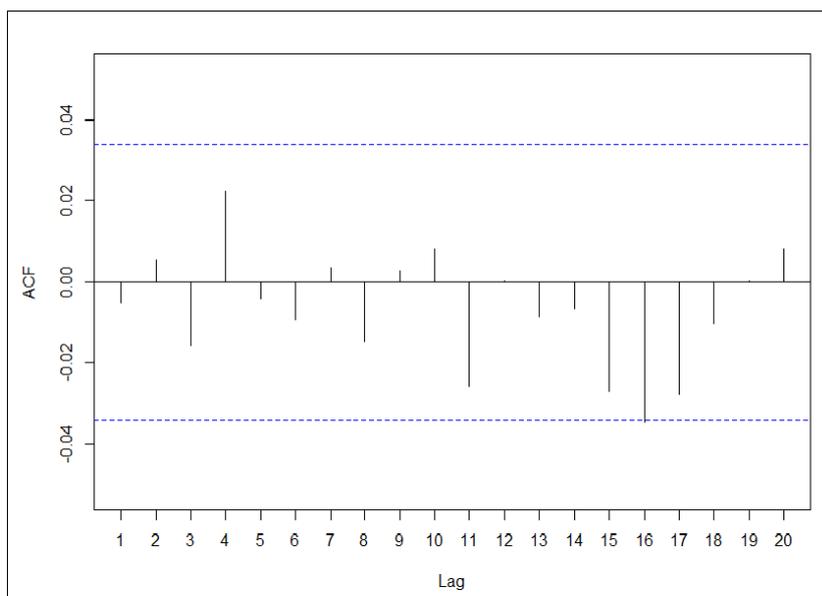


Gráfico A.8: FAC dos resíduos para o subsistema Sul
Fonte: ONS, elaboração própria.

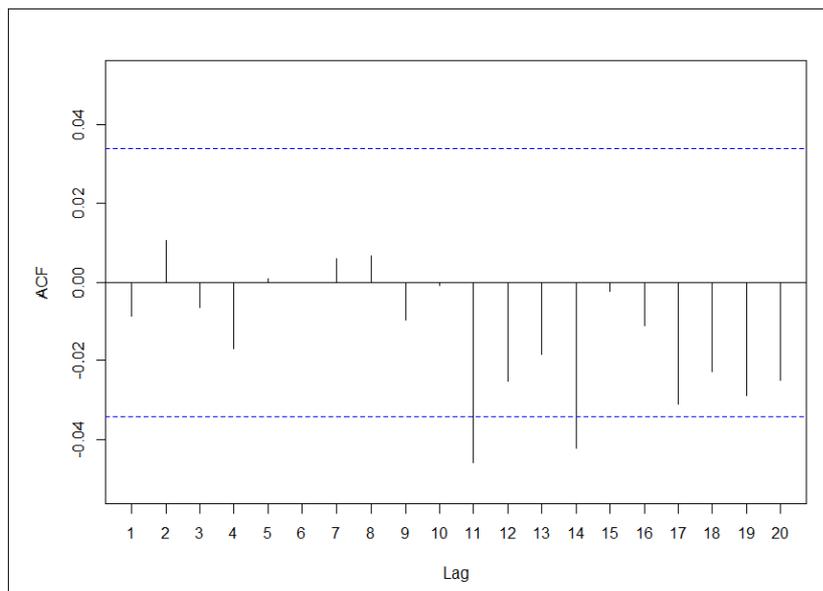


Gráfico A.9: FAC dos resíduos para o subsistema Sudeste / Centro-Oeste

Fonte: ONS, elaboração própria.

Através da análise dos gráficos A.7, A.8 e A.9, pode-se dizer que as auto-correlações para a série de resíduo dos 3 subsistemas são muito baixas, bem próximas de zero. Para os três subsistemas, pelo menos um dos 20 lag's ultrapassou o limite inferior, indicando que para este lag há auto-correlação. Porém, em todos os casos, este lag auto-correlacionado é alto, não influenciando no ajuste da Carga no tempo t . A hipótese de que os resíduos não apresentam auto-correlação serial é confirmada com o teste de Ljung-Box, pois as três séries de resíduos apresentaram um P-valor alto, maior do que o nível de significância usual de 5% (conforme tabela A.4), dessa forma, não se rejeita a hipótese nula de que os resíduos seguem um ruído branco.

Tabela A.4: Teste de Ljung-Box para os resíduos dos modelos.

Subsistema	Est. Teste	P-valor	Decisão
Nordeste	0,09	0,77	Não Rejeitar H_0
Sul	0,09	0,77	Não Rejeitar H_0
Sudeste / Centro-Oeste	0,25	0,62	Não Rejeitar H_0

Fonte: ONS, elaboração própria.