



Pedro Guilherme Costa Ferreira

A Estocasticidade Associada ao Setor Elétrico Brasileiro e uma Nova Abordagem para a Geração de Afluências Via Modelos Periódicos Gama

Tese de Doutorado

Tese apresentada ao programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio como parte dos requisitos parciais para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Prof. Reinaldo Castro Souza

Rio de Janeiro

Março de 2013



Pedro Guilherme Costa Ferreira

A Estocasticidade Associada ao Setor Elétrico Brasileiro e uma Nova Abordagem para a Geração de Afluências Via Modelos Periódicos Gama

Tese apresentada como requisito parcial para obtenção do grau de Doutor 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. Alexandre Zanini
UFJF

Prof. Fernando Salgueiro Perobelli
UFJF

Prof. Joari Paulo da Costa
ONS

Profa. Monica Barros
ENCE

Prof. Nivalde José de Castro
UFRJ

Prof. Plutarcho Maravilha Lourenço
Centro de Pesquisas de Energia Elétrica

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 14 de março de 2013

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

Pedro Guilherme Costa Ferreira

Graduação em Economia, com ênfase em Econometria Espacial e Mestrado em Economia Aplicada (Modelos Econométricos com Insumo-Produto), ambos pela Universidade Federal de Juiz de Fora (2008). Atualmente, Doutorando em Engenharia Elétrica (Métodos de Apoio à Decisão) pela Pontifícia Universidade Católica do Rio de Janeiro.

Ficha Catalográfica

Ferreira, Pedro Guilherme Costa

A estocasticidade associada ao setor elétrico brasileiro e uma nova abordagem para a geração de afluições via modelos periódicos Gama / Pedro Guilherme Costa Ferreira; orientador: Reinaldo Castro Souza. – 2013.

138 f. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2013.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Setor elétrico brasileiro. 3. Modelos periódicos autorregressivos. 4. Modelos Gama. I. Souza, Reinaldo Castro. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

À minha família, à Luciana, ao Pedrinho e ao povo brasileiro, por ter investido durante tantos anos nos meus estudos. Espero retribuir tal investimento.

Resumo

Ferreira, Pedro Guilherme Costa; Souza, Reinaldo Castro (Orientador). **A estocasticidade associada ao Setor Elétrico Brasileiro e uma nova abordagem para a geração de aflúências via Modelos Periódicos Gama.** Rio de Janeiro, 2013. 138p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Essa tese discute, basicamente, três importantes tópicos relacionados ao Setor Elétrico Brasileiro (SEB), a saber, a estocasticidade intrínseca ao setor, as peculiaridades metodológicas que norteiam o módulo de geração de energias afluentes, peça chave às atividades de planejamento da expansão, operação e precificação de curto prazo, e a “atenção” para a necessidade de um modelo Gama, mais adequado às características da série histórica de energia a ser modelada. Nesta tese é apresentado ao leitor uma visão diferente do Setor Elétrico Brasileiro, evidenciando-se a intrínseca relação entre a estocasticidade e as atividades executadas pelo SEB, um importante setor de infraestrutura do país. Percebe-se que as séries sintéticas de energia/vazão são cruciais para determinar qual é a melhor maneira de operar o setor, subsidiam decisões sobre quando se deve ou não expandi-lo, evitando-se assim custo e/ou perdas desnecessárias. E, ainda, são um fator preponderante na determinação do preço de curto prazo de energia elétrica, dado que a quantidade simulada/prevista de água nos reservatórios no futuro será um dos determinantes do preço da energia no curto prazo. Após este preâmbulo apresenta-se o Módulo de Geração de energias Afluentes, destacando-se como o mesmo está ligado ao Setor Elétrico Brasileiro e à Programação Dinâmica Dual Estocástica (PDDE), enfatizando quais as “ressalvas” que esta ligação estabelece. Ainda neste tópico, levantam-se algumas peculiaridades do módulo, como por exemplo, a questão das configurações intercorrelacionadas, característica não discutida até então, e pontos que levantam discussão com relação à teoria de Séries Temporais, como por exemplo, a definição da ordem do modelo, a questão da estacionariedade e a necessidade de tratar possíveis *outliers*. Ao unir as especificidades do SEB com a característica da série histórica de energia relacionada à distribuição Gama, aborda-se, seguindo as peculiaridades do SEB, uma nova metodologia que prescinde a normalidade e a possível geração de cenários negativos. Ao adotar tal metodologia, concluiu-se que a utilização dos modelos Gama, da forma que foram propostos por (Fernandez & Salas, 1986) não podem ser usados pelo SEB, pois os resultados da geração dos cenários mostraram-se não factíveis. Por fim, a importância desse trabalho vai além do que foi discutido em sua essência, ele abre um leque de possibilidades e discussões sobre como está sendo feito o planejamento da operação, o planejamento da expansão e a determinação do preço spot da energia elétrica. Nesse sentido a ideia desse trabalho é ter como resultado intangível a criação de uma agenda de pesquisa sobre a modelagem estocástica que envolve o SEB.

Palavras-chave

Setor Elétrico Brasileiro; Modelos Periódicos Autorregressivos; Modelos Gama.

Abstract

Ferreira, Pedro Guilherme Costa; Souza, Reinaldo Castro (Advisor). **The stochasticity associated with Brazilian Electrical Sector and a new approach to generate Natural Inflow Energy via Periodic Gama Model.** Rio de Janeiro, 2013. 138p. PhD Thesis - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis discusses basically three important topics related to the Brazilian Electric Sector (BES), namely, the intrinsic stochasticity of the sector, the methodological peculiarities that guide the module of affluent energy, a key planning activities of the expansion, operation and pricing of short-term, and "attention" to the necessity for a Gamma model, more suitable to the characteristics of the time series of affluent energy. This thesis presents the reader with a different view of the Brazilian Electric Sector, revealing the intrinsic relationship between stochasticity and the activities performed by BES, a major infrastructure sector in the country. It can be seen that the synthetic series of energy/flow rates are crucial to determine which is the best way to operate the industry, subsidize decisions about when to expand it or not, thus avoiding the cost and/or unnecessary losses. And yet, they are a major factor in determining the short-term price of electricity, since the amount simulated / predicted water reservoirs in the future will be one of the determinants of the price of energy in the short term. After this preamble presents the Module of Affluent energy, especially as it is linked to the Brazilian Electric Sector and Stochastic Dual Dynamic Programming (SDDE), which emphasizes the "attentions" that establishes this link. Although this topic, raises some peculiarities of the module, for example, the issue of intercorrelated settings, feature not discussed so far, and points that raise discussion about the theory of time series, such as the definition of the model order, the question of stationarity and the need to deal with possible outliers. By joining the especifications of BES with the characteristic energy of the time series related to the Gamma distribution is approached, following the peculiarities of BES, a new methodology that dispenses normality and possible generation of negative scenarios. By adopting this method, it was concluded that the use of Gamma models, which have been proposed in the form of (Fernandez & Salas, 1986) can not be used for SEB, as the results showed the generation of the scenarios is not feasible. Finally, the importance of this thesis goes beyond what was discussed in its essence, it opens up a range of possibilities and discussions on how it is being done the operation planning, expansion planning and determining the spot price of electricity. In this sense the idea of this thesis is to deliver intangible creating a research agenda on the stochastic modeling involving SEB.

Keywords

Brazilian Electrical Sector; Periodic Autoregressive Models; Gama Models.

Sumário

1. Introdução	14
1.1. Considerações Iniciais	14
1.2. Contexto do Trabalho	15
1.3. Objetivo do Trabalho	16
1.4. Relevância do Trabalho	18
1.5. Organização	18
2 . Como a estocasticidade orienta o Setor Elétrico Brasileiro	21
2.1. Do Modelo Antigo ao Modelo Atual	21
2.2. O Setor Elétrico sob o Novo Modelo	27
2.2.1. O planejamento da operação do Sistema Elétrico	31
2.2.2. O Planejamento da Expansão do Sistema Elétrico	39
2.2.3. Comercialização de Energia Elétrica	42
3 . Referencial Teórico	45
3.1. Métodos e Modelos aplicados no contexto de simulação de vazões/energia	45
3.2. Modelos de Box & Jenkins	49
3.3. O Modelo PAR(p)	53
4 . O modelo de geração de Energias Afluentes (MGEA) no Contexto do SEB	59
4.1. Peculiaridades do MGEA	60
4.1.1. A montagem das configurações do Parque Gerador	61
4.1.2. Modelo “PAR(p) Interconfigurações”	66
4.1.3. As equações do Modelo Autorregressivo Periódico e a PDDE	75
4.1.4. Geração de séries sintéticas e a questão das afluências negativas	77
4.1.5. A estrutura não-linear do PAR (p)	81
4.2. Estudo de caso: Modelagem e Geração de Energias Afluentes no contexto do SEB	84
4.2.1. Ajuste do Modelo “PAR(p) Interconfigurações”	85
4.2.1.1. A questão da estacionariedade	85
4.2.1.2. A questão da identificação da ordem e os parâmetros	86
4.2.1.3. Análise dos Ruídos	90
4.2.2. Ajuste dos cenários	97
4.2.2.1. Séries sintéticas e teste de média dos cenários	97
4.2.2.2. Teste de Variância dos cenários	99
5 . Processo Periódico Autorregressivo Gama	102
5.1. O modelo teórico	103
5.1.1. Modelo Periódico Autorregressivo Aditivo Gama (PAGAR(1))	103
5.1.2. Modelo Periódico Autorregressivo Multiplicativo Gama (PMGAR(1))	107
5.1.3. Modelo Periódico Autorregressivo Gama (PGAR(1))	111
5.1.3.1. Definição	111

5.1.3.2. Estimação dos parâmetros dos modelos	113
5.2. Análise Empírica	114
5.2.1. A série de ENA e a distribuição Gama	114
5.2.2. A modelagem Gama periódica	118
6 . Considerações finais e trabalhos futuros	128
7 . Referências	131

Lista de figuras

Figura 2.1 – Agentes institucionais do Setor Elétrico Brasileiro	28
Figura 2.2 – Sistema Interligado Nacional	30
Figura 2.3 – Problema de decisão da operação	32
Figura 2.4 – Critério de Planejamento da Operação	33
Figura 2.5 – Cadeia de Modelos em vigor no SEB	35
Figura 2.6 – Relação entre os módulos do NEWAVE	38
Figura 2.7 – Principais entradas e saídas do programa DECOMP	39
Figura 2.8 – A decisão do Investimento	40
Figura 2.9 – Critério de planejamento da expansão	40
Figura 3.1 – Estágios da modelagem Box & Jenkins	52
Figura 4.1 – Usinas instaladas e que entraram em operação de acordo com cada configuração	64
Figura 4.2 – Pré-estudo, Pós-estudo e as diferentes configurações de energia	65
Figura 4.3 – FAC-IC <i>mainstream</i>	68
Figura 4.4 – FAC-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo	69
Figura 4.5 – FAC-IC cruzando com as configurações anteriores	69
Figura 4.6 – FAC-IC <i>mainstream</i>	71
Figura 4.7 – FAC-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo	71
Figura 4.8 – FAC-IC cruzando com as configurações anteriores	72
Figura 4.9 – Algoritmo para testar a estacionariedade “Intercorrelações”	86
Figura 4.10 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Sudeste/Centro-Oeste	88
Figura 4.11 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Sul	88
Figura 4.12 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Nordeste	89
Figura 4.13 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Norte	89

Figura 4.14 – FAC dos resíduos – Subsistema Sudeste/Centro-Oeste	91
Figura 4.15 – FAC dos resíduos – Subsistema Sul	91
Figura 4.16 – FAC dos resíduos – Subsistema Nordeste	91
Figura 4.17 – FAC dos resíduos – Subsistema Norte	91
Figura 4.18 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Sudeste/Centro-Oeste	92
Figura 4.19 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Sul	93
Figura 4.20 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Nordeste	93
Figura 4.21 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Norte	93
Figura 4.22 – Histograma e estatísticas dos Resíduos – Subsistema Sudeste/Centro-Oeste	94
Figura 4.23 – Histograma e estatísticas dos Resíduos – Subsistema Sul	94
Figura 4.24 – Histograma e estatísticas dos Resíduos – Subsistema Nordeste	94
Figura 4.25 – Histograma e estatísticas dos Resíduos – Subsistema Norte	95
Figura 4.26 – Gráfico dos resíduos para detecção de <i>outliers</i> – Subsistema Sudeste/Centro-Oeste	95
Figura 4.27 – Gráfico dos resíduos para detecção de <i>outliers</i> – Subsistema Sul	96
Figura 4.28 – Gráfico dos resíduos para detecção de <i>outliers</i> – Subsistema Nordeste	96
Figura 4.29 – Gráfico dos resíduos para detecção de <i>outliers</i> – Subsistema Norte	96
Figura 4.30 – Séries Sintéticas e médias dos cenários	97
Figura 4.31 – Séries Sintéticas e médias dos cenários	98
Figura 4.32 – P-valores do Teste t	98
Figura 4.33 – P-valores do Teste t	99
Figura 4.34 – P-valores do Teste de Levene	99
Figura 4.35 – P-valores do Teste de Levene	100
Figura 5.1 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Sudeste/Centro-Oeste	115

Figura 5.2 – Teste do Qui-Quadrado (λ^2) para série de ENA mensal do subsistema Sudeste/Centro-Oeste	116
Figura 5.3 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Sul	116
Figura 5.4 – Teste do Qui-Quadrado (λ^2) para série de ENA mensal do subsistema Sul	116
Figura 5.5 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Nordeste	117
Figura 5.6 – Teste do Qui-Quadrado (λ^2) para série de ENA mensal do subsistema Nordeste	117
Figura 5.7 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Norte	118
Figura 5.8 – Teste do Qui-Quadrado (λ^2) para série de ENA mensal do subsistema Norte	118
Figura 5.9 – Estudo da restrição (5.22a) para a série histórica do Subsistema Sudeste/Centro-Oeste	120
Figura 5.10 – Estudo da restrição (5.22a) para a série histórica do Subsistema Sul	120
Figura 5.11 – Estudo da restrição (5.22a) para a série histórica do Subsistema Nordeste	120
Figura 5.12 – Estudo da restrição (5.22a) para a série histórica do Subsistema Norte	121
Figura 5.13 – Restrição (5.22a) corrigida para a série histórica do Subsistema Sul	123
Figura 5.14 – Cenários e testes - Subsistema Sudeste/Centro-Oeste	1255
Figura 5.15 – Cenários e testes - Subsistema Nordeste	125
Figura 5.16 – Cenários e testes - Subsistema Nordeste	126
Figura 5.17 – Cenários e testes - Subsistema Norte	126

Lista de tabelas

Tabela 4.1 – Exemplo de Novas Configurações	64
Tabela 4.2 – Testes de Estacionariedade	86

Lista de quadros

Quadro 2.1 – Comparativo entre os Modelos Energéticos	25
Quadro 3.1 – Artigos aplicados ao problema de simulação/previsão de vazões/ENAs	50
Quadro 3.1 – Artigos aplicados ao problema de simulação/previsão de vazões/ENAs (continuação)	51

*“O mais importante da vida
não é a situação em que estamos,
mas a direção para a qual nos movemos”.*

Oliver Wendell Holmes

1

Introdução

1.1

Considerações Iniciais

O Setor Elétrico Brasileiro (SEB) ocupa a 9ª posição mundial em termos de geração energética, produzindo um total de aproximadamente 470 TWh (OECD, 2012). Em 2011, os setores de geração, distribuição e transmissão tiveram um faturamento de aproximadamente R\$ 15 bilhões (ABINEE, 2012) e a eletricidade é um insumo que atualmente chega à casa de mais de 99% dos brasileiros.

Para se ter uma ideia do tamanho do setor, estima-se que, no quadriênio 2011 a 2014, os investimentos do setor elétrico serão na ordem de R\$100 bilhões (PAC, 2012), ocupando o 2º lugar em termos de investimentos dos setores vinculados à infraestrutura. Somente o BNDES (Banco Nacional de Desenvolvimento) estima investir cerca de R\$ 60 bilhões no setor, 10% do total que será investido pelo Banco no quadriênio (BNDES, 2012).

Afora a magnitude econômica do SEB, graças a um processo histórico de aproveitamento dos recursos naturais para geração de energia, desenvolveu-se o Sistema Interligado Nacional (SIN). O SIN, de dimensões continentais, permite, respeitado certas configurações técnicas, que a energia gerada em qualquer ponto do país seja consumida por diferentes consumidores, em regiões distintas. A interligação entre as regiões resulta em melhor aproveitamento dos recursos. O SEB resultante é de grande porte, complexo, e modelos computacionais constituem ferramentas (necessárias) para apoiar o planejamento e operação do SIN.

Neste contexto, ao unir um sistema que envolve bilhões de reais em geração, distribuição e transmissão, eminentemente hidráulico, interligado nacionalmente, misto em termos de investimentos, isto é, mescla empresas públicas e privadas e, em termos de estrutura de mercado, monopólios, principalmente na área de transmissão e grandes oligopólios, geração e distribuição, surgem uma gama de problemas e desafios que requerem tomadas de decisão em diferentes horizontes

de tempo, isto é, longo (5 a 10 anos), médio (5 a 1 ano), curto (até 1 ano) e curtíssimo prazo (programação diária).

Tais problemas são de diferentes ordens e envolvem tomadas de decisão diretamente relacionadas aos três pilares que sustentam o Setor Elétrico Brasileiro: planejamento da expansão, planejamento e programação da operação e determinação do preço da energia no curto prazo. Funções essas desempenhadas pela Empresa de Pesquisa Energética (EPE), pelo Operador Nacional do Sistema Elétrico (ONS) e Câmara de Comercialização de Energia Elétrica (CCEE), respectivamente.

Dentro dessa conjuntura, algumas questões surgem, tais como: deve-se investir no aumento da capacidade do sistema ou poupar e aguardar um momento de maior expansão da atividade econômica? Quanto se deve gerar entre energia hidráulica e térmica de forma a atender a demanda? Qual é o momento certo para economizar água e usar combustíveis fósseis? Qual deve ser o preço *spot* da energia para que sejam pagos todos os fatores de produção e para que a modicidade tarifária seja atendida?

Essas e outras questões fazem parte do cotidiano do Setor Elétrico Brasileiro e respondê-las não é uma tarefa trivial, sendo necessários amplo planejamento e uma gestão sincronizada. Ressalte-se que a estocasticidade está presente nas atividades dos três pilares do SEB: planejamento da expansão, planejamento e programação da operação e precificação de curto prazo.

1.2

Contexto do Trabalho

Para planejar a expansão, a operação e determinar o preço da energia elétrica no curto prazo, o Setor Elétrico Brasileiro desenvolveu uma cadeia de modelos, na qual, à medida em que o horizonte de planejamento fica menor, a representação das incertezas diminui e a representação das características físicas aumenta. O modelo de mais longo prazo desta cadeia é o modelo NEWAVE. Nele, em linhas gerais, a determinação da estratégia de operação, isto é, a definição da repartição de geração hidráulica ou térmica, é objeto de um processo de otimização estocástico que utiliza o algoritmo de Programação Dinâmica Dual

Estocástica – PDDE e determina uma função de custo futuro (FCF), que será utilizada no acoplamento dos demais modelos da cadeia.

Tal processo de otimização tem como variável estocástica a Energia Natural Afluente (ENA), gerada por meio de um modelo Periódico Autorregressivo (PAR(p)), ajustado a partir dos parâmetros estimados do histórico da mesma. Um ponto peculiar que deve ser ressaltado é que, para evitar a geração de cenários negativos, o modelo PAR(p) adotado pelo SEB utiliza uma distribuição lognormal para produzir as séries sintéticas simuladas via Monte Carlo.

O sistema de planejamento da expansão, operação e precificação do SEB vem sendo questionado ao longo dos últimos anos quanto à sua eficácia na correta representação dos cenários sintéticos de ENAs/vazões, e suas possíveis consequências, como distúrbios no preço “justo” do PLD (Preço de Liquidação das Diferenças), por exemplo. Tal fato ficou mais evidente, quando, por exemplo, no mês de junho de 2010, a ANEEL expediu um despacho (despacho nº 2.654, de 24 de Junho) determinando a revisão das previsões de vazões do Programa Mensal de Operação (PMO), resultando em recálculo do PLD que em algumas semanas apresentou aumento significativo e em outras, redução expressiva.

Assim, embora com relação ao preço da energia seja possível identificar uma corrente no setor a favor de que o preço seja determinado pelo mercado e não por modelos computacionais, este trabalho insere-se no contexto dos três pilares do SEB mencionados na seção anterior e tem como propósito levantar pontos e propor aperfeiçoamentos que devem ser discutidos e observados pelos agentes do SEB.

1.3

Objetivo do Trabalho

Conforme se observou, há uma gama de problemas e questões que podem ser mais bem equacionadas que envolvem o Setor Elétrico Brasileiro. Embora todas essas questões sejam relevantes, o objetivo dessa tese não é discorrer sobre todas elas e tampouco procurar soluções para resolvê-las, mas sim procurar saídas para alguns dos problemas.

Uma vez que muitas tomadas de decisão do SEB estão relacionadas ao problema estocástico, e mais especificamente, ao problema de geração de séries

sintéticas de vazões/ENAs para subsidiar o processo de otimização e simulação que auxilia nas decisões do setor, este trabalho insere-se num ponto ainda mais específico na cadeia de modelos utilizados pelo SEB, que é a geração de séries sintéticas de ENA, com periodicidade mensal, as quais, no contexto do SEB, são geradas pelo Modelo de Energias Afluentes (MGEA) e utilizadas pelo módulo NEWAVE.

Neste contexto, um dos objetivos centrais dessa tese é propor uma abordagem para a geração de séries sintéticas de energia utilizando a distribuição Gama. Tal abordagem se faz relevante por três pontos principais: (i) a série histórica de ENA comporta-se de maneira muito próxima à distribuição Gama e utilizar tal distribuição à série de dados e aos resíduos poderá permitir resultados mais próximos ao “perfil” do histórico; (ii) conforme demonstrado no capítulo 4, ao utilizar a distribuição lognormal para evitar a geração de cenários negativos, surge um problema metodológico de não linearidade dos modelos PAR(p) que serão utilizados na PDDE. Para evitar possíveis distorções metodológicas advindas dessa perda de linearidade, propõe-se um modelo que já apresenta uma formulação não linear e que extrapola o problema séries sintéticas negativas, que é o modelo Gama; e, (iii) novas abordagens para geração dos cenários podem ser úteis em simulações da Função de Custo Futuro, gerada pelo NEWAVE, uma vez que as distribuições de probabilidade apresentam características distintas.

O segundo objetivo é abordar o Modelo de Geração de Energias Afluentes (MGEA) no contexto do SEB, em outras palavras, procurar-se-á definir como são feitas as montagens das configurações do parque gerador; como funciona o Modelo “PAR(p) Interconfigurações”, explicitado pela primeira vez neste trabalho; discutir as equações do modelo autorregressivo periódico e a PDDE; a geração de séries sintéticas; e a questão das afluições negativas e a estrutura não linear do PAR(p). Ainda, abordar, sucintamente, possíveis pontos de melhoria e conflitos com a teoria Box & Jenkins.

Por fim, um objetivo mais geral, mas não menos relevante, é mostrar a relação “*Estocasticidade x Setor Elétrico Brasileiro*”, propor alguns pontos de atenção aos agentes que irão trabalhar na modelagem de séries sintéticas e levantar possíveis novos desafios ao setor.

1.4

Relevância do Trabalho

“Os modelos lineares Gaussianos podem ser uma ferramenta muito útil para modelar séries temporais hidrológicas, mas, como qualquer outro modelo, apresentam vantagens e desvantagens.”

Esse comentário, feito por (Fernandez & Salas, 1986), retrata bem o que está sendo proposto nesse trabalho e, de uma maneira geral, o que acontece na maioria dos trabalhos de pesquisa, isto é, embora o modelo “PAR(p) Interconfigurações” e a geração das séries sintéticas via lognormal tenham sido úteis para atender as demandas do SEB, estas metodologias apresentam vantagens e desvantagens que devem ser discutidas e melhoradas.

Dessa forma, a relevância deste trabalho está vinculada, em primeiro lugar, ao fato de que o tema proposto está intrinsecamente correlacionado com o Setor Elétrico Brasileiro que é um importante setor de infraestrutura do país. Ainda, ao discutir possíveis desvantagens do atual modelo e alguns pontos metodológicos, este trabalho contribui para que o setor avance no sentido de melhorar os métodos empregados. Por fim, ao propor uma nova modelagem baseada no contexto “Interconfigurações” e nos modelos periódicos, este trabalho contribui de uma maneira mais específica com o problema de geração de séries sintéticas de energia.

Assim, a inovação e a contribuição desta tese está relacionada aos três pontos que serão abordados pela mesma: apresentar a relação entre a estocasticidade e as diferentes tomadas de decisão do setor elétrico, levantar e discutir pontos referentes a metodologia utilizada pelo SEB para a modelagem das séries históricas de ENA e propor uma nova abordagem para a geração de cenários baseada nas características do histórico.

1.5

Organização

Antes de retratar a organização do trabalho propriamente dita é importante ressaltar que o mesmo foi dividido objetivando deixar claro para o leitor quais são os três pilares que sustentam o Setor Elétrico Brasileiro e de que forma os mesmos

estão interligados. Em outras palavras, é importante explicitar qual a relação entre o problema abordado e o SEB, os pontos que devem ser abordados e questionados com relação ao que é feito atualmente pelo setor e, por fim, apresentar uma nova proposta de geração de séries sintéticas, a saber, a modelagem periódica autorregressiva Gama.

Com esse intuito, além dessa parte introdutória, que mostra a dimensão do Setor Elétrico Brasileiro e levanta algumas questões que o setor precisa responder, o presente trabalho está organizado da seguinte forma.

O segundo capítulo, intitulado “Como a estocasticidade orienta o Setor Elétrico Brasileiro” apresenta a relação existente entre o problema estocástico e as questões do setor elétrico. O capítulo começa retratando os diferentes modelos de gestão do setor, chegando ao Novo Modelo do Setor Elétrico, adotado em 2004. A partir desse ponto, retratam-se quais são as estruturas desse Novo Modelo e como os três principais agentes “operacionais” do setor são afetados pelo problema estocástico.

Dando sequência, como nos capítulos subsequentes serão abordados modelos de séries temporais aplicados ao problema de geração de séries sintéticas de vazão/ENAs, o capítulo 3 apresenta um referencial teórico sobre os modelos da família Box & Jenkins no contexto de simulação de vazões /ENAs. Em seguida, aborda a modelagem Box & Jenkins, fundamental para o entendimento dos modelos $PAR(p)$ *mainstream*.

O capítulo 4 trata do Modelo de Geração de Energias Afluentes (MGEA) no contexto do Setor Elétrico Brasileiro, a ideia desse capítulo é retratar peculiaridades do MGEA, tais como, a montagem das configurações do parque gerador, o modelo “ $PAR(p)$ Interconfigurações”, as equações do modelos autorregressivo periódico e a PDDE, a geração de séries sintéticas e a questão das afluências negativas, a estrutura não linear do $PAR(p)$, e por fim, um estudo de caso abordando a modelagem e geração de energias afluentes no contexto do SEB.

Dando sequência ao trabalho, no capítulo 5 abordam-se os Modelos Periódicos Autorregressivos Gama, tanto na parte metodológica, quanto com relação aos resultados. É importante ressaltar que toda abordagem feita nesse capítulo respeitará as peculiaridades de modelagem adotadas pelo Setor Elétrico Brasileiro.

Por fim, o último capítulo fica reservado às conclusões e à agenda de pesquisas.

2

Como a estocasticidade orienta o Setor Elétrico Brasileiro

A ideia desse capítulo é, em primeiro lugar, apresentar ao leitor uma visão do Setor Elétrico Brasileiro como um todo, identificando o processo de transformação pelo qual passou ao longo das duas últimas décadas, chegando ao Novo Modelo adotado pelo setor.

Com a inserção do Novo Modelo, mostra-se como a boa modelagem da variável estocástica é determinante para o bom funcionamento dos três pilares fundamentais do Setor Elétrico Brasileiro, quais sejam planejamento da expansão, planejamento e programação da operação e processo de contabilização e liquidação das transações de energia no mercado de curto prazo.

Este capítulo está dividido da seguinte forma: na próxima seção explicitam-se quais foram os modelos adotados pelo SEB até chegar ao modelo atual. A seguir, descreve-se o setor elétrico sob o prisma do Novo Modelo, identificando quais são as relações entre os três pilares fundamentais do SEB e a estocasticidade.

2.1

Do Modelo Antigo ao Modelo Atual

A procedência da crise de oferta de energia elétrica no Brasil está relacionada com quatro causas principais: (a) esgotamento do modelo de financiamento e estrutura de mercado estatal, responsável pela expansão do setor desde os anos 1960; (b) falhas no planejamento da transição do modelo estatal para o modelo privado; (c) problemas contratuais e regulatórios; e (d) falta de coordenação entre os órgãos governamentais (Pires, et al., 2002).

Quanto à falência do modelo estatal, essa se deu principalmente por duas razões: em primeiro lugar, a crise fiscal do Estado, com o fim da capacidade de investimento da União nos níveis necessários para expansão do sistema (as empresas eram predominantemente estatais). Em segundo lugar, um regime regulatório inadequado, que não estimulava a busca da eficiência e do baixo custo

na geração, uma vez que as tarifas eram reguladas nos segmentos de geração, distribuição e transmissão.

Em relação à crise fiscal do Estado, que teve início na década de 1980 e reduziu a quantidade de recursos da União para investimentos¹, é importante destacar que outros fatores contribuíram para o aprofundamento desse problema. Primeiro, a elevação do custo marginal de expansão do setor, dado que as novas bacias hidrelétricas a serem exploradas se situavam mais distantes dos centros consumidores. Nesse sentido, proporcionalmente mais recursos tinham de ser investidos para construção da mesma quantidade de geração.

Segundo, tem-se a deterioração do valor real das tarifas, cujo patamar de preço passou a não refletir o aumento dos custos setoriais. As tarifas, além de serem equalizadas para todo o país, foram utilizadas como instrumento de controle inflacionário. Esse processo culminou com a descapitalização e decorrente inadimplência dos diversos agentes setoriais. Finalmente, com a consolidação da democracia e com o advento da estabilidade monetária, as demandas sociais impuseram aos governos a necessidade de maior critério na aplicação dos recursos da União (Pires, et al., 2002).

Essas razões fizeram com que os investimentos das empresas estatais não fossem capazes de acompanhar as necessidades de crescimento da demanda no país. Então um grande número de obras de expansão do SEB ficou paralisada e/ou não cumpriu o cronograma previamente estabelecido, piorando a situação financeira do Setor por conta da elevação de seus custos devido à dilatação do cronograma de obras.

Do ponto de vista regulatório, a inexistência de estímulos à busca de eficiência produtiva fez com que as empresas não tivessem incentivos para reduzirem seus custos. Segundo (Schaeffer, et al., 2003), na década de 1970, por exemplo, as tarifas eram equalizadas em todo o território nacional como incentivo ao desenvolvimento energético de certas regiões, fazendo com que empresas superavitárias e deficitárias compensassem, mediante transferências, os ganhos e perdas provenientes do esforço individual de cada uma delas. Em 1993, devido ao início do processo de privatização, as tarifas voltaram a ser fixadas pelas concessionárias distribuidoras de energia elétrica, justificando a necessidade das

¹ Até 1995 o financiamento do setor elétrico era, basicamente, com recursos públicos (Quadro 2.1) (CCEE, 2012).

empresas adequarem suas tarifas ao mercado que atendiam, de modo a atingir um patamar satisfatório de rentabilidade.

Quanto às falhas de transição do modelo estatal para o privado, ao iniciar as privatizações em meados de 1995, ao contrário do que era esperado, as empresas privatizadas não investiram na expansão do parque gerador. Assim, o efeito mais importante dessas privatizações, segundo (Fernandes, et al., 2005), foi uma rápida recuperação dos níveis tarifários (defasados até então), a fim de aumentar a atratividade das empresas privadas para o setor.

Em 1995, com a lei nº 8.987, de 13 de fevereiro de 1995, chamada “Lei das Concessões” inicia-se o processo de privatização e reforma do setor. Segundo (Pêgo & Campos Neto, 2008), além de criar condições para maior participação do capital privado, a nova Lei introduziu a competição na construção de novos projetos mediante regulamentação do regime de licitação das concessões, antes exclusivo às concessionárias estaduais e federais.

A reforma para o modelo de Livre Mercado (1995 a 2002), (quadro 2.1), teve prosseguimento com a aprovação do Decreto nº 1.503, de 25 de maio de 1995, o qual incluiu o Sistema Eletrobrás² no Programa Nacional de Desestatização (PND) e orientou as privatizações nos segmentos de geração e distribuição. Ainda, de acordo com os autores, a reestruturação foi reforçada com a Lei nº 9.074, de 7 de setembro de 1995, a qual estabeleceu as bases legais para que os grandes consumidores pudessem comprar energia livremente (Consumidores Livres), e o Decreto nº 2.003/1996 regulamentou o produtor independente de energia, peça-chave no sistema, além de autorizar a venda da produção excedente do autoprodutor.

Dando sequência à reforma, em 1996 foi implantado o Projeto de Reestruturação do Setor Elétrico Brasileiro (Projeto RE-SEB). As principais conclusões do projeto foram, (i) a necessidade de implementar a desverticalização das empresas de energia elétrica, isto é, dividi-las nos segmentos de geração, transmissão e distribuição; (ii) incentivar a competição nos segmentos de geração e comercialização, e (iii) manter sob regulação os setores de distribuição e

² A Eletrobrás controla grande parte dos sistemas de geração e transmissão de energia elétrica do Brasil por intermédio de seis subsidiárias: Chesf, Furnas, Eletrosul, Eletronorte, CGTEE e Eletronuclear. A empresa possui ainda 50% da Itaipu Binacional e também controla o Centro de Pesquisas de Energia Elétrica (CEPEL), o maior de seu gênero no Hemisfério Sul (ELETROBRÁS, 2011).

transmissão de energia elétrica, considerados como monopólios naturais, sob regulação do Estado (CCEE, 2012).

Observou-se também a necessidade de criação de um novo órgão regulador e fiscalizador de todas as relações no setor³ (Agência Nacional de Energia Elétrica - ANEEL), de um operador para o sistema elétrico nacional, o qual deveria controlar de forma integrada toda a operação do sistema elétrico e que iniciou seu funcionamento em agosto de 1998 (Operador Nacional do Sistema Elétrico - ONS) e de um ambiente para a realização das transações de compra e venda de energia elétrica (Mercado Atacadista de Energia Elétrica – MAE)⁴, o qual começou a atuar, com muitas restrições, a partir de 2000.

(Goldemberg & Prado, 2003) destacam que o insucesso da reforma do setor elétrico (Modelo de Livre Mercado) aconteceu não apenas devido à falta de recursos externos⁵ ou às resistências políticas encontradas no próprio governo, mas, e principalmente, às falhas de gerenciamento estratégico, coordenação e planejamento do sistema elétrico induzidas pela adoção de uma reforma calcada em experiências de outros países e inadequada às características brasileiras e ao sistema predominantemente hidrelétrico.

Em 2001, devido às falhas da reforma, a insustentabilidade do modelo e a problemas hídricos (baixo volume de chuvas), o setor elétrico sofreu uma grave crise de abastecimento que culminou em um plano nacional de racionamento de energia elétrica, onde foram afetadas todas as categorias de consumidores (residencial, industrial, comercial, rural, poder público, iluminação pública, serviço público e consumo próprio). Como a crise foi tratada pelo Governo tardiamente, as medidas de resultados imediatos foram as de controle do consumo. Quanto ao governo, este se concentrou na viabilização dos projetos de construção das usinas termelétricas e reforçou o orçamento de investimento das empresas estatais.

Em 2003, os problemas financeiros se agravaram junto às distribuidoras levando o governo a viabilizar a implantação de um programa de antecipação de

³ No caso do setor elétrico, o órgão regulador se chamava DNAEE – Departamento Nacional de Energia Elétrica (1967 a 1997), sendo substituído pela ANEEL em 26/12/1996, pela Lei nº 9427.

⁴ Mercado de curto prazo (spot) criado para permitir o encontro de contas para as inevitáveis diferenças entre quantidades contratadas e quantidades efetivamente consumidas ou geradas (CCEE, 2012).

⁵ Havia grande dependência do setor externo para o financiamento do setor elétrico, tal fato, confirmou-se com a crise do México em março de 1995 e a crise Asiática em outubro de 1997, causando grande fuga de capitais ((Canuto, 2000); (Marreco, 2007)).

recursos via BNDES para as concessionárias de distribuição. Além das medidas de emergenciais de curto prazo, foi elaborado um Novo Modelo institucional para o setor energético (instituído pela Lei N° 10.848, de 15 de março e decreto N° 5.163 de 30 de julho de 2004), com o propósito de corrigir as falhas que ocasionaram a crise, com foco principal nas questões de modicidade tarifária, universalização do acesso e na retomada do planejamento energético (Quadro 2.1).

O Novo Modelo definiu a criação de uma entidade responsável pelo planejamento do setor elétrico a longo prazo (a Empresa de Pesquisa Energética – EPE), uma instituição com a função de avaliar permanentemente a segurança do suprimento de energia elétrica (o Comitê de Monitoramento do Setor Elétrico – CMSE) e uma instituição para dar continuidade às atividades do MAE (Mercado Atacadista de Energia), relativas à comercialização de energia elétrica no Sistema Interligado (a Câmara de Comercialização de Energia Elétrica - CCEE).

Quanto à comercialização de energia, foram instituídos dois ambientes para celebração de contratos de compra e venda de energia: o Ambiente de Contratação Regulada (ACR), do qual participam Agentes de Geração e de Distribuição de energia; e o Ambiente de Contratação Livre (ACL), do qual participam Agentes de Geração, Comercializadores, Importadores e Exportadores de energia e Consumidores Livres (CCEE, 2012).

O Novo Modelo do setor elétrico previa um conjunto de medidas a serem observadas pelos Agentes, como a exigência de contratação de totalidade da demanda por parte das distribuidoras e dos consumidores livres, nova metodologia de cálculo do lastro para venda de geração, contratação de usinas hidrelétricas e termelétricas que assegurem melhor equilíbrio entre garantia e custo de suprimento, bem como o monitoramento permanente da continuidade e da segurança de suprimento, visando detectar desequilíbrios conjunturais entre oferta e consumo (CCEE, 2012).

Em termos de modicidade tarifária, o modelo previa a compra de energia elétrica pelas distribuidoras no ambiente regulado por meio de leilões – observado o critério de menor custo, objetivando a redução do custo de aquisição da energia elétrica a ser repassada para a tarifa dos consumidores cativos.

Observa-se no (quadro 2.1), um resumo das principais mudanças entre os modelos pré-existentes e o modelo atual, que acabaram por resultar em transformações nas atividades e de alguns agentes do setor.

Quadro 2.1 – Comparativo entre os Modelos Energéticos

Características	Modelo Antigo (até 1995)	Modelo de Livre Mercado (1995 a 2003)	Novo Modelo (a partir de 2004)
<i>Financiamento</i>	recursos públicos	recursos públicos e privados	recursos públicos e privados
<i>Estrutura das empresas</i>	Verticalizadas	geração, distribuição, transmissão e comercialização	geração, distribuição, transmissão e comercialização
<i>Relação Público vs Privado</i>	Empresas predominantemente Estatais	Abertura e ênfase na privatização das empresas	Convivência entre empresas Estatais e Privadas
<i>Estrutura de mercado</i>	Monopólios – competição inexistente	Competição na geração e comercialização	Competição na geração e comercialização
<i>Consumidores</i>	Cativos	cativos e livres	cativos e livres
<i>Estrutura tarifária</i>	Tarifas reguladas em todos os segmentos	Preços livremente negociados na geração e comercialização	Ambiente de Contratação Livre e Regulada
<i>Mercado</i>	Regulado	Livre	Convivência entre o mercado livre e o regulado
<i>Planejamento da Expansão</i>	Grupo Coordenador do Planejamento dos Sistemas Elétricos	Conselho Nacional de Política Energética (CNPE)	Empresa de Pesquisa Energética (EPE)
<i>Contratação</i>	100% do mercado	95% do mercado (até dez/2004)	100% do mercado + reserva
<i>Sobras/Déficits do balanço energético</i>	Rateados entre compradores	liquidados no MAE	liquidados na CCEE

Fonte: (CCEE, 2012)

2.2

O Setor Elétrico sob o Novo Modelo

Sob o Novo Modelo, o Governo Federal definiu um conjunto de agentes institucionais, com competências e atribuições bem definidas, objetivando garantir o bom funcionamento setorial e alcançar os três objetivos almejados, a saber, modicidade tarifária, segurança no suprimento e universalização do acesso (Tolmasquim, 2011).

Segundo o autor, os agentes podem ser classificados em três níveis, seguindo a natureza jurídica do ente e suas competências institucionais. Assim, definem-se agentes que executam atividades do governo, que executam atividades regulatórias e entidades de direito privado que executam atividades especiais.

As atividades do governo são exercidas pelo Conselho Nacional de Política Energética (CNPE), pelo Ministério de Minas e Energia (MME) e pelo Comitê de Monitoramento do Setor Elétrico (CMSE). As atividades de regulação são exercidas pela Agência Nacional de Energia Elétrica (ANEEL) e, adicionalmente, entes operacionais executam as atividades técnicas, as quais sejam planejamento da expansão do setor (EPE), planejamento e programação da operação (ONS) e a viabilização das atividades de comercialização (CCEE). Na (Figura 2.1) pode-se observar a relação entre os agentes institucionais.

Como já destacado na introdução desse trabalho e discutido nas próximas seções, o tema tratado nessa tese interfere diretamente nas atividades da EPE, do ONS e da CCEE que utilizam séries sintéticas de ENA⁶.

Dando continuidade, o próximo passo desse estudo é mostrar ao leitor como entidades operacionais, com diferentes finalidades, estão interligadas à estocasticidade associada ao módulo NEWAVE, que serve de suporte tanto para o planejamento da expansão (EPE), quanto para a operação do sistema (ONS) e valoração do preço da energia de curto prazo (CCEE).

⁶ Em linhas gerais, denomina-se ENA (Energia Natural Afluente) a quantidade de eletricidade que pode ser gerada pelo parque hidrelétrico com a água que chega às centrais hidrelétricas. Esta energia é estimada assumindo que o nível dos reservatórios esteja num patamar médio de 65% de sua capacidade total e uma política de operação. Vale lembrar que este valor pode mudar de acordo com a política de operação (Terry, et al., 1986).

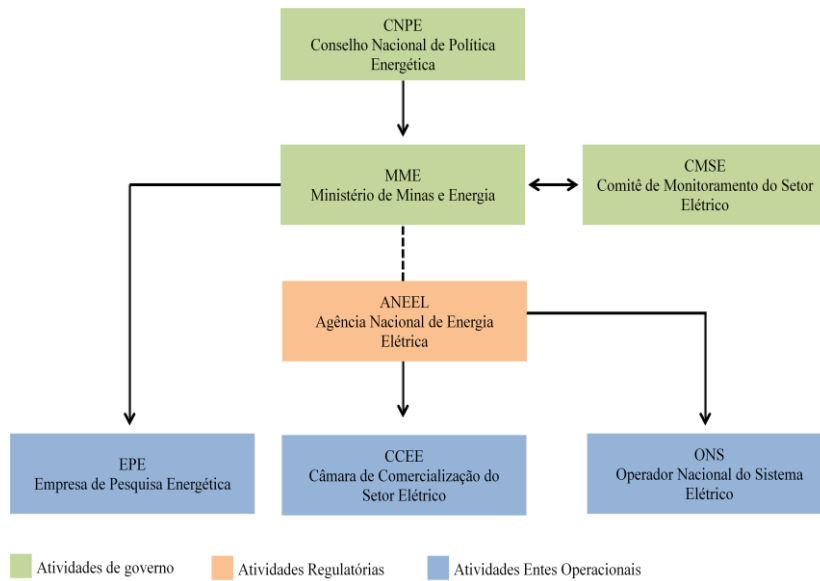


Figura 2.1 – Agentes institucionais do Setor Elétrico Brasileiro

Fonte: (Ministério de Minas e Energia, 2013)

Para alcançar tal objetivo de maneira clara e consistente, em primeiro lugar apresentam-se as características do setor elétrico brasileiro abordando pontos relacionados à execução das três atividades mencionadas acima, em seguida, abordam-se separadamente tais atividades, mostrando como o Módulo de Energias Afluentes, vinculado ao modelo NEWAVE, auxilia na execução das mesmas.

O setor elétrico brasileiro apresenta características que fazem com que a operação e o planejamento do sistema sejam uma atividade complexa e de difícil implementação. Segundo (Marreco, 2007), tanto os aspectos técnicos quanto econômicos devem ser salientados. Isto porque, o setor energético possui características que devem ser ressaltadas em relação a outros setores, tais como, a de monopólio natural (transmissão e distribuição) e de indústria de rede.

No caso do setor elétrico, a subaditividade da produção elétrica e a especificidade dos ativos de transporte tornam ineficientes pressões competitivas visando aumentar o número de agentes em certas atividades. Isto porque, em monopólio natural, a regulação da qualidade e do preço dos serviços de energia é que importam para um atendimento eficiente do mercado; sobretudo em indústrias de rede, onde os custos se reduzem quando se aumenta o número de consumidores conectados à rede, ou seja, os custos marginais de longo prazo tendem a ser

decrecentes para as atividades de transmissão e distribuição, por exemplo. Outra especificidade adicional é a incapacidade de estocagem de energia elétrica.

Para o caso brasileiro, destacam-se ainda: (i) base hidrelétrica, com usinas em cascata e proprietários distintos; (ii) elevada intensidade de capital; (iii) longo prazo de maturação dos investimentos; (iv) grandes interligações e, (v) grandes incertezas (hidrológicas, crescimento da demanda, preços dos combustíveis, dentre outros).

Com relação à elevada intensidade de capital e ao longo período de maturação dos investimentos, destaca-se que o tempo entre a decisão de construção de uma usina hidrelétrica e a sua efetiva entrada em operação, pode ser superior a dez anos. Outra característica são os custos afundados (ou irrecuperáveis), devido ao alto grau de especialidade das atividades vinculadas ao setor (*e.g.* linhas de transmissão de energia elétrica). Destaca-se, nesse ponto, a importância do planejamento de longo prazo feito pela EPE.

Outro atributo técnico fundamental do setor é o equilíbrio físico, o qual requer a coordenação do sistema, isto porque suas partes operam com forte interdependência. A geração de energia elétrica pode ser obtida por meio de várias tecnologias com diferentes custos e impactos sócio-ambientais. No caso das hidrelétricas, por exemplo, o insumo é um fluxo aleatório baseado no regime de chuvas, o que implica em grande complexidade no processo de planejamento de curto e médio prazos, tarefa essa executada pelo ONS.

Ainda, mesmo tendo um sistema eficiente de planejamento da demanda e da oferta de energia elétrica, há diferenças entre a energia efetivamente produzida e consumida pelo mercado. Para resolver este problema, há um mercado de energia de curto prazo que tem como objetivo final resolver as diferenças entre a oferta e a demanda, o qual é operado pela CCEE.

Graças a um processo histórico de aproveitamento dos recursos naturais para geração de energia, desenvolveu-se o Sistema Interligado Nacional (Figura 2.2). O SIN, de dimensões continentais, permite, respeitado certas configurações técnicas, que a energia gerada em qualquer ponto do país seja consumida por diferentes consumidores, em regiões distintas⁷. A interligação entre as regiões resulta em melhor aproveitamento dos recursos.

⁷ Apenas 3,4% da capacidade de produção de eletricidade do país encontram-se fora do SIN, em pequenos sistemas isolados localizados principalmente na região amazônica.

do preço *spot*⁸ da energia evidenciam a proeminência de se estudar o setor. Essas decisões, muitas vezes, são tomadas perante um ambiente de incertezas e necessitam de processos sistemáticos de apoio à decisão, em especial sobre as perspectivas do futuro.

Para confirmar o que foi abordado até o momento, as próximas três seções abordam o planejamento da expansão, o planejamento da operação, e a determinação do preço *spot* de energia separadamente, identificando características peculiares a cada segmento, e os relacionando com a estocasticidade inerente ao setor.

Para facilitar a exposição dos argumentos retratados anteriormente, primeiramente trata-se do planejamento da operação do sistema, uma vez que tal atividade está diretamente relacionada ao tema proposto nessa tese. Dando sequência, abordam-se o planejamento da expansão e a determinação do preço *spot* da energia, mostrando como a estocasticidade interfere nessas atividades.

2.2.1

O planejamento da operação do Sistema Elétrico

O sistema de geração de energia elétrica do Brasil, com aproximadamente 105.000 MW instalados, é basicamente hidrelétrico. Em 2009, a participação dessa fonte energética foi de aproximadamente 80.000 MW incluindo a parte brasileira da UHE Itaipu (7.000 MW) e a importação não consumida pelo sistema elétrico Paraguaio (aproximadamente 6.000 MW), o que representa aproximadamente 70% da capacidade instalada de geração (e, em condições normais, é responsável pela produção de aproximadamente 90% da eletricidade consumida no país). Por sua vez, o setor termoeletrico e o setor termonuclear representam 18% e 2% da capacidade de geração, respectivamente (EPE, 2012).

Neste contexto, em um sistema hidrotérmico em que se tem, de um lado, a disponibilidade futura de geração hídrica incerta, a um custo de geração considerado zero, e de outro, a geração térmica futura supostamente conhecida, a um custo de geração elevado, fica evidente a necessidade de decidir a cada instante o nível de repartição térmica/hidráulica.

⁸ Preço de curto prazo da energia elétrica determinado pela CCEE.

Na operação de sistemas hidrotérmicos⁹, ao contrário do que acontece em sistemas puramente térmicos onde o problema de planejamento da operação pode ser resolvido apenas encontrando uma combinação entre as plantas que minimizam os custos dos combustíveis¹⁰, a tomada de decisão é acoplada no tempo, isto é, uma decisão tomada hoje terá consequências no futuro (Terry, et al., 1986).

Por exemplo, se ocorrer um despacho hidrelétrico expressivo antes de um período seco, corre-se o risco de despachar térmicas de custo mais elevado no futuro; por outro lado, um despacho térmico anterior a um período úmido poderá provocar vertimentos, incorrendo em desperdício de energia (Figura 2.3).

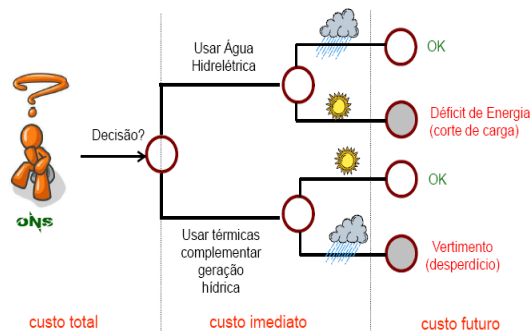


Figura 2.3 – Problema de decisão da operação

Fonte: (ONS, 2012)

Dessa forma, a operação equilibrada¹¹ do sistema envolverá um compromisso entre deplecionar (usar água) e não deplecionar (usar térmica) os reservatórios. A variável de decisão é o volume de água armazenado no fim do período da operação (volume final).

Tal decisão tem um custo imediato, associado à geração térmica (Função de custo Imediato - FCI) e um custo futuro, associado à expectativa de despacho térmico, indicado pela função de custo futuro (FCF). O custo total é a soma desses custos e a decisão ótima é obtida anulando-se a derivada do custo total em relação ao volume armazenado final (Pereira, et al., 1998).

⁹ O planejamento deverá contemplar também os usos múltiplos das águas nas cachoeiras, como por exemplo, a manutenção de níveis de navegação, irrigação, controle de cheias e saneamento.

¹⁰ Obviamente há outros fatores adicionais que podem tornar o problema mais complexo, como por exemplo, perdas de energia, limites de transmissão, custos de início de operação (Pereira, et al., 1998).

¹¹ O sistema é equilibrado quando é possível gerar energia suficiente para atender a demanda, com um pequeno risco de falha (Kelman, 2001).

Como se pode observar na (Figura 2.4), a FCI cresce à medida que o volume armazenado final aumenta, isto é, quanto maior este volume, maior o gasto com térmicas. A FCF tem comportamento contrário a FCI, ou seja, quanto maior o estoque final, menor é a expectativa de gastos com combustíveis no futuro.

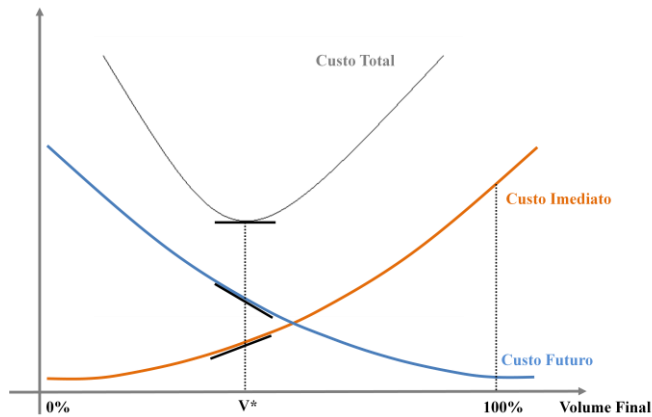


Figura 2.4 – Critério de Planejamento da Operação

Fonte: (Pereira, et al., 1998)

Dessa forma, a coordenação do planejamento da operação de um sistema hidrotérmico, como o SIN, pode ser desenhada como um problema de otimização de grande porte, com acoplamento temporal e espacial da operação, dinâmico, estocástico, interconectado e não linear. Sua solução requer que o problema seja decomposto em uma cadeia de modelos acoplados que considerem horizontes de longo prazo (probabilidade de armazenamento futuro de energia, valor futuro esperado de geração térmica etc.), médio prazo (contratos anuais para reserva na ponta etc.) e curto prazo (controle de fluxo, despacho horário etc.) ((Pereira & Pinto, 1982); (Pereira & Pinto, 1983); (Lepecki & Kelman, 1985); (Terry, et al., 1986)).

O Setor Elétrico Brasileiro adota, para a solução deste problema, uma cadeia de modelos acoplados que consideram diferentes horizontes de planejamento: longo prazo, médio prazo, curto prazo e programação diária (Maceira, et al., 2002). O acoplamento entre os modelos é feito através da função de custo futuro da operação energética.

Mais especificamente, define-se como planejamento da operação os esforços para delinear o comportamento do sistema em um horizonte de até cinco

anos, promovendo o aproveitamento racional dos recursos a fim de garantir a qualidade e a segurança no atendimento à demanda do mercado, os requisitos operativos do sistema hidrotérmico e a minimização dos custos operativos.

A programação da operação tem por objetivo estabelecer a operação de curto prazo do sistema hidrotérmico fornecendo decisões operativas do sistema de geração que sejam factíveis ao sistema de transmissão e que respeitem as metas estabelecidas no planejamento da operação.

Para solucionar os modelos adotados na cadeia de planejamento energético do SEB são usados algoritmos de Programação Dinâmica Dual Estocástica (PDDE) (Pereira & Pinto, 1991); (Pereira & Pinto, 1985); (Pereira, 1989); (Shapiro, et al., 2012)). Esta metodologia utiliza a técnica de Decomposição de Benders ((Benders, 1962); (Pereira & Pinto, 1983)) buscando encontrar estratégias ótimas para a operação dos subsistemas interligados, enquanto as afluições são tratadas a partir de um modelo autorregressivo periódico (Hipel & McLeod, 1994).

A PDDE é utilizada para a determinação da política operativa que minimiza o valor esperado do custo esperado da operação para o horizonte de até cinco anos. A estratégia é verificada através do processo de simulação que utiliza séries de energias afluentes representativas de cada subsistema, que podem ser históricas ou sintéticas. O parque gerador hidrelétrico é representado por reservatórios equivalentes, no qual um conjunto de usinas hidrelétricas é agregado num único reservatório equivalente que recebe, armazena e descarrega energia (Energia Natural Afluente) ao invés de água; e o termelétrico através de custos de operação das térmicas e geração mínima e máxima. O modelo NEWAVE fornece como resultado uma função de custo futuro que é acoplada ao módulo de curto prazo no final do horizonte de planejamento (Figura 2.5).

Objetivando reduzir o esforço computacional exigido pelos modelos de otimização, os modelos utilizados para o planejamento da operação no médio prazo utilizam a agregação das usinas em reservatórios equivalentes de energia. Cada subsistema (Sudeste/Centro-Oeste, Sul, Norte e Nordeste¹²) do SIN é

¹² Atualmente, há uma discussão no SEB para aumentar o número de subsistemas considerando a incorporação ao SIN, por exemplo, das bacias dos rios Madeira, Tapajos e Xingu, contudo, neste trabalho serão considerados apenas os quatro subsistemas tradicionais.

representado por um reservatório equivalente que é composto pelas usinas de sua região.

Porém, para verificar se a política de operação obtida pelo modelo de decisão estratégica será viável, é necessário desagregar a solução obtida para os reservatórios equivalentes de energia em usinas individualizadas, ou seja, verificar se as usinas que compõem o sistema equivalente serão capazes de atender ao montante de geração hidráulica determinado para o sistema pelo modelo de decisão estratégica (programação de curto prazo). Em seguida, ajusta-se a proposta de despacho semanal (e diário) no contexto de programação diária.

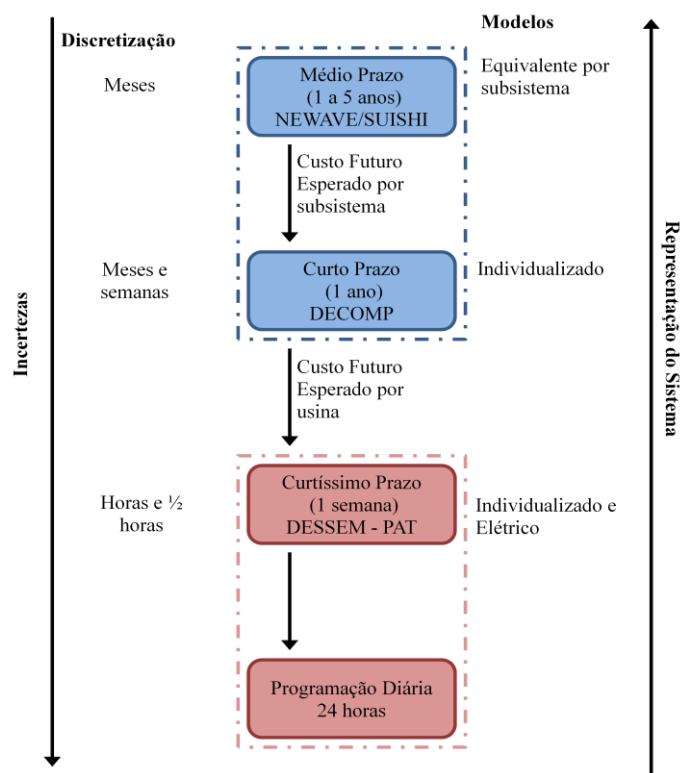


Figura 2.5 – Cadeia de Modelos do Setor Elétrico Brasileiro

Fonte: (ONS, 2012)

É importante destacar que neste estudo o foco será no módulo estocástico que dá suporte ao modelo de planejamento NEWAVE. Conforme observado, tal modelo é utilizado, em linhas gerais, para otimizar o uso da água e, para isso, depende da estocasticidade, a qual é representada por séries sintéticas geradas a partir de modelos de séries temporais que estão contemplados no módulo de Energias Afluentes.

No curto prazo, os aspectos de qualidade e atendimento às demandas consistem do atendimento das metas de geração definidas no planejamento de médio prazo e da verificação da factibilidade operativa em termos de restrições dos equipamentos de geração e transmissão. Nesse horizonte de planejamento, o objetivo do modelo DECOMP é minimizar o valor esperado do custo total do sistema, conforme poderá ser mais bem observado nas seções posteriores.

Quando o horizonte de estudo é reduzido para o curto e curtíssimo prazo (DECOMP e DESSEM-PAT), a representação do sistema é refinada às usinas hidrelétricas, expressando suas características operativas e restrições hidráulicas e energéticas. Na programação da operação o horizonte é de até duas horas com discretização em horas ou trinta minutos (Figura 2.5). O objetivo é obter o despacho ótimo do sistema hidrotérmico que contemple os aspectos energéticos, hidráulicos e elétricos do sistema. O horizonte de curto e curtíssimo prazo (programação diária) permite considerar os valores de afluências conhecidos ou determinísticos (Maceira, et al., 2002).

Por último, é importante observar que à medida que diminui o intervalo de tempo considerado, as incertezas com relação às vazões\ENAS diminuem, chegando a uma extrapolação determinística, e o nível de detalhes de representação do sistema aumenta, na medida em que se consideram características do sistema individualizado.

Seguindo a lógica proposta no início dessa seção, apresentam-se os modelos NEWAVE, enfatizando a ligação entre tal modelo e o modelo de geração de série sintética (Módulo de Energias Afluentes); e o modelo DECOMP, importante para a determinação do preço da energia de curto prazo e que está acoplado ao modelo NEWAVE pela função de custo futuro.

Modelo NEWAVE

O NEWAVE é um modelo de otimização para o planejamento de médio/longo prazo, com discretização mensal e representação a sistemas equivalentes. Seu objetivo é determinar a estratégia de geração hidráulica e térmica em cada estágio que minimiza o valor esperado do custo de operação para todo o período de planejamento, a um determinado nível de confiabilidade.

Um dos principais resultados desse modelo são as funções de custo futuro, essenciais para a determinação do “valor da água” e conseqüentemente dos impactos da utilização da água armazenada nos reservatórios. No NEWAVE, a carga e a função de custo de déficit podem ser representadas em patamares¹³ e permite-se a consideração de limites de interligação entre os subsistemas (CCEE, 2012).

O NEWAVE é composto por quatro módulos operacionais: (i) Módulo de cálculo do sistema equivalente – calcula os subsistemas equivalentes de energia e, baseado nas séries históricas de vazões e nas características desses aproveitamentos, calcula a Energia Natural Afluente; (ii) Módulo de Energias Afluentes – estima os parâmetros do modelo estocástico e gera séries sintéticas de energia afluente, que serão utilizadas para o cálculo da política de operação hidrotérmica; (iii) Módulo de cálculo da política de operação hidrotérmica – determina a política de operação mais econômica para os subsistemas equivalentes, considerando as incertezas nas afluências futuras, a indisponibilidade dos equipamentos e os patamares de demanda; e, (iv) Módulo de simulação da operação – simula a operação do sistema ao longo do período do planejamento para diferentes cenários de sequências hidrológicas (CEPEL, 2011a).

Conforme se pode observar na (Figura 2.6) estes quatro módulos estão interligados. Por exemplo, o módulo de Energias Afluentes recebe séries históricas de ENAs do módulo de cálculo do sistema equivalente, estima o modelo $PAR(p)$, gera séries sintéticas para cada subsistema e fornece para os módulos de política de operação e de simulação os parâmetros gerados pelo modelo e as séries sintéticas de energia. De posse das variáveis estocásticas e dos parâmetros do modelo, o módulo de cálculo da política da operação, utilizando-se do ferramental de PDDE, determina a política ótima de operação e o módulo de simulação avalia a política calculada.

¹³ É importante observar que tais patamares podem ser de diferentes naturezas, isto é, (i) Patamar de carga, o qual é uma agregação dos valores de energia ao longo de um intervalo de tempo (e.g. mês) de forma a separar demandas baixas (carga leve), médias e alta. Em linhas gerais, é o período de tempo em que as características de consumo de energia elétrica tendem a ser semelhantes; e, (ii) Patamar da curva de custo de déficit, o qual como se fossem consideradas térmicas com custos diferentes, cada uma com uma determinada capacidade de “geração”, por exemplo, 5% da demanda do estágio para a primeira térmica déficit (ou o “primeiro patamar”). Para maiores detalhes ver: Boletim Técnico, (EPE, 2009).

Dado que as demais variáveis, tais como previsão de mercado, custo e disponibilidade e limites de transmissão entre os subsistemas são fixas, fica evidente que a única variável estocástica que é utilizada para a otimização do sistema hidrotérmico brasileiro são as séries sintéticas de energia geradas pelo módulo estocástico de energias afluentes. Em outras palavras, os cenários fornecidos pelo Módulo de Energias Afluentes, o qual engloba a etapa (ii) do módulo NEWAVE, em última instância, irão influenciar os três pilares básicos do SEB: planejamento da expansão, planejamento e programação da operação e a determinação do preço da energia.

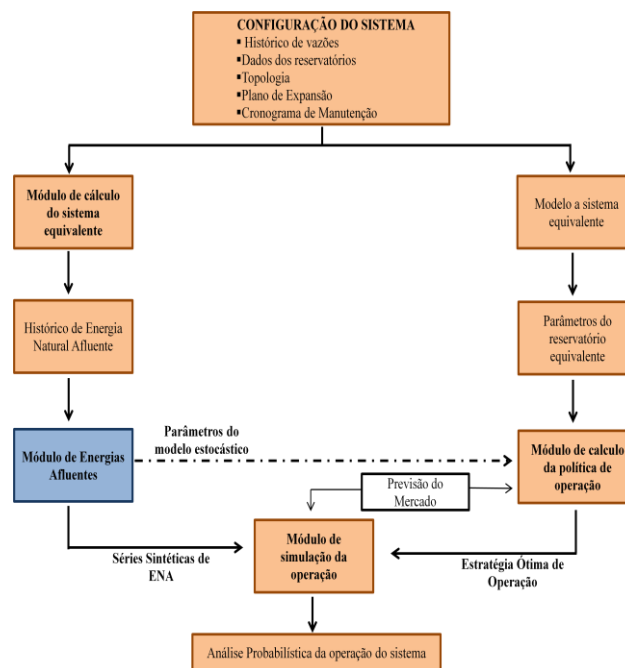


Figura 2.6 – Relação entre os módulos do NEWAVE

Fonte: (CEPEL, 2011a)

Modelo DECOMP

Assim como o NEWAVE, o modelo DECOMP também procura obter uma operação ótima do sistema hidrotérmico, mas num horizonte de tempo mais curto. O DECOMP resolve o problema do planejamento da operação no curto prazo de um sistema hidrotérmico. Conforme se pode observar na (Figura 2.7), utilizando a função de custo futuro obtida no NEWAVE e informações sobre carga, vazão, disponibilidades e limites de transmissão entre os submercados, o módulo

DECOMP obtêm o resultado otimizado para o planejamento do primeiro mês em base semanal. Suas principais características são o planejamento de curto prazo com discretização semanal no primeiro mês de estudo.

É importante observar que as vazões previstas e a aleatoriedade das vazões do restante do período são obtidas através de uma árvore de possibilidades e de um parque gerador individualizado (ao contrário do formato agregado utilizado no NEWAVE). Conforme observado, o módulo DECOMP é utilizado para a operação de curto prazo do SEB e, como acontece com o NEWAVE, depende fortemente dos modelos estocásticos, isto é, a estocasticidade influencia de maneira indireta o DECOMP através das FCF e influenciam diretamente na consideração da árvore de cenários de afluições.

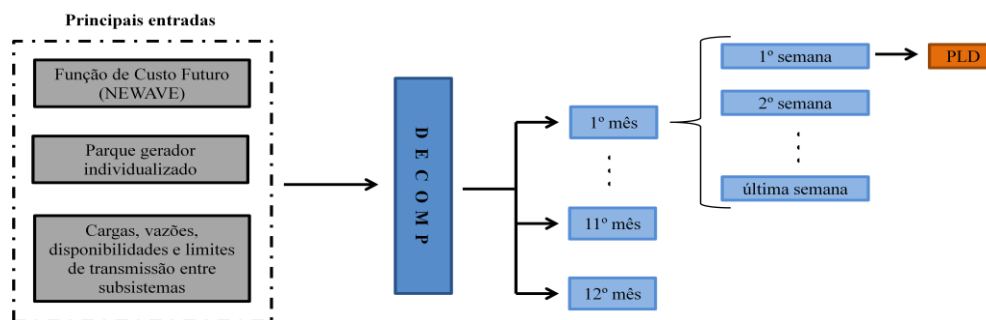


Figura 2.7 – Principais entradas e saídas do programa DECOMP

Fonte: (CEPEL, 2011b)

2.2.2

O Planejamento da Expansão do Sistema Elétrico

Segundo (Tolmasquim, 2011), assim como o planejamento da operação envolve o compromisso entre o uso imediato e o uso futuro da água, no caso do planejamento da expansão, tem-se a relação entre o uso futuro e imediato do capital financeiro disponível para a expansão.

Isto é, os agentes precisam decidir entre investir agora e correr o risco de ociosidade do sistema devido ao crescimento aquém do esperado da demanda, ou então, postergar o investimento e arriscar-se ao racionamento, conforme se pode observar na (Figura 2.8).

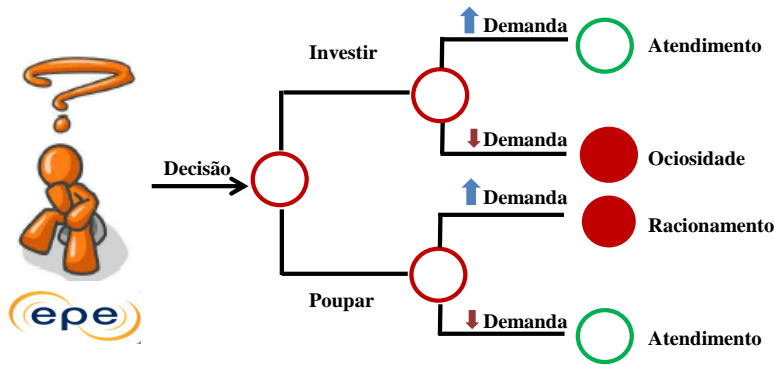


Figura 2.8 – A decisão do Investimento

Fonte: (ONS, 2012)

Seguindo a mesma lógica do uso da água na operação, a decisão de expandir a capacidade do sistema diminui o estoque de capital, com um custo imediato (custo de expansão, conhecido) e um custo futuro (custo de déficit, estimado).

Dessa forma, a decisão de “poupar” tem baixo custo imediato e alto custo futuro, este devido ao aumento de consumo de combustíveis e racionamentos. Por outro lado, a decisão de “investir” tem alto custo imediato e baixo custo futuro (Figura 2.9)¹⁴ (Tolmasquim, 2011).

Similar ao que acontece no processo de operação, no caso do planejamento, dado uma demanda prevista, há várias alternativas de expansão. Cada alternativa equivale a um nível de confiabilidade R (eixo das abscissas), onde o custo da expansão dependerá da demanda que é única e do nível de confiabilidade, R , desejado pelo planejador.

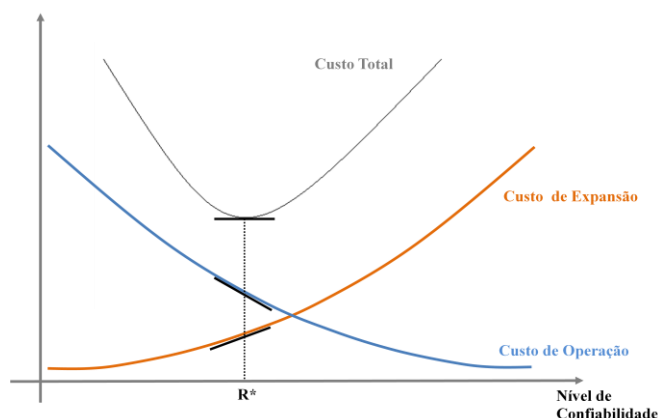


Figura 2.9 – Critério de planejamento da expansão

Fonte: (Tolmasquim, 2011)

¹⁴ Embora a análise seja apresentada de forma binária (ou investe ou não investe), na prática, a decisão de expansão comporta diversos níveis de investimento, que caracterizam sequências temporais alternativas de empreendimentos de geração e transmissão.

Assim como acontece com a operação, a solução para o processo de planejamento é dada por exercício de otimização, tendo como variável de decisão o nível de confiabilidade (R) e, como função objetivo o custo total (CT - soma dos custos de expansão e de operação). Vale lembrar que, o nível de confiabilidade (R^*) e o plano de expansão são ótimos no ponto de mínimo custo total (Tolmasquim, 2011).

No plano ótimo, a derivada do custo de expansão (CE) em relação à demanda (D) representa o Custo Marginal de Expansão (CME), o qual está ligado ao custo de uma carga incremental, com expansão da capacidade. A derivada do custo de operação (CO) de relação à demanda representa o Custo Marginal de Operação (CMO), representando o custo para atender a carga incremental, sem expansão da capacidade.

Matematicamente, o último parágrafo pode ser mais bem entendido derivando e igualando a zero (Condição de Primeira Ordem) a equação de Custo Total ($\partial CT/\partial D = \partial CE/\partial D + \partial CO/\partial D = 0$). Supondo “ponto mínimo”, observa-se que o Custo Total é mínimo quando o ($CMgE = |CMgO|$). Ainda, observando a identidade matemática, é fácil verificar que aumentos Custo Marginal de Expansão deverão ser compensados por quedas no CMO .

Assim, dado que a demanda é crescente ao longo do tempo, a expansão do SEB acontece quando o módulo do CMO iguala-se ao CME . Ao expandir o sistema, o CMO decresce em relação ao CME e reinicia outro ciclo.

Como no planejamento da operação, para o planejamento da expansão utiliza-se o programa NEWAVE para dar suporte as tomadas de decisão. O cálculo do CMO é realizado por meio de simulações da operação do sistema, com base no modelo NEWAVE. Já o cálculo do CME é estimado com base nos resultados dos leilões de energia nova e na crença de que o lance vencedor do empreendimento mais caro do certame, que revela a disposição de investir dos agentes, constitui uma boa aproximação do mesmo.

Neste ponto, cabe ressaltar a dicotomia entre o cálculo dos custos marginais, isto é, por um lado, tem-se o CMO , calculado através do modelo NEWAVE que, conforme observado nessa tese, apresenta uma série de limitações e sua principal variável de decisão é baseada nas Energias Naturais Afluentes simuladas. Por outro, o CME é estimado mesclando decisões mercadológicas, tais como, WACC

(*weighted average cost of capital*), Risco Brasil, juros e parâmetros do modelo NEWAVE.

Por fim, é importante observar que, dado o tempo para a maturação dos investimentos no setor elétrico e as graves consequências de um racionamento, o planejamento da geração precisa atender ao critério de confiabilidade definido pelo Conselho Nacional de Política Energética (CNPE), segundo o qual o risco anual de déficit não deva superar 5%¹⁵ em nenhum subsistema. Ainda, para atender ao critério de economicidade, é preciso haver igualdade entre o CMO e o CME.

Em resumo, a estocasticidade das afluições está intrinsecamente ligada ao planejamento de expansão do SEB pois, conforme observado, a decisão de expandir o sistema está intimamente ligada ao CMO, o qual é calculado com base no modelo NEWAVE e tem como variável estocástica as séries sintéticas de energia¹⁶.

2.2.3

Comercialização de Energia Elétrica

O Novo Modelo dividiu o mercado brasileiro de energia em dois ambientes de comercialização: Ambiente de Contratação Regulada (ACR), o qual visa atender as demandas dos consumidores cativos e são representados por contratos a termo derivados de energia para o *Pool* de mercado; e o Ambiente de Contratação Livre (ACL), voltado para empresas com maior volume de consumo, no qual os contratos bilaterais são livremente negociados, conforme regras e procedimentos de comercialização específicos¹⁷ (Castro & Leite, 2010).

¹⁵ Para isso, simula-se no NEWAVE a operação futura do sistema, considerando 2000 cenários hidrológicos. O critério é atendido se não ocorrerem déficits em pelo menos 1900 (95%) dos dois mil cenários gerados.

¹⁶ Há outras variáveis estocásticas que podem influenciar no planejamento da expansão, tais como, demanda de energia, atraso em obras, custos operacionais, etc.

¹⁷ É importante observar que o modelo elétrico vigente determina a obrigatoriedade de lastro por parte dos agentes do setor, isto é, exige a contratação da totalidade da demanda por parte das distribuidoras, comercializadores e consumidores livres, e ainda, exige dos agentes vendedores de energia lastro para os contratos de compra e venda. Como a demanda é prevista, a demandante só será penalizada se o montante contratado for 3% inferior ao montante consumido. Por outro lado, caso montante contratado ultrapasse o montante consumido em até 2% a CCEE compra esse energia ao preço do PLD, caso ultrapasse esse percentual a concessionária deverá vendê-la no ACL.

Contudo, devido ao atributo físico da energia elétrica necessitar de equilíbrio instantâneo entre demanda e oferta, a oferta prevista *ex-ante* não necessariamente iguala a demanda observada, fazendo-se necessário o equilíbrio instantâneo em dois pontos: o fornecimento e a contabilização financeira (Rodrigues, 2007).

Quanto ao primeiro, o ONS, responsável pelo planejamento da operação do sistema, centraliza o despacho das usinas por meio da agregação dos empreendimentos de geração e transmissão de forma a impetrar uma gerência mais efetiva no uso dos recursos, e consequente, minimização do custo da energia. Neste caso, uma geradora, mesmo tendo toda a sua energia contratada, pode não ter que ofertá-la ao sistema devido às decisões do operador.

O segundo é função da Câmara de Comercialização de Energia Elétrica, a qual contabiliza as diferenças entre o que foi produzido ou consumido e o que foi contratado. As diferenças positivas ou negativas são liquidadas no mercado de curto prazo (*spot*), valoradas ao PLD (Preço de Liquidação das Diferenças), que é determinado semanalmente para cada patamar de carga e para cada submercado, tendo como base o CMO do subsistema. O PLD é limitado por um preço mínimo e por um preço máximo.

Como observado, o preço nesse mercado não obedece à relação econômica entre oferta e demanda dos agentes, mas é determinado por um conjunto de modelos computacionais (*e.g.* NEWAVE, DECOMP), operados pelo ONS e pela CCEE. As expectativas quanto ao consumo futuro de eletricidade e quanto ao regime futuro para a ENA têm papel determinante no uso da energia acumulada nos reservatórios hidrelétricos, e por consequência, também no preço da energia no mercado de curto prazo (*spot*). Logo, o custo mínimo esperado num dado horizonte de tempo deve levar em consideração diferentes cenários de afluência que, por sua vez, levam a diferentes decisões operativas.

O modelo NEWAVE calcula a política ótima de operação do sistema, considerando custos presentes e futuros. A predominância de usinas hidroelétricas no parque gerador brasileiro torna a questão da aleatoriedade das vazões importante no problema de otimização do custo total de operação no horizonte de tempo considerado, pois o custo futuro é função de afluências futuras, aleatórias, enquanto o custo imediato é função do despacho atual das usinas térmicas e hidráulicas, este último considerado zero. Deve-se notar, no entanto, que este

custo marginal só será zero (correspondendo ao custo da água, ou seja, da geração hidráulica), em situações de reservatórios cheios no presente e no futuro, isto é, decorrentes de séries de afluência com hidrologia muito favorável (Barros, et al., 2009).

Assim, após discorrer sobre o planejamento da operação, o planejamento da expansão e sobre a determinação do preço *spot* da energia, tornou-se evidente a forte relação existente no setor elétrico brasileiro entre a estocasticidade e os três pilares do setor, isto é, as séries sintéticas de energia são cruciais para determinar qual é a melhor maneira de operar o setor, subsidiam decisões sobre quando se deve ou não expandi-lo, evitando-se assim custo e/ou perdas desnecessárias. E, ainda, são um fator preponderante na determinação do preço de curto prazo de energia elétrica, dado que a quantidade simulada/prevista de água nos reservatórios no futuro irá determinar o preço da energia no curto prazo.

No próximo capítulo são abordados modelos da família Box & Jenkins no contexto de simulações de vazões/energia, conceitos fundamentais sobre a teoria de séries temporais e os modelos Autorregressivos Periódicos (PAR(p)), os quais são utilizados na geração de séries sintéticas utilizadas no modelo NEWAVE.

3

Referencial Teórico

O capítulo 3 inicia-se fazendo uma revisão dos métodos e modelos aplicados ao problema de simulação de vazões/ENAs, objetivando contextualizar o leitor quanto à aplicação que será feita nesse trabalho.

Nas duas seções seguintes abordam-se, sucintamente, métodos e modelos aplicados no contexto de simulação de vazões/ENAs e o modelo PAR(p). Ambas as seções fazem-se necessárias, porque, em primeiro lugar, a aplicação dos modelos Gama propostos nesta tese fazem parte do conjunto de modelos aplicados à solução do problema de simulação/previsão. E, segundo, porque os modelos PAR(p) são usados pelo SEB para a simulação de séries sintéticas e suas limitações e peculiaridades para atender as necessidades do setor serão abordadas no capítulo 4 dessa tese.

3.1

Métodos e Modelos aplicados no contexto de simulação de vazões/energia

Segundo (Salas & Obeysekera, 1982), modelos de séries temporais aplicados às séries hidrológicas são uma importante ferramenta para o planejamento e a operação de sistemas hidrotérmicos. Há na literatura diversos trabalhos dedicados à solução deste problema e aplicados em diferentes horizontes de tempo, isto é, modelos anuais, mensais, semanais e diários.

Entre os métodos mais referenciados para aplicação no problema de simulação/previsão de vazões/ENAs podem-se citar: (i) Autorregressivo Médias Móveis (ARMA(p,q)), (ii) Autorregressivo (AR(p)), (iii) ARMA (p,q) multivariado; (iv) ARMA Periódico (PARMA(p,q)), (v) Autorregressivo Periódico (PAR(p)); (vi) Modelos memória longa (ARFIMA(p,d,q)); (vii) Gama Autorregressivo (GAR(1)), (xiii) Periódico Autorregressivo Aditivo Gama (PAGAR(1)); (ix) Periódico Autorregressivo Multiplicativo Gama (PMGAR(1)) e (x) Periódico Autorregressivo Gama (PGAR(1)).

Quanto à aplicação dos modelos AR a este tipo de problema e, de maneira geral, dos modelos da família Box & Jenkins, tem-se, como precursores (Thomas & Fiering, 1962), os quais propuseram um modelo AR de ordem 1 para modelar vazões, e (Hannan, 1955), que utilizou este mesmo modelo para modelar precipitações mensais em rios da Austrália no período de 1859 a 1952 (Quadro 3.1).

Desde então, muitos autores dedicaram-se a aplicar tal metodologia para a solução deste problema e propuseram pontos de melhoria na aplicação, como por exemplo, na estimação dos parâmetros e na definição da ordem dos modelos. Alguns importantes trabalhos que utilizaram modelos AR(p) e ARMA(p,q) para modelar séries hidrológicas anuais, mensais e diárias foram (Beard, 1965), (Beard, 1967); (Stedinger & Taylor, 1982); (Stedinger, et al., 1985); (Matalas, 1967); (Srikanthan & McMahon, 1978); (Salas, et al., 1982) e (Salas & Obeysekera, 1982). Por exemplo, (Beard, 1965) e (Beard, 1967) abordam as vantagens estatísticas relacionadas à simulação de vazões para um posto de aflúências e, dando sequência ao trabalho, expande tal análise para um número ilimitado de estações considerando que as mesmas são correlatadas (quadro 3.1).

Mais recentemente, (Bacanli, 2012) aplica tanto os modelos AR(p) e ARMA(p,q) para modelar médias mensais de vazões no rio Goksu, localizado no sul da Turquia. O interessante desse artigo é que o autor aborda três diferentes métodos para a seleção do modelo mais adequado, os quais sejam função de autocorrelação, variância residual mínima e o critério de informação de Akaike. Tal análise torna-se interessante sob o ponto de vista de que deve-se considerar mais de um critério para definir a ordem do modelo, ao contrário do que acontece no Setor Elétrico Brasileiro.

Na classe de modelos periódicos, destaque para os modelos PAR (*Periodic Autoregressive*) e PARMA (*Periodic ARMA*). O modelo PAR (p) ajusta para cada período da série um modelo AR (p). De maneira similar, um PARMA (p,q) consiste num modelo ARMA (p,q) para cada período em estudo.

Mesmo com a afirmação de (Rasmussen, et al., 1996) de que a extrapolação dos modelos PAR(p) para os modelos PARMA (p,q) não é uma tarefa trivial e pode não ser justificável dado o bom desempenho dos modelos autorregressivos, há diversos trabalhos que aplicaram a metodologia PARMA(p,q), como, (Salas, et al., 1982); e (Tesfaye, 2005); (Shao & Lund, 2004); (Vecchia, 1985a) e (Vecchia,

1985b). Em (Vecchia, 1985b), o autor discute a estimação dos parâmetros do modelo PARMA via métodos dos momentos e máxima verossimilhança, concluindo que o segundo método é mais adequado. (Tsfaye, 2005), propõe um modelo parcimonioso para a estimação dos parâmetros do modelo PARMA e aplica tal método a séries mensais de vazões dos rios Frazer (British Columbia) e Salt (Arizona) (quadro 3.1).

No Brasil, pode-se citar como exemplo (Maceira, et al., 1999), no qual os autores fazem previsões de vazões semanais baseados no modelo ARMA, considerando parâmetros periódicos e não periódicos. O interessante dessa aplicação é que os autores utilizam o software PREVIVAZ (PREVIVAZ, 2000), desenvolvido pelo Centro de Pesquisas em Energia Elétrica (CEPEL, 2011b) e utilizado pelo Operador Nacional do Sistema (ONS, 2012) no planejamento da operação.

O modelo periódico autorregressivo (PAR(p)) é a ferramenta utilizada pelo SEB (NEWAVE) para a geração dos cenários de afluência que serão utilizadas na PDDE para dar origem às funções de custo futuro ((CEPEL, 2011a); (Maceira, et al., 2005)). Como resultado da adoção de tais modelos pelo SEB há no Brasil uma diversidade de trabalhos que abordam tal metodologia para os mais diversos tipos de aplicação e/ou discussão metodológica.

Basicamente, esses trabalhos podem ser divididos em três grandes grupos: (a) trabalhos que discutem a metodologia; (b) trabalhos que utilizam o PAR(p) para comparar com um método “concorrente” e, (c) trabalhos que utilizam tal modelo como subsídio para discussões sobre o processo de otimização. Entre eles podem ser citados: ((Souza, et al., 2012); (Oliveira & Souza, 2011) (Ferreira, et al., 2012); (Marien, 2001); (Maceira & Damázio, 2004) (Maceira, et al., 2005); (Pereira, 2011); (Campos, 2010); (Ballini, et al., 1997); (Lima & Lall, 2010); (Sarnaglia, et al., 2008); (Vieira, et al., 1998); (Penna, 2009); (Shapiro, et al., 2012)) (Quadro 3.1).

Apesar da ampla aplicação dos modelos PAR(p) no Brasil, é importante lembrar que os mesmos foram originalmente introduzidos por (Thomas & Fiering, 1962) e, desde então, vêm sendo amplamente discutidos na literatura internacional. (Jones & Brelsford, 1967) e (Hipel & McLeod, 1994) contribuíram significativamente na formalização de tais modelos, paralelamente, vários autores contribuíram para a evolução do método, como por exemplo, na discussão da

“melhor” maneira para estimar os parâmetros (*e.g.* método dos momentos, máxima verossimilhança e métodos de inferência Bayesiana) e diversos estudos de caso. ((Vecchia, 1985b); (Pagano, 1978); (Salas, et al., 1982)).

Conforme observado, Salas J. D. contribuiu significativamente para a evolução na aplicação dos modelos da família Box & Jenkins no contexto de geração de séries sintéticas ((Fernandez & Salas, 1986); (Salas & Obeysekera, 1982) (Salas, et al., 1982); (Salas, et al., 2007); (Salas, 2012)). Em (Fernandez & Salas, 1986), os autores argumentam que, embora os modelos autorregressivos que pressupõem distribuição gaussiana dos resíduos tenham sido úteis na simulação de séries sintéticas, transformações (*e.g.* (Box & Cox, 1964) feitas nas séries históricas ((Stedinger, 1981); (Hirsch, 1979)) podem não levar a uma simulação não realista da assimetria dessas séries.

Os modelos periódicos autorregressivos Gama surgem na literatura basicamente com dois objetivos: o primeiro, para auxiliar na modelagem e geração de séries sintéticas hidrológicas. O segundo, baseado nas características dessas séries, para evitar transformações estatísticas necessárias para cumprir o requisito de normalidade. Fruto do segundo objetivo, a não transformação da série e a geração das séries sintéticas utilizando a distribuição Gama possibilitam a correção de um problema recorrente quando se usa a pré-suposição de normalidade, que é a manutenção da assimetria nas séries sintéticas simuladas.

É importante lembrar que, antes de (Fernandez & Salas, 1986), outros trabalhos foram elaborados para tentar resolver o problema da assimetria, como (McMahon & Miller, 1971), onde os autores utilizam o ruído da série para estimar a assimetria. Contudo, embora a assimetria possa ser reproduzida por essa aproximação, (Bobeé & Robitaille, 1975) e, mais tarde, (Fernandez & Salas, 1990) mostram que há um viés na assimetria estimada devido à dependência temporal presente na série.

Na década de 1980, (Gaver & Lewis, 1980) e (Lawrance, 1982) apresentaram um modelo Gama autorregressivo de ordem 1, conhecido como GAR (1). Segundo (Fernandez & Salas, 1986), em termos hidrológicos, foi a generalização mais importante até aquele momento, contudo, embora úteis, são mais adequados para séries anuais, as quais apresentam menor assimetria e tendem a ser mais bem comportadas.

Em 1986, (Fernandez & Salas, 1986) propõem modelos com parâmetros periódicos que pressupõe distribuição gama (e.g. Modelo Periódico Autorregressivo Aditivo Gama (PAGAR(1)), Modelo Periódico Autorregressivo Multiplicativo Gama (PMGAR(1)) e Modelo Periódico Autorregressivo Gama (PGAR(1)), extrapolando o problema das transformações para atingir normalidade e modelando a assimetria de forma mais realista.

Dessa forma, dentro do contexto de modelos periódicos da família Box & Jenkins, um dos objetivos desse trabalho é testar como um modelo que pressupõe distribuição gama se comporta na tarefa de modelar e simular séries sintéticas de Energia Natural Afluente - ENAs inseridas na realidade do SEB. Conforme será observado no decorrer deste trabalho, há peculiaridades na modelagem adotada pelo SEB que serão respeitadas nessa tese.

3.2

Modelos de Box & Jenkins¹⁸

Diferente dos modelos de regressão, nos quais Y_t é explicado por k regressores $X_1, X_2, X_3 \dots X_k$, os modelos de séries temporais do tipo Box & Jenkins permitem que Y_t seja explicado por valores passados, ou defasados, do próprio Y e dos termos de erro estocásticos. Devido a essa característica, os modelos ARIMA (autorregressivo integrado e de médias móveis) são chamados por alguns autores de modelos “aleatórios”, dado que são ateoréticos (Gujarati, 2006).

¹⁸ É importante observar que, por ser uma técnica amplamente utilizada, esta seção preocupou-se apenas em passar a ideia da modelagem Box & Jenkins. Para maiores detalhes ver: (Box, et al., 2008); (Hamilton, 1994).

Quadro 3.1 – Artigos aplicados ao problema de simulação/previsão de vazões/ENAs

Título	Autor(es)	Ano	Comentários
Mathematical Synthesis of streamflow sequences for the analysis of river basins by simulation	Thomas, H.A., Fiering, M.B.	1962	Modelo AR(1)/PAR(1) – considerados os precursores do modelo PAR(p), os autores aplicam modelos da família Box & Jenkins para o problema de simulação de vazões/ENAs.
A test for singularities in Sydney rainfall	Hannan, E. J.	1955	Modelo AR(1) – apesar de ter sido publicado antes, não é tão citado quanto Thomas e Fiering (1962).
Use of interrelated records to simulate streamflow	Beard, L. R.	1965	Modelo AR(p) – o autor aborda possíveis melhorias na simulação para múltiplos postos.
ARMA model identification of hydrologic Times Series	Salas, J. D., Obeysekera J. T. B.	1982	Modelo ARMA (p,q) – os autores discutem 3 métodos para a definição da ordem dos modelos ARMA (p,q) e fazem aplicações com séries de vazões anuais para os rios Saint Lawrence e Nilo.
Stochastic Modeling of Monthly Streamflow Data	Bacanli, U. G.	2012	Modelos ARMA(p,q) e AR(p) – define três diferentes critérios para ajustar o “melhor” modelo a série de dados.
Periodic Autoregressive - Moving Average (PARMA) modeling with applications to water resources	Vecchia, A.V.	1985	Modelo PARMA(p,q) – autor mostra vantagens da estimação dos parâmetros via máxima verossimilhança e faz uma aplicação para o Rio Caroni, Venezuela.
Seasonal Time Series Models and Their Application to the Modeling of River Flows	Tesfaye, Y. G.	2005	Modelo PARMA(p,q) – excelente tese de doutorado sobre os modelos PARMA (p,q)
Estimation of ARMA models with sazonal parameters	Salas <i>et. al.</i>	1982	Modelo PARMA(p,q) – sugerem a estimação dos parâmetros do modelo PARMA (p,q) utilizando equações de Yule-Waker sazonais.
Periodic ARMA Models Applied to Weekly Streamflow Forecasts	Maceira <i>et. al.</i>	1999	Modelo PARMA(p,q) – os autores apresentam previsões de vazões semanais baseados no modelo ARMA, considerando parâmetros periódicos e não periódicos.

Quadro 3.1 – Artigos aplicados ao problema de simulação/previsão de vazões/ENAs (continuação)

Título	Autor(es)	Ano	Comentários
Modeling of aggregated hydrologic time series	Obeyskera J. T. B., Salas, J. D.	1986	Modelos PARMA(p,q) e PAR(p) – discutem as condições de estacionariedade e invertibilidade dos modelos PAR(p) e PARMA(p,q).
Modelos de Memória Longa para a geração de cenários hidrológicos sintéticos	Pereira, G. A. A.	2011	Modelo SARFIMA (p,d,q) (P,D,Q)_s – dissertação aplicada à realidade do SEB que busca modelar processos <i>long memory</i> nas séries de ENAs.
Optimal operation of hydrothermal systems with Hydrological Scenario Generation through Bootstrap and Periodic Autoregressive Models	Souza <i>et. al.</i>	2012	Modelo PAR(p) – os autores propõem a indentificação da ordem e a geração dos cenários via Bootstrap e utilizam tais séries na Programação Dinâmica Estocástica (PDE).
Periodic Models	Hipel, K. W; McLeod, A. I.	1994	Modelo PAR(p) – excelente discussão e definição dos modelos PAR(p).
Geração de cenários sintéticos de energia e vazão para o planejamento da operação energética	Maceira <i>et. al.</i>	2005	Modelo PAR(p) – artigo discute a modelagem PAR(p) no contexto do SEB.
Gamma-Autoregressive Models for Stream-Flow Simulation	Fernandez, B., Salas, J.D.	1990	Modelo Gama – aplicam o modelo GAR(1) à modelagem de séries de energia anuais.
A new autoregressive time series model in exponential variables (NEAR(1))	Lawrance, A. J. Lewis, P. A. W	1981	Modelo Gama – apresentaram o Novo modelo Exponencial Autorregressivo (NEAR(1)) que serviu de base para o modelo GAR(1).
Gamma-Autoregressive Models for Stream-Flow Simulation	Fernandez, B., Salas, J.D.	1986	Modelo Gama – os autores apresentam os modelos PAGAR(1), PMGAR(1), PGAR(1) e comparam os resultados do modelo PGAR(1) com o de outros modelos periódicos.
Stochastic modeling of periodic streamflow series with gamma distributions	Fernandez, B.	1984	Modelo Gama – trabalho que apresenta formalmente os modelos adotados nessa tese.

O primeiro passo da metodologia consiste em identificar, caso seja necessário, a ordem de homogeneidade “d”, ou seja, trata-se de identificar o número de vezes que a série original deve ser diferenciada para se tornar uma série estacionária. Este procedimento pode ser feito através da observação do próprio gráfico da série, da função de autocorrelação, ou, mais usual e confiável, através dos testes de Raiz Unitária¹⁹.

O próximo passo é a identificação do modelo (Figura 3.1). Para isso são utilizados: (i) conceitos de função de autocorrelação (FAC) e autocorrelação parcial (FACP), onde, para a identificação da ordem observa-se o comportamento de ambas; (ii) critérios de informação, como por exemplo, *Akaike information criterion* (Akaike, 1987), *Schwarz criterion* (Schwarz, 1978) entre outros.

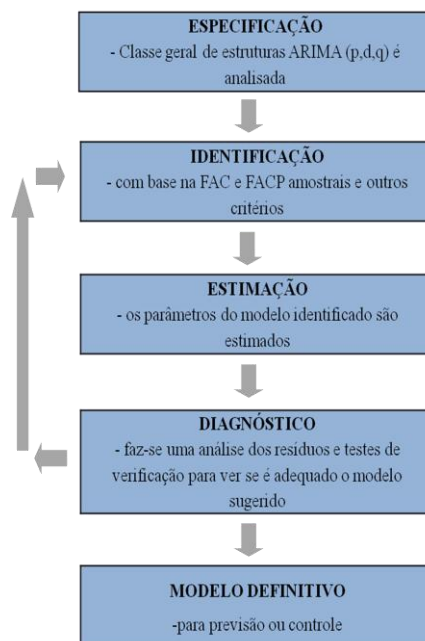


Figura 3.1 – Estágios da modelagem Box & Jenkins

Fonte: (Hamilton, 1994)

De maneira geral, para se identificar a ordem p de um modelo AR (p), por exemplo, observa-se se a FAC decresce exponencialmente e se a FACP apresenta um corte. Se isto acontece, o *lag* onde este corte ocorre fornece a ordem p ($p=\text{lag}$ do corte).

¹⁹ Uma descrição detalhada destes testes pode ser encontrada em (Hamilton, 1994).

Por outro lado, para um modelo MA (q), a FAC e a FACP apresentam comportamento inverso ao de um modelo puramente autorregressivo. Isto significa que, para um modelo MA, a FACP decresce exponencialmente, e a FAC é que apresenta um corte. Da mesma forma, o *lag* onde este corte ocorre fornece a ordem q do modelo MA.

Quanto aos critérios de informação, o modelo mais adequado geralmente é aquele que minimiza o valor do critério de informação observado.

Após a identificação da ordem do modelo, é necessário obter as estimativas dos parâmetros desse modelo (Figura 3.1). A técnica utilizada para as estimativas é a da máxima verossimilhança, sendo que para modelos “AR puros” o método dos momentos aproxima-se significativamente do método de máxima verossimilhança.

Por fim, identificado o modelo e estimado os parâmetros, fazem-se os testes de aderência para verificar a adequabilidade final do modelo: testes para os resíduos e os testes de sobrefixação.

Nos testes para os resíduos, procura-se constatar se, depois de ajustado o modelo, o resíduo gerado por este modelo é um resíduo branco, ou seja, se o modelo foi capaz de explicar satisfatoriamente o comportamento da série de forma que o erro não apresente nenhuma estrutura de correlação linear. Este fato consiste num dos indicadores de eficiência explicativa do modelo.

O teste de sobrefixação, por sua vez, consiste simplesmente em se gerar modelos de ordens superiores à identificada, de forma que se possa reforçar a adequabilidade do modelo estimado.

3.3

O Modelo PAR(p)

Seja uma série temporal com n anos e s períodos sazonais em cada ano (no caso deste trabalho $s=12$ meses) e $Y_{r,m}$ uma realização desse processo no ano r e no mês m ($r = 1, 2, \dots, n$; $m = 1, 2, \dots, s$).

O modelo autorregressivo periódico, PAR(p), onde p é a ordem do modelo ($p = [p_1, p_2, \dots, p_s]$), é, de maneira geral, obtido a partir da definição de um modelo AR para cada período sazonal. Matematicamente tem-se:

$$Y_{m,r} - \mu_m = \sum_{i=1}^{p_m} \phi_i^m (Y_{m-i,r} - \mu_{m-i}) + a_{m,r} \quad (3.1)$$

μ_m Média sazonal do período m .

ϕ_i^m i -ésimo coeficiente autorregressivo do período m .

$a_{m,r}$ Série de ruídos independentes com média zero e variância $\sigma_a^{2(m)}$.

Seja L o operador de defasagem, tal que $LY_{m,r} = Y_{m-1,r}$. Tem-se $L^k Y_{m,r} = Y_{m-k,r}$. Assim, o modelo descrito em (3.1) pode ser reescrito como:

$$\phi^m(L)(Y_{m,r} - \mu_m) = a_{m,r} \quad (3.2a)$$

onde:

$$\phi^m(L) = 1 - \phi_1^m L - \phi_2^m L^2 - \dots - \phi_{p_m}^m L^{p_m} \quad (3.2b)$$

é o operador AR de ordem p_m para o mês m .

Quanto a estimação dos parâmetros do modelo PAR(p), têm-se que a função de autocovariância de lag k para o mês m é definida como:

$$\gamma_k^m = E[(Y_{m,r} - \mu_m)(Y_{m-k,r} - \mu_{m-k})] \quad (3.3)$$

Multiplicando-se a (equação 3.1) por $(Y_{m-k,r} - \mu_{m-k})$ e tirando o seu valor esperado, tem-se:

$$E[(Y_{m,r} - \mu_m)(Y_{m-k,r} - \mu_{m-k})] = \sum_{i=1}^{p_m} \phi_i^m E[(Y_{m-i,r} - \mu_{m-i})(Y_{m-k,r} - \mu_{m-k})] + E[a_{m,r}(Y_{m-k,r} - \mu_{m-k})] \quad (3.4)$$

que pode ser reescrita como:

$$\gamma_k^m = \sum_{i=1}^{p_m} \phi_i^m \gamma_{k-i}^m \quad (3.5)$$

pois $E[a_{m,r}(Y_{m-k,r} - \mu_{m-k})] = 0$ para $k > 0$ e $n=1,2,\dots,s$

Fixando-se m e variando-se $k = 1,2,\dots, p_m$ em (3.5), obtém-se o conjunto de (equações 3.6), denominadas Equações de Yule-Walker.

$$\begin{cases} \gamma_1^m = \phi_1^m \gamma_0^{m-1} + \phi_2^m \gamma_{-1}^{m-2} + \phi_3^m \gamma_{-2}^{m-3} + \dots + \phi_{p_m}^m \gamma_{-(p_m-1)}^{m-p_m} \\ \gamma_2^m = \phi_1^m \gamma_{-1}^{m-1} + \phi_2^m \gamma_0^{m-2} + \phi_3^m \gamma_{-1}^{m-3} + \dots + \phi_{p_m}^m \gamma_{-(p_m-2)}^{m-p_m} \\ \vdots \\ \gamma_{p_m}^m = \phi_1^m \gamma_{-(p_m-1)}^{m-1} + \phi_2^m \gamma_{-(p_m-2)}^{m-2} + \phi_3^m \gamma_{-(p_m-3)}^{m-3} + \dots + \phi_{p_m}^m \gamma_0^{m-p_m} \end{cases} \quad (3.6)$$

Para possibilitar que o (sistema 3.6) possa ser reescrito de forma mais simples, evitando a representação das autocovariâncias com *lags* negativos, pode-se usar a relação $\gamma_{-j}^{m-k} = \gamma_{+j}^{m-k+j}$. As incógnitas deste sistema são os parâmetros $\phi_1^m, \phi_2^m, \dots, \phi_{p_m}^m$ do modelo PAR(p) para o mês m . O sistema apresentado em (3.6), chamado de sistema Yule-Waker Periódico, pode ser reescrito como:

$$\begin{bmatrix} \phi_1^m \\ \phi_2^m \\ \vdots \\ \phi_{p_m}^m \end{bmatrix} = \begin{bmatrix} \gamma_0^{m-1} & \gamma_1^{m-1} & \gamma_2^{m-1} & \dots & \gamma_{p_m-1}^{m-1} \\ \gamma_1^{m-1} & \gamma_0^{m-2} & \gamma_1^{m-2} & \dots & \gamma_{p_m-2}^{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{p_m-1}^{m-1} & \gamma_{p_m-2}^{m-2} & \gamma_{p_m-3}^{m-3} & \dots & \gamma_0^{m-p_m} \end{bmatrix}^{-1} \begin{bmatrix} \gamma_1^m \\ \gamma_2^m \\ \vdots \\ \gamma_{p_m}^m \end{bmatrix} \quad (3.7)$$

Objetivando escrever o sistema de Yule-Walker em função das

$$\frac{Y_{r,m} - \mu_m}{\sigma_m} = \sum_{i=1}^{p_m} \phi_i^m \left(\frac{Y_{m-i,r} - \mu_{m-i}}{\sigma_{m-i}} \right) + a_{m,r} \quad (3.8)$$

Daí, a função de autocorrelação periódica é definida para o mês m e *lag* k como:

$$\rho_k^m = \frac{\gamma_k^m}{\sqrt{\gamma_0^m \gamma_0^{m-k}}} \quad (3.9)$$

Multiplicando-se a (equação 3.9) por $\frac{(Y_{m-k,r} - \mu_{m-k})}{\sigma_{m-k}}$ e tirando o seu valor esperado, tem-se:

$$E \left[\frac{(Y_{m,r} - \mu_m)(Y_{m-k,r} - \mu_{m-k})}{\sigma_m \sigma_{m-k}} \right] = \sum_{i=1}^{p_m} \phi_i^m E \left[\frac{(Y_{m-i,r} - \mu_{m-i})(Y_{m-k,r} - \mu_{m-k})}{\sigma_{m-i} \sigma_{m-k}} \right] \quad (3.10)$$

Dado que $\gamma_0^m = \sigma_m^2$ e $\gamma_0^{m-k} = \sigma_{m-k}^2$ substituindo na (equação 3.10) tem-se:

$$\rho_k^m = \sum_{i=1}^{p_m} \phi_i^m \rho_{k-i}^{m-i} \quad (3.11)$$

Valendo-se de que a relação $\rho_{-j}^{m-k} = \rho_j^{m-k+j}$ é verdadeira, pode-se reescrever a (equação 3.11), fixando-se m e variando-se os *lags* na forma matricial de Yule-Walker para as correlações.

$$\begin{bmatrix} \phi_1^m \\ \phi_2^m \\ \vdots \\ \phi_{p_m}^m \end{bmatrix} = \begin{bmatrix} \rho_0^{m-1} & \rho_1^{m-1} & \rho_2^{m-1} & \cdots & \rho_{p_m-1}^{m-1} \\ \rho_1^{m-1} & \rho_0^{m-2} & \rho_1^{m-2} & \cdots & \rho_{p_m-2}^{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p_m-1}^{m-1} & \rho_{p_m-2}^{m-2} & \rho_{p_m-3}^{m-3} & \cdots & \rho_0^{m-p_m} \end{bmatrix}^{-1} \begin{bmatrix} \rho_1^m \\ \rho_2^m \\ \vdots \\ \rho_{p_m}^m \end{bmatrix} \quad (3.12)$$

Caso a ordem dos modelos fosse previamente conhecida, os parâmetros poderiam ser obtidos resolvendo o sistema de (equação 3.12). Contudo, como deseja-se descobrir a ordem do modelo, a (equação 3.12) pode ser resolvida sucessivamente para diferentes valores de p_m , determinando-se os parâmetros $\phi_{p_m,1}^m \dots \phi_{p_m,p_m}^m$ para cada modelo de ordem p_m .

Na prática, há diversas maneiras de identificar a ordem dos modelos Box & Jenkins. Uma alternativa, seria aumentar o valor de p_m de 1 até um $(k+1)$ valor no qual o parâmetro $\phi_{k+1,k+1}^m$ não seja mais significativo. Neste caso, o último valor de p_m para o qual o parâmetro ϕ_{p_m,p_m}^m foi significativo será a ordem do modelo ($p_m=k$). Os parâmetros do modelo serão $\phi_{k,1}^m \dots \phi_{k,k}^m$.

Neste ponto, é importante ressaltar alguns pontos com relação à identificação da ordem e estimação dos parâmetros do modelo. Com relação à identificação da ordem, usualmente considera-se o intervalo de confiança de 95% para $\phi_{k,k}^m$ dado por $1,96/\sqrt{N}$, onde N é o número de anos da série. Tal intervalo é uma aproximação elaborada por (Bartlett, 1946) e foi abordada por (Neto & Souza, 1996), o qual definiu um processo quase ruído branco para processos identificados por essa aproximação. Em 2011, (Oliveira & Souza, 2011) levantaram esse problema para o caso de modelos periódicos autorregressivos.

Ainda, além do problema da aproximação de (Bartlett, 1946), não existe um consenso quanto à maneira de identificar-se a ordem do modelo PAR(p). Em outras palavras, além da possibilidade descrita anteriormente, a qual, neste

trabalho será conhecida como identificação da esquerda para direita (E-D), ressaltam-se pelo menos mais três maneiras de executar tal tarefa²⁰.

Na segunda maneira, identificação da direita para a esquerda (D-E), ao contrário do que acontece na identificação E-D, a ordem do modelo é definida pelo valor de k^* tal que para todo $k > k^*$ ϕ_{kk}^m não é significativo. Exemplificando: se a FACP indica que o *lag* 6 é significativo e para todo *lag* maior que 6 não é, a ordem escolhida será 6, independentemente da ocorrência de valores ϕ_{kk}^m não significantes para k menores do que 6, digamos 4 e 5. Tal critério é importante porque essa é a forma com que o Setor Elétrico Brasileiro identifica a ordem dos modelos.

Uma terceira maneira para identificar a ordem dos modelos $PAR(p)$, foi abordada por (Oliveira, 2010). Nesta, o autor utiliza técnicas de reamostragem Bootstrap para definir a ordem do modelo. Observou-se que a técnica proposta aproximou-se bastante do critério de identificação E-D e propiciou bons resultados quanto à geração de cenários.

Outra possibilidade muito abordada na literatura é definir a ordem dos modelos minimizando o módulo dos Critérios de Informação, como por exemplo, os critérios (Akaike, 1987) e (Schwarz, 1978). A vantagem desse método é o equilíbrio entre o ajuste do modelo e a *Lex Parsimoniae*. Particularizando para os modelos de séries temporais, com esse critério, parâmetros estimados somente serão utilizados caso contribuam realmente para o ajuste da modelagem à série de dados. De acordo com a literatura mais recente de séries temporais, este é o método mais adequado para identificação de modelos da família ARIMA, como destacado em (Lütkepohl, 1991). Em (Lund, et al., 2006) os autores mostram a importância desse tipo de identificação para modelos periódicos.

Por fim, com relação à estimação dos parâmetros, é importante ressaltar que a técnica apresentada nesse trabalho, que é a mesma utilizada pelo Setor Elétrico, é a estimação via Método dos Momentos, isto é, igualam-se os momentos amostrais aos momentos populacionais e definem-se os parâmetros do modelo. No caso da metodologia Box & Jenkins, conforme observado anteriormente, tal processo é feito por meio das equações de Yule-Walker.

²⁰ Vale lembrar que existem mais maneiras de identificar a ordem de modelos da família Box & Jenkins, como por exemplo, através da FACE, utilizando técnicas de inteligência artificial etc.

Segundo (Hipel & McLeod, 1994) apesar da estimação via Máxima Verossimilhança apresentar resultados estatísticos mais robustos com relação aos parâmetros estimados, para modelos autorregressivos que não apresentam a parte MA (*Moving Average*), como é o caso dos modelos PAR(p), a estimação via Método dos Momentos é uma boa aproximação deste tipo de estimação.

No próximo capítulo discute-se o Modelo de Geração de Energias Afluentes (MGEA) no contexto do SEB, destacando suas peculiaridades, sua relação com a PDDE e um estudo de caso, respeitando as premissas adotadas pelo Setor Elétrico Brasileiro.

4

O modelo de geração de Energias Afluentes (MGEA) no Contexto do SEB

Conforme observado neste trabalho, o módulo NEWAVE é dividido em quatro partes: (i) Módulo de cálculo do sistema equivalente, (ii) Módulo de Energias Afluentes, (iii) Módulo de cálculo da política de operação hidrotérmica e (iv) Módulo de simulação da operação.

Uma vez que o objetivo central dessa tese é propor uma nova abordagem para a geração de séries sintéticas, nessa seção discutem-se pontos relacionados ao módulo de Energias Afluentes implementado no modelo NEWAVE. Tal investida faz-se relevante, pois conhecer as peculiaridades adotadas pelo setor, levantar possíveis entraves à geração de bons cenários hidrológicos e limitações relativas à fase de otimização (PDDE) feita no módulo de cálculo da política de operação hidrotérmica é essencial para uma boa comparação entre o método que está em vigor e a nova proposta.

Dentro deste contexto, este capítulo procura responder qual é a utilidade do MGEA, como o mesmo está ligado à PDDE e quais são as “ressalvas” que essa ligação estabelece. Ainda, objetiva-se levantar algumas peculiaridades do módulo, como por exemplo, a questão das configurações intercorrelacionadas e pontos que levantam discussão com relação à teoria de Séries Temporais, como por exemplo, a definição da ordem do modelo, abordada no capítulo 3.

Esse capítulo é apresentado da seguinte forma: em primeiro lugar, apresentam-se as particularidades do MGEA no contexto do SEB, como sua ligação com a PDDE, a montagem das configurações do parque gerador, o modelo “PAR(p) Interconfigurações” (PAR(p)-IC), as equações do modelo PAR(p) e a PDDE, a geração de séries sintéticas e a questão das afluições negativas e a estrutura não linear do PAR (p). Num segundo momento, expõe-se um estudo de caso adotando os critérios definidos pelo SEB, onde serão levantados pontos relativos à teoria de Séries Temporais e resultados que servirão de base de comparação à abordagem Gama proposta.

4.1

Peculiaridades do MGEA

A adoção de critérios probabilísticos em diversas atividades do planejamento e operação de sistemas hidrotérmicos criou a necessidade da modelagem probabilística de afluências. Em estudos energéticos, critérios de suprimentos são baseados em índices de risco, estimados a partir da simulação da operação energética para diversos cenários (sequências) de afluências aos aproveitamentos hidroelétricos ou subsistemas (CEPEL, 2011b).

Segundo (Kelman & Pereira, 1977) a série histórica de afluências, o único cenário disponível, é insuficiente para compor uma amostra de tamanho necessário para estimar índices de riscos com incertezas aceitáveis. Para solucionar tal problema, estimam-se modelos baseados nas séries históricas e geram-se cenários igualmente prováveis capazes de reproduzir as características das séries históricas. Assim, a informação contida na série histórica pode ser extraída de maneira mais completa, permitindo a avaliação de riscos e incertezas pertinentes a um sistema hidrotérmico.

Neste contexto, um modelo de geração de energias afluentes (“Modelo PAR(p) Interconfigurações”) foi adotado pelo SEB com o objetivo de gerar séries sintéticas para os diferentes módulos que auxiliam na operação do sistema hidrotérmico brasileiro. O foco principal deste trabalho são as séries sintéticas de energia utilizadas na fase de otimização (PDDE) para o planejamento de médio prazo, e, conforme já observado, tal modelo está inserido no âmbito do módulo NEWAVE.

Para fornecer séries de energia à fase de otimização do módulo NEWAVE, o MGEA é dividido em três partes principais: (i) transformar as séries de vazões em Energias Naturais Afluentes²¹, agregá-las por subsistema e “montar” as séries históricas de energias correspondentes às configurações²² do parque gerador; (ii) ajuste do modelo “PAR(p) Interconfigurações”; e, (iii) geração das séries sintéticas.

²¹ Para maiores detalhes ver: (Terry, et al., 1986); (CEPEL, 2011b).

²² Em linhas gerais, a configuração do parque gerador é determinada a partir de uma composição fixa de UHEs, UTEs, LTs etc, em um dado instante. Para esta configuração, a série de ENA é calculada a partir da série de vazões históricas disponível (e.g. jan/1931 a dez/2011).

A primeira parte, “montagem” das séries históricas de energias correspondentes às configurações do parque gerador, apesar de não apresentar conceitos estatísticos ou de séries temporais, é fundamental para o entendimento das análises subsequentes, uma vez que a abordagem do modelo PAR(p) feita pelo SEB diferencia-se da versão *mainstream* descrita na literatura ((Thomas & Fiering, 1962); (Hipel & McLeod, 1994)) e abordada no capítulo 3 desse trabalho.

Ao abordar a parte (ii) do modelo de geração de energias afluentes discutem-se a ideia do modelo “PAR(p) Interconfigurações” e pontos relativos à teoria Box & Jenkins e a PDDE, por fim, discutem-se algumas questões relacionadas à geração de cenários e o comportamento dos mesmos assumindo ruídos lognormal.

4.1.1

A montagem das configurações do Parque Gerador

Conforme foi discutido anteriormente, embora o objetivo desse trabalho não seja construir as configurações topológicas feitas pelo SEB, essa seção serve de preâmbulo para o entendimento dos demais pontos que serão abordados nessa tese. Portanto, para compreender o processo de modelagem do “PAR(p) Interconfigurações”, é preciso ter claro quais são os impactos das mudanças de configurações topológicas na modelagem do problema.

A Câmara de Comercialização de Energia Elétrica disponibiliza em seu site mensalmente o Programa Mensal de Operação (PMO)²³ elaborado pelo ONS, no qual é possível observar quais usinas entrarão em operação nos próximos anos, qual será a sua capacidade de geração e produtividade. Durante a fase de entrada em operação de uma usina, são observadas três etapas: período de enchimento de volume morto, período de submotorização e período de completa motorização. No período de enchimento de volume morto, somente o impacto da entrada em operação do reservatório da usina é levado em consideração no processo de otimização, pois uma parcela de afluência à usina hidrelétrica é direcionado para o enchimento do reservatório e não existe geração de energia elétrica.

²³ Disponível em: <http://www.ccee.org.br>

No período de submotorização, o volume morto do reservatório já se encontra cheio e a usina gera energia para o SIN, mas como o número de máquinas em operação não é suficiente para a usina gerar a sua energia assegurada, diz-se que a usina ainda não atingiu sua potência de base e portanto está submotorizada. No período de submotorização, a usina ainda não é considerada no processo de otimização e sua geração média projetada é deduzida do mercado.

A partir do mês que a usina começa a operar com número de máquinas suficiente para a geração de sua energia assegurada, diz-se que a usina encontra-se motorizada e esta passa a ser considerada no processo de otimização. Portanto, a cada mês em que novas unidades geradoras hidrelétricas se tornam operacionais, tem-se uma nova configuração hidrelétrica do sistema, isto é, a série histórica de ENA para o período de jan/1931 a dez/2011, por exemplo, sofrerá um incremento caso uma nova usina entre em operação em um determinado mês/ano no futuro. Logo, durante o processo de entrada de uma nova usina, a série histórica de ENA sofre alterações, pois a configuração topológica do sistema é alterada.

Como observado em (Terry, et al., 1986) e (CEPEL, 2011b), as ENAs são basicamente uma função das vazões e da capacidade de geração das usinas. Uma vez que existe o histórico de vazões e tem-se o conhecimento sobre capacidade de geração do conjunto de usinas mês a mês (usinas já instaladas e usinas que irão entrar em operação), o MGEA calcula os históricos de ENAs para cada uma das configurações.

Resumindo, no contexto do módulo MGEA, configuração é uma série histórica de energia correspondente a um dado período de tempo fixo (por exemplo, jan/1931 à dez/2011). Caso uma nova usina entre em operação são acrescentadas energias correspondentes à nova usina para todo o período de tempo, dando origem a uma nova configuração do parque gerador. A junção de cada configuração forma o conjunto de configurações que, no limite, podem atingir 60 (total de meses do período de planejamento da operação). Dessa forma, o número de configurações vai ser igual ao número de meses em que entrou em operação uma nova usina.

Ainda, neste contexto de configuração, existem dois termos definidos pelo SEB que são o pré-estudo e o pós-estudo. O pré-estudo corresponde, basicamente, a primeira configuração e serve como “ponto inicial” para a estimação dos

modelos, já o pós-estudo é a última configuração do sistema do período de planejamento e os modelos definidos nessa configuração serão replicados, caso deseje-se criar cenários dez/trinta anos à frente, por exemplo para tratar dos efeitos de horizonte infinito.

No contexto do módulo NEWAVE, os períodos de pré-estudo e pós-estudo são períodos nos quais as configurações do parque gerador e os dados de demanda são considerados constantes. Estes períodos servem para eliminar os efeitos do estado inicial (pré-estudo) e para obter uma informação acerca da FCF a partir do período de interesse (FCF diferente de zero). Estes períodos são utilizados tanto no planejamento da expansão quanto no planejamento da operação. A diferença é que para o planejamento da expansão, o que se quer é definir as configurações do parque gerador para todo o período de planejamento, que compreende 10 anos ou mais, enquanto que no planejamento da operação a configuração é supostamente conhecida.

Para entender melhor o que foi dito anteriormente será utilizado um exemplo didático. Observando a (Tabela 4.1) e a (Figura 4.1), as quais indicam a inserção ou não de novas usinas no parque gerador, verifica-se que nos meses de janeiro e fevereiro de 2013 (configuração 1) não há a entrada de nenhuma nova usina em operação, por outro lado, no mês de março (configuração 2) há a entrada de uma nova usina e nos dois meses seguintes, correspondentes às configurações 3 e 4, verifica-se que 5 novas usinas entraram em operação. Por fim, em junho de 2013 (configuração 5) mais 8 novas usinas começaram a operar. Neste exemplo didático, considera-se que quando uma usina entra em operação, ela já passa operar diretamente como uma usina motorizada, ou seja, não estão sendo considerados os períodos de enchimento de volume morto e submotorização.

De posse desses dados, o modelo de geração de energias afluentes irá gerar diferentes séries de ENAs para uma dada série histórica, isto é, admitindo-se que a série histórica de vazões naturais compreenda o período de jan/1931 à dez/2011, no mês de mar/2013 (Tabela 4.1), por exemplo, será construída uma série histórica de ENAs (jan/1931 à dez/2011) correspondente à configuração de mar/2013. É importante lembrar que, mesmo que a usina só entre em operação em mar/2013, sua contribuição hídrica será acrescida em todos os meses daquela configuração, inclusive janeiro e fevereiro, por exemplo.

Tabela 4.1 – Exemplo de Novas Configurações

	Jan/2013	Fev/2013	Mar/2013	Abr/2013	Mai/2013	Jun/2013
Novas Usinas	0	0	1	5	5	8
Configuração	1	1	2	3	4	5

Em outras palavras, a partir da “série histórica inicial”, a qual denominou-se pré-estudo, constrói-se novas séries históricas para o período de jan/1931 à dez/2011, incorporando nas mesmas a entrada de novas usinas (Figura 4.1) como ocorreu no mês de mar/2013, por exemplo.

Observando a (Figura 4.1), verifica-se que as configurações são as mesmas para o Pré-Estudo e para os meses de jan/2013 e fev/2013. Já em mar/2013, como haverá a entrada de uma nova usina, a ENA calculada para esse subsistema será a ENA da configuração 1 (fev/2012) mais a variação na ENA provocada por essa nova usina ($ENA(\text{conf } 2) = ENA(\text{conf } 1) + \Delta ENA_{NOVA \text{ USINA}}$), e assim, sucessivamente.

Observe que, neste exemplo, durante o período de planejamento da operação (jan/2013 à dez/2016), ocorreram 34 mudanças de configuração, isto é, neste período, novas usinas entraram em operação em 34 diferentes meses, conforme verificado na (Figura 4.1).

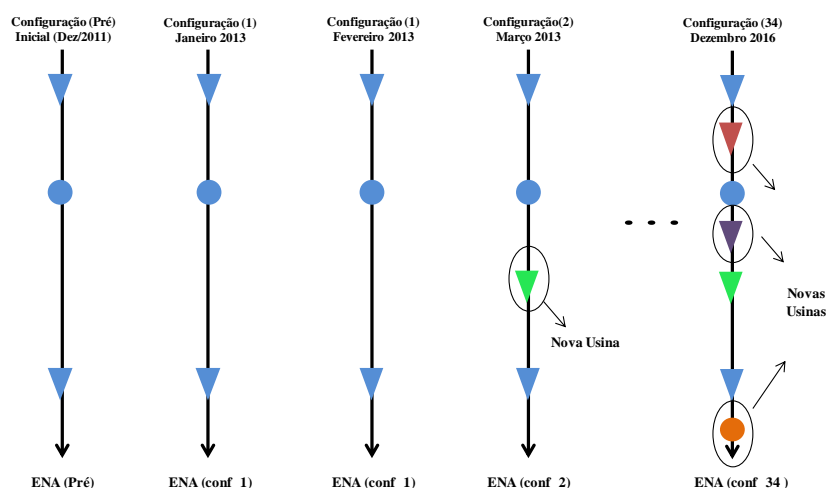


Figura 4.1 – Usinas instaladas e que entraram em operação de acordo com cada configuração

Apesar da lógica apresentada acima ser a utilizada pelo SEB, para fins didáticos, a base de dados que será utilizada para estimar o modelo “PAR(p)

Interconfigurações” ficará conforme apresentada na (Figura 4.2), isto é, assumindo o exemplo citado acima em que se tem o PMO de janeiro e que se queira gerar cenários sintéticos 5 anos à frente, trabalhar-se-á com 60 diferentes configurações²⁴ do parque gerador (obviamente, caso não ocorra o incremento novas usinas pode haver configurações iguais), mais as configurações de pré-estudo e pós-estudo (Figura 4.2).

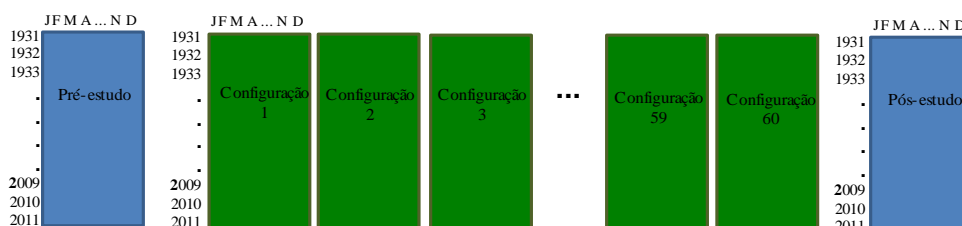


Figura 4.2 – Pré-estudo, Pós-estudo e as diferentes configurações de energia

Como observado, para estudos de planejamento da operação, o período de interesse são os cinco primeiros anos. No contexto do NEWAVE, adicionalmente, em geral consideram-se cinco anos adicionais, perfazendo um total de 10 anos de cenários gerados, para levar em conta os efeitos de final de horizonte, isto é, o sistema deve continuar operando após os cinco anos iniciais. Para o planejamento da expansão, em geral o horizonte é maior, 10 anos de interesse, podendo chegar a 30. De qualquer forma, em ambos os estudos, operação e planejamento, sempre se consideram cinco anos adicionais, destinados a construir uma função de custo futuro para a fronteira do horizonte de interesse.

Outra peculiaridade da geração de séries sintéticas que deve ser ressaltada nesse tópico é a questão “tendência hidrológica”. Os dados disponibilizados pelo ONS para a composição do histórico de vazões às usinas hidrelétricas, o qual será utilizado para o cálculo da série de ENAs normalmente inicia-se em janeiro de 1931 e termina em dezembro do ano igual a dois anos anteriores aos dados do PMO. Exemplo, para o PMO de outubro de 2013, o histórico de vazões fornecido pelo ONS inicia-se em janeiro de 1931 e termina em dezembro de 2011.

²⁴ Costuma-se referenciar as configurações com o nome do mês e ano em que a mesma corresponde, isto é, caso esteja-se trabalhando com o PMO de janeiro de 2012, por exemplo, referencia-se a configuração 1 como configuração de jan/2012, a configuração 2 como configuração de fev/2012 e assim sucessivamente até chegar a última configuração que, neste caso, seria a configuração de dez/2017 (configuração 60).

Como a estrutura do modelo $PAR(p)$, dada por (4.4), requer a utilização de até onze períodos anteriores, no processo de geração de séries sintética são oferecidas duas opções ao usuário dos modelos: (a) Consideração da tendência hidrológica (b) Sem consideração da tendência hidrológica. Na opção de consideração da tendência hidrológica, o ONS fornece as ENAs calculadas para os onze meses anteriores ao início, por exemplo, para o PMO de outubro de 2013, o ONS fornecerá um arquivo contendo as ENAs de dezembro de 2012 a setembro de 2013. Na opção sem consideração da tendência hidrológica, o modelo NEWAVE irá considerar que as ENAs dos onze meses anteriores ao início do estudo é igual a MLT (média de longo termo) de cada mês, ou irá obter tantas séries de ENAs passadas quantas serão as séries utilizadas no passo *Foward* (200, atualmente).

4.1.2

Modelo “ $PAR(p)$ Interconfigurações”

Conforme verificado no capítulo 3, há diversas aplicações do modelo $PAR(p)$ para diferentes tipos de problemas envolvendo questões hidrológicas. Ainda, neste mesmo capítulo verificou-se a abordagem *mainstream* do modelo $PAR(p)$ apresentada por (Hipel & McLeod, 1994), na qual, de posse de uma série temporal (*e.g.* jan/1931 a dez/2010) calculam-se os momentos amostrais, a Função de Autocorrelação, a Função de Autocorrelação Parcial, define-se a ordem dos modelos, estimam-se os parâmetros e faz-se a previsão e/ou simulação.

Entretanto, como será demonstrado a seguir, o Setor Elétrico Brasileiro não trabalha dessa forma. Para estimar o modelo $PAR(p)$, o qual foi denominado nessa tese como Modelo “ $PAR(p)$ Interconfigurações”, o setor não considera uma série histórica de dados “constante” no tempo, isto é, dadas as diferentes configurações de séries temporais (seção 4.1.1), calculam-se os momentos amostrais, a FAC, a FACP, define-se a ordem dos modelos, estimam-se os parâmetros e faz-se a previsão e/ou simulação.

Nessa seção apresenta-se o $PAR(p)$ em sua forma intercorrelacionada, ou seja, no formato onde é possível incorporar o efeito da entrada de novas usinas no parque gerador nacional. Como poderá ser observado, a formulação matemática é

muito similar a apresentada na (seção 3.3) dessa tese, diferenciando-se apenas no “cruzamento” das configurações.

O modelo $PAR(p)$ é empregado nas séries de ENAs para que seja possível gerar séries sintéticas e parâmetros para a PDDE. Como dito anteriormente, é necessário também que seja possível incorporar os efeitos de futuras usinas que entrarão em operação no horizonte de estudo. Para que isto seja possível, o $PAR(p)$ deverá ser estimado com base em diferentes ENAs simultaneamente. Em outras palavras, haverá um “cruzamento” entre as configurações do parque gerador.

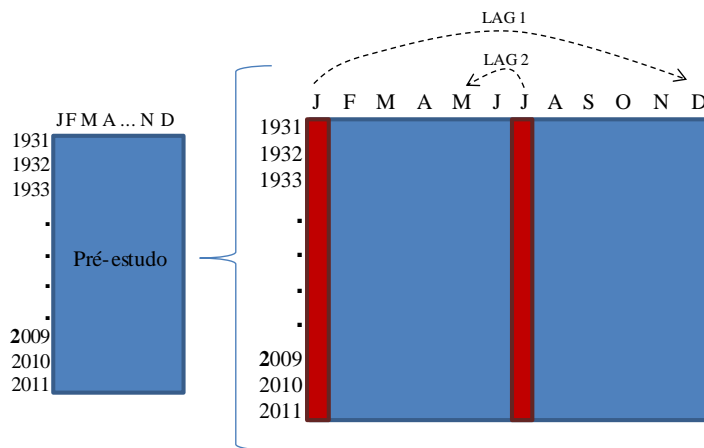
Dado o ineditismo dessa demonstração, avaliou-se que antes de partir para a formulação do modelo, torna-se útil seguir o exemplo apresentado na seção anterior aplicado à estimação do modelo “ $PAR(p)$ -IC”. Recapitulando, assume-se que se deseja estimar modelos para a geração de cenários 5 anos à frente, cujo PMO é de janeiro de 2012 e que a série de ENAs compreende o período de jan/1931 à dez/2011, dessa forma, as configurações do parque gerador são as mesmas apresentadas na (Figura 4.2) deste trabalho.

Admitindo que o leitor tenha um conhecimento prévio da teoria Box & Jenkins e dos modelos $PAR(p)$ e seguindo as fases da modelagem Box & Jenkins apresentadas na (Figura 3.1), este exemplo inicia-se abordando a estimação da “função de autocorrelação interconfigurações (FAC-IC)”.

A FAC-IC será estimada de acordo com três diferentes algoritmos: (i) FAC-IC seguindo o modelo $PAR(p)$ *mainstream*, onde são considerados os dados da configuração do pré-estudo apenas; (ii) FAC-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo; e (iii) FAC-IC cruzando apenas com as configurações anteriores.

(i) FAC-IC seguindo o modelo $PAR(p)$ mainstream

Para a configuração do pré-estudo são estimados modelos $PAR(p)$ para cada mês, nesse caso, a FAC-IC é estimada apenas com os dados do pré-estudo, conforme ilustrado na (Figura 4.3), ou seja, o *lag* 1 de janeiro, por exemplo, é a correlação entre o mês de janeiro do pré-estudo com o mês de dezembro, também do pré-estudo. O *lag* 2 será a correlação entre janeiro com novembro e assim sucessivamente.

Figura 4.3 – FAC-IC *mainstream*

Outra peculiaridade adotada pelo SEB que merece destaque é a correlação entre meses antecedentes em anos coincidentes, como ocorre entre janeiro e dezembro de 1931, por exemplo. Na abordagem tradicional, excluir-se-ia o ano de 1931, faria as correlações entre jan/1932 e dez/1931, jan/1933 e dez/1932 e assim sucessivamente, perdendo-se as observações jan/1931 e dez/2011, neste exemplo.

No caso do SEB, para não perder a correlação com a última observação, adotou-se o seguinte procedimento: depois do último janeiro, neste caso, jan/2011, faz-se uma média entre todos os janeiros, adiciona-se esse elemento na série e o correlaciona com o último dezembro, em outras palavras, é como se fosse criado um valor jan/2012.

(ii) *FAC-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo*

Esse segundo procedimento para o cálculo da FAC-IC correlaciona os dados das configurações com elas mesmas e com os dados do pré-estudo. Nesse procedimento, ao contrário do que acontece em (i), olha-se apenas para a esquerda para calcular as correlações, isto é, por exemplo, o *lag* 1 de fevereiro da configuração 1, será a correlação entre fevereiro da configuração 1 e janeiro do pré-estudo e o *lag* 2 será a correlação entre fevereiro e dezembro, da configuração 1 e do pré-estudo, respectivamente.

Observe agora as correlações de fevereiro da configuração 2, a correlação de lag 1 será dada pela relação entre fevereiro da configuração 2 e janeiro da configuração 1, já o *lag* 2 é dado pela relação entre fevereiro da configuração 2 e

dezembro do pré-estudo (Figura 4.4). O *lag* 3 será a correlação entre fevereiro da configuração 2 e novembro do pré-estudo e assim, sucessivamente. Verifique que, como não há configurações anteriores ao pré-estudo, as correlações superiores à segunda defasagem sempre remontarão ao pré-estudo, o que não irá acontecer quando houver mais configurações, próximo caso.

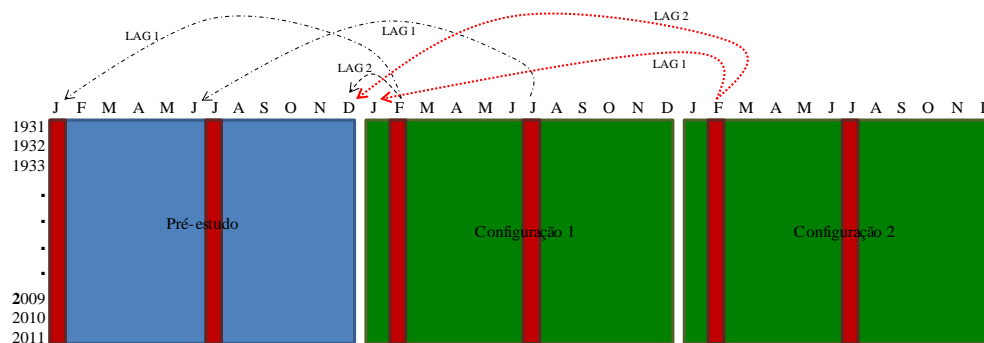


Figura 4.4 – FAC-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo

(iii) FAC-IC cruzando apenas com configurações anteriores

A terceira e última lógica para o cálculo da FAC-IC é apresentado na (Figura 4.5). Conforme pode-se observar as correlações dos mês de fevereiro da configuração 26 são dadas pela relação entre fevereiro (configuração 26) com janeiro (configuração 25), *lag* 1, fevereiro (configuração 26) com dezembro (configuração 24), *lag* 2, fevereiro (configuração 26) com novembro (configuração 23), *lag* 3 e assim sucessivamente.

É importante lembrar que, apesar dos modelos PAR(p) não poderem ter ordem maiores que 6 (valor fixado pela ANEEL), as FAC-IC são calculadas para defasagens superiores a este valor, pois, para o cálculo da FACP-IC tais valores serão necessários.

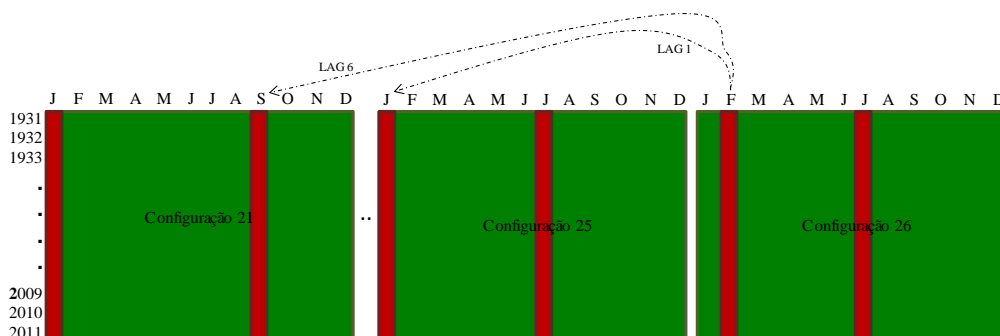


Figura 4.5 – FAC-IC cruzando com as configurações anteriores

Após o cálculo da FAC-IC, calcula-se a função de autocorrelação parcial – interconfigurações (FACP-IC). No caso da FACP-IC, como ocorreu na FAC-IC há também o cruzamento de informações entre as configurações e a lógica para a construção da FACP-IC é parecida com a apresentada anteriormente, a principal diferença é que, para o caso da FACP-IC, são os valores das correlações que importam e não mais os valores das ENAs.

Em outras palavras, como apresentado pela teoria Box & Jenkins, para encontrar as ordens de defasagens ϕ_{kk} , para $k= 1,2, \dots$ é preciso resolver um sistema de equações que tem como variáveis os valores das funções de autocorrelações parciais. No caso da FACP-IC, a ideia é a mesma, a única diferença é que os valores de ρ_k irão variar de acordo com as diferentes configurações.

Seguindo a mesma lógica apresentada para o cálculo da FAC-IC, expõe-se o cálculo da FACP-IC, assumindo $k=2$ e 3.

(i) *FACP-IC seguindo o modelo PAR(p) mainstream*

Para o caso da FACP-IC seguindo o modelo PAR(p) *mainstream*, assim como acontece com a FAC-IC os valores da autocorrelação não variam com as configurações, isto é, estimam-se FACP's estáticas (equação 4.1).

$$\begin{bmatrix} \phi_1^{jan, pré-estudo} \\ \phi_2^{jan, pré-estudo} \end{bmatrix} = \begin{bmatrix} \rho_0^{dez, pré-estudo} & \rho_1^{dez, pré-estudo} \\ \rho_1^{dez, pré-estudo} & \rho_0^{nov, pré-estudo} \end{bmatrix}^{-1} \begin{bmatrix} \rho_1^{jan, pré-estudo} \\ \rho_2^{jan, pré-estudo} \end{bmatrix} \quad (4.1)$$

Para ilustrar, observe a (Figura 4.6) onde pode-se verificar que para calcular a autocorrelação parcial ϕ_{22} utilizam-se apenas correlações estimadas a partir dos dados históricos do pré-estudo.

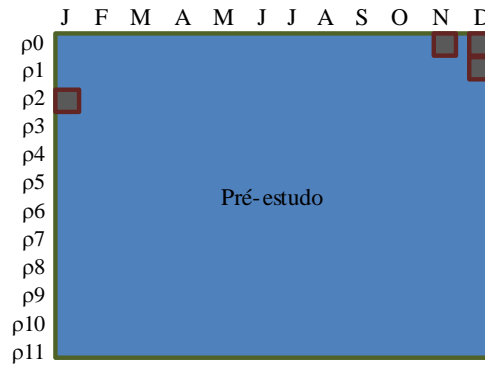


Figura 4.6 – FACP-IC *mainstream*

(ii) FACP-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo

Para o cálculo da FACP-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo, observa-se que utilizam-se dados das correlações dos meses de janeiro (configuração 1), dezembro e novembro, os dois últimos pertencentes ao pré-estudo.

$$\begin{bmatrix} \phi_{31}^{fev,config2} \\ \phi_{32}^{fev,config2} \\ \phi_{33}^{fev,config2} \end{bmatrix} = \begin{bmatrix} \rho_0^{jan,config1} & \rho_1^{jan,config1} & \rho_2^{jan,config1} \\ \rho_1^{jan,config1} & \rho_0^{dez,pré-estudo} & \rho_1^{dez,pré-estudo} \\ \rho_2^{jan,config1} & \rho_1^{dez,pré-estudo} & \rho_0^{nov,pré-estudo} \end{bmatrix}^{-1} \begin{bmatrix} \rho_1^{fev,config2} \\ \rho_2^{fev,config2} \\ \rho_3^{fev,config2} \end{bmatrix} \quad (4.2)$$

Como ilustração, observa-se na (Figura 4.7) como estes dados estariam distribuídos matricialmente onde as linhas são preenchidas valores de ρ_k e as colunas pelos meses.



Figura 4.7 – FACP-IC com o cruzamento dos dados das configurações iniciais e dados do pré-estudo

(iii) FACP-IC cruzando apenas com configurações anteriores

Para calcular a FACP-IC cruzando apenas com configurações anteriores, observa-se para o cálculo do ϕ_{22} do mês de fevereiro (configuração 14) que as correlações necessárias são as correlações de *lag* 0 e 1 do mês de janeiro (configuração 13) e a correlação 0 (variância) do mês de dezembro (equação 4.3).

$$\begin{bmatrix} \phi_1^{fev,config14} \\ \phi_2^{fev,config14} \end{bmatrix} = \begin{bmatrix} \rho_0^{jan,config13} & \rho_1^{jan,config13} \\ \rho_1^{jan,config13} & \rho_0^{dez,config12} \end{bmatrix}^{-1} \begin{bmatrix} \rho_1^{fev,config14} \\ \rho_2^{fev,config14} \end{bmatrix} \quad (4.3)$$

A (Figura 4.8) retrata a disposição dos dados necessários para o cálculo da correlação parcial de ordem 2 no mês de fevereiro da configuração 14.



Figura 4.8 – FACP-IC cruzando com as configurações anteriores

Seguindo a lógica proposta por Box & Jenkins (Figura 3.1), após estimar a FAC-IC e a FACP-IC, defini-se a ordem dos modelos. Conforme discutido na (seção 3.3) desta tese, há diferentes maneiras de identificar a ordem do modelo e, no caso do SEB, a ordem é definida da direita para esquerda (D-E) e não maior que 6 (seis).

No caso da identificação da ordem não há grande diferença entre o PAR(p)-IC e o PAR(p) *mainstream*, contudo, no caso do PAR(p)-IC são estimados tantos modelos quanto o número de meses em que há mudança de configuração. No caso do PAR(p) estático são estimados apenas 12 modelos.

Após identificar a ordem estimam-se os parâmetros do modelo. Conforme observado por (Hipel & McLeod, 1994), para modelos AR o método dos momentos apresenta resultados semelhantes a estimação por máxima verossimilhança, logo os parâmetros estimados são não viesados e consistentes. O

SEB também utiliza o método dos momentos para estimar os parâmetros do modelo PAR(p)-IC.

Por fim, faz-se os diagnósticos do modelo (Figura 3.1), isto é, testa-se pressupostos básicos da teoria Box & Jenkins, tais como, ruídos sem dependência linear e gaussianos. No caso do SEB, não há essa preocupação com a modelagem, o que pode acarretar em modelos e parâmetros não confiáveis.

Quanto à formulação matemática do modelo, para facilitar a notação, assume-se que a configuração estará definida de acordo com o seu mês de referência, por exemplo, se o modelo é autorregressivo de ordem dois referente ao mês de janeiro de 2011, assume-se que tal modelo refere-se à configuração a qual está inserida este mês e que os meses de dezembro e novembro (*lags* 1 e 2) correspondem às suas respectivas configurações. Assim, o modelo PAR(*p*) “Interconfigurações” pode ser descrito matematicamente, de maneira genérica, por:

$$\left(\frac{Y_{r,m,c_m} - \mu_{m,c_m}}{\sigma_{m,c_m}} \right) = \sum_{i=1}^{p_{m,c_m}} \phi_i^{m,c_m} \left(\frac{Y_{r,m-i,c_{m-i}} - \mu_{m-i,c_{m-i}}}{\sigma_{m-i,c_{m-i}}} \right) + a'_{m,c_m} \quad (4.4)$$

onde

Y_{r,m,c_m} é o valor da série temporal no ano r , mês m e configuração c_m , referente ao mês m .

$Y_{m-i,r,c_{m-i}}$ é o valor da série temporal no ano r , mês $(m-i)$ e configuração c_{m-i} , referente ao mês $(m-i)$.

μ_{m,c_m} média sazonal do período m referente à configuração c_m

$\mu_{m-i,c_{m-i}}$ média sazonal do período $(m-i)$ referente à configuração c_{m-i} .

σ_{m,c_m} desvio-padrão sazonal do período m referente à configuração c_m

$\sigma_{m-i,c_{m-i}}$ desvio-padrão sazonal do período $(m-i)$ referente à configuração c_{m-i}

ϕ_i^{m,c_m} i -ésimo coeficiente autorregressivo do período m referente à configuração c_m

p_{m,c_m} ordem do operador autorregressivo do período m referente à configuração c_m

a'_{m,c_m} Série de ruídos do modelo

A metodologia ajusta, portanto, um modelo autorregressivo de ordem p para cada um dos períodos (meses) das séries hidrológicas históricas para cada uma das configurações existentes.

Com relação à identificação da ordem dos modelos e a estimação dos parâmetros, a lógica é a mesma descrita na (seção 3.3), a única diferença é que no caso do PAR(p)-IC, as correlações também são calculadas baseadas nas diferentes configurações, conforme observado no exemplo anterior.

Dessa forma, as formulações apresentadas no capítulo 3 passam a depender das diferentes configurações, assim, pode-se reescrever, de maneira genérica, a (equação 3.12) da seguinte maneira:

$$\begin{bmatrix} \phi_1^{m,c_m} \\ \phi_2^{m,c_m} \\ \vdots \\ \phi_{p_m}^{m,c_m} \end{bmatrix} = \begin{bmatrix} \rho_0^{m-1,c_{m-1}} & \rho_1^{m-1,c_{m-1}} & \rho_2^{m-1,c_{m-1}} & \cdots & \rho_{p_m-1}^{m-1,c_{m-1}} \\ \rho_1^{m-1,c_{m-1}} & \rho_0^{m-2,c_{m-2}} & \rho_1^{m-2,c_{m-2}} & \cdots & \rho_{p_m-2}^{m-2,c_{m-2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p_m-1}^{m-1,c_{m-1}} & \rho_{p_m-2}^{m-2,c_{m-2}} & \rho_{p_m-3}^{m-3,c_{m-3}} & \cdots & \rho_0^{m-p_m,c_{m-p_m}} \end{bmatrix}^{-1} \begin{bmatrix} \rho_1^{m,c_m} \\ \rho_2^{m,c_m} \\ \vdots \\ \rho_{p_m}^{m,c_m} \end{bmatrix} \quad (4.5)$$

Dado a (equação 4.5), estima-se a função de autocorrelação parcial, define-se a ordem dos modelos e os parâmetros que serão usados para gerar as séries sintéticas de energia e para a formulação da PDDE no módulo NEWAVE.

Após apresentar o modelo PAR(p)-IC é importante destacar, em primeiro lugar, o motivo pelo qual o SEB o desenvolveu, e, em segundo lugar, levantar possíveis “riscos” aos quais o setor está submetido ao adotar tal metodologia.

Com relação ao desenvolvimento dessa técnica de cruzamento entre as configurações, a mesma se justifica pois, dessa forma, o uso do modelo de série temporal no algoritmo de otimização PDDE torna-se viável. Em outras palavras, essa técnica “facilita” as atualizações das equações do modelo PAR(p) na PDDE.

Quanto aos “riscos” de se utilizar essa técnica, podem-se destacar pelo menos três principais: (i) mudança no processo gerador da série histórica; (ii) a não trivialidade de se testar a estacionariedade das séries históricas de ENAs; e, (iii) a questão das observações atípicas (*outliers*).

Esses “riscos” podem colocar em cheque requisitos fundamentais a Teoria Box & Jenkins e devem ser levantados, abordados e testados pelo SEB, uma vez que podem levar a estimação de funções de autocorrelação e autocorrelação

parciais viesadas e, conseqüentemente parâmetros não confiáveis para a geração das séries sintéticas.

Com relação ao processo gerador da série histórica, toda a teoria Box & Jenkins é calcada no pressuposto de que existe uma série temporal que segue um mesmo processo estocástico. Ao violar esse pressuposto, pode-se incorrer em problemas, como por exemplo, estimar parâmetros que não representam adequadamente a estrutura linear da série temporal estudada e gerar séries sintéticas assumindo distribuições dos resíduos que não refletem os momentos amostrais da série estudada.

A não trivialidade de se testar a estacionariedade das séries históricas de ENAs advém do fato de que a série temporal se modifica a cada estágio. Essa característica pode levar a estimação de parâmetros viesados e não estacionários.

Quanto aos *outliers*, os mesmos acontecem quando alguma observação atípica não é bem explicada pelo modelo. A inserção de novas configurações à série histórica poderá levar à potencialização da ocorrência de *outliers*²⁵, uma vez que o valor da ENA de um determinado mês poderá ser muito maior que seu valor histórico. E, conforme destacado por (Justel, et al., 2001), observações atípicas podem influenciar de forma não trivial a FAC, FACP, estimativas dos parâmetros dos modelos, diagnósticos de normalidade, a capacidade de geração de séries sintéticas etc.

Por fim, essa seção mostrou como o modelo PAR(p) está inserido no contexto do SEB e ainda, levantou as motivações para o uso das técnicas empregadas e os “cuidados” que a mesma requer. Prosseguindo, nas próximas seções são abordadas mais peculiaridades do modelo PAR(p)-IC.

4.1.3

As equações do Modelo Autorregressivo Periódico e a PDDE

Conforme abordado nessa tese, para o planejamento da operação energética foi proposta a técnica de Programação Dinâmica Dual Estocástica (PDDE). Tal metodologia utiliza a técnica de Decomposição de Benders (Benders, 1962) buscando encontrar estratégias ótimas para a operação dos subsistemas

²⁵ A análise de *outliers* deve ser feita mesmo quando não há a inserção de novas configurações.

interligados, enquanto as afluições são tratadas a partir de um modelo autorregressivo periódico.

Dado a natureza estocástica, o problema é dividido em vários subproblemas, tantos quantos forem os estágios considerados no horizonte de estudo. A função objetivo do subproblema de um determinado estágio t corresponde a minimizar o custo de operação presente, t , e o custo futuro, que representa o custo de operação desde o estágio seguinte, $t+1$, até o último estágio do horizonte de estudo. A Função de Custo Futuro é representada pelas Equações de Cortes de Benders no caso da PDDE. ((Pereira & Pinto, 1991); (Benders, 1962); (Pereira & Pinto, 1983); (CEPEL, 2011a)).

A utilização dessas equações implica em uma restrição aos parâmetros estimados pelo MGEA. Em outras palavras, um coeficiente autorregressivo negativo na equação da afluição de um mês específico pode provocar a presença de um coeficiente positivo associado à afluição prevista para um mês anterior, no corte de Benders que descreve a função de custo futuro deste mês. Um corte de Benders deste tipo leva a uma inconsistência ao se atualizar a previsão da afluição para o mês em curso ((Maceira & Damázio, 2004); (CEPEL, 2011a)).

Segundo (CEPEL, 2011a), coeficientes positivos nos cortes de Benders têm origem nos coeficientes autorregressivos negativos do modelo estocástico de afluições. Contudo, a presença de coeficientes autorregressivos negativos não necessariamente implica em coeficientes positivos nos cortes de Benders.

A ocorrência de um coeficiente autorregressivo negativo gerar coeficiente positivo associado à afluição prevista para o mês em curso no corte de Benders pode ser avaliada a partir do exame da equação de regressão da afluição de um mês futuro qualquer, em função da afluição prevista para o mês em curso e das afluições passadas.

Como argumenta (CEPEL, 2011a), visando conciliar os objetivos originais do modelo PAR(p) com a sua incorporação no modelo de otimização NEWAVE de forma explícita, deve-se buscar uma ordem suficientemente grande, mas que não produza coeficientes positivos nos cortes de Benders associados à afluição prevista para o mês em curso.

Uma solução para o problema, quando ele ocorre, é analisar novamente a função de autocorrelação parcial do mês em questão procurando o próximo coeficiente significativo. A ordem deste novo coeficiente será o novo valor de

tentativo para a ordem p do modelo PAR(p) do mês em foco. Com este valor de p , os novos coeficientes são estimados e uma análise dos coeficientes, como a que foi mencionada acima, é realizada. Este procedimento é repetido até que o conjunto de parâmetros de todos os meses não produzam contribuição negativa em qualquer mês futuro.

4.1.4

Geração de séries sintéticas e a questão das afluições negativas²⁶

O histórico de vazões e/ou ENAs disponível para o ajuste dos modelos é uma série temporal que, na verdade, consiste em apenas uma das possíveis realizações do processo estocástico gerador. O objetivo é, portanto, dado um modelo autorregressivo ajustado à série original, aproximar o comportamento estocástico e, sinteticamente, gerar tantas novas séries temporais quanto se queira, diferentes do histórico original, mas igualmente possíveis do ponto de vista estatístico.

O modelo PAR(p) é utilizado para aproximar este processo estocástico, isto é, o modelo deve permitir que, artificialmente, se façam tantos sorteios quantos forem necessários para o estudo em foco.

Esta seção apresenta os conceitos matemáticos e estatísticos utilizados para geração de cenários hidrológicos sintéticos da forma como estão implementados no modelo de geração energias afluentes, isto é, utilizando o modelo “PAR(p) Interconfigurações”. Matematicamente tem-se:

$$\left(\frac{Y_{r,m,c_m} - \mu_{m,c_m}}{\sigma_{m,c_m}} \right) = \varphi_1^{m,c_m} \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right) + \dots$$

$$+ \varphi_{p_m}^{m,c_m} \left(\frac{Y_{r,m-p,c_{m-p_m}} - \mu_{m-p_m,c_{m-p_m}}}{\sigma_{m-p_m,c_{m-p_m}}} \right) + a_t \quad (4.6)$$

Manipulando a (expressão 4.6) pode-se reescrevê-la da seguinte maneira:

²⁶ Seção baseada em (CEPEL, 2011a) com a inserção da ideia do cruzamento entre as configurações.

$$\begin{aligned}
Y_{r,m,c_m} = & \mu_{m,c_m} + \varphi_1^{m,c_m} \sigma_{m,c_m} \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right) + \dots \\
& + \varphi_{p_m}^{m,c_{m-p_m}} \sigma_{m,c_m} \left(\frac{Y_{r,m,c_{m-p_m}} - \mu_{m-p_m,c_{m-p_m}}}{\sigma_{m-p_m,c_{m-p_m}}} \right) + \sigma_{m-p_m,c_{m-p_m}} a_t
\end{aligned} \tag{4.7}$$

Um dos problemas frequentes em sistemas com reservatórios em cascata, onde as vazões incrementais podem ser muito pequenas, é que o procedimento para estimar as vazões incrementais numa cascata de reservatórios pode resultar em valores negativos para as vazões incrementais devido à consideração aproximada dos efeitos de propagação das vazões ao longo do rio. Dessa forma, ao gerar cenários sintéticos, para obter Y_{r,m,c_m} positivo é necessário impor a seguinte restrição (equação 4.8):

$$\begin{aligned}
Y_{r,m,c_m} = & \mu_{m,c_m} + \varphi_1^{m,c_m} \sigma_{m,c_m} \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right) + \dots \\
& + \varphi_{p_m}^{m,c_{m-p_m}} \sigma_{m,c_m} \left(\frac{Y_{r,m,c_{m-p_m}} - \mu_{m-p_m,c_{m-p_m}}}{\sigma_{m-p_m,c_{m-p_m}}} \right) + \sigma_{m-p_m,c_{m-p_m}} a_t > 0
\end{aligned} \tag{4.8}$$

Colocando a (inequação 4.8) em função de a_t :

$$\begin{aligned}
a_t > & - \left(\frac{\mu_{m,c_m}}{\sigma_{m,c_m}} \right) - \varphi_1^{m,c_m} \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right) - \dots \\
& - \varphi_{p_m}^{m,c_{m-p_m}} \left(\frac{Y_{r,m,c_{m-p_m}} - \mu_{m-p_m,c_{m-p_m}}}{\sigma_{m-p_m,c_{m-p_m}}} \right)
\end{aligned} \tag{4.9}$$

Chamando o lado direito da (inequação 4.9) de Δ , tem-se:

$$a_t > \Delta \tag{4.10}$$

Muitos pesquisadores consideram que os resíduos a_t apresentam distribuição Normal (Fernandez & Salas, 1986) e uma possível não-normalidade pode ser corrigida por transformações não-lineares (Box e Cox, 1965). Entretanto, como as

séries sintéticas produzidas serão utilizadas em modelos que calculam as estratégias ótimas de operação de um sistema multireservatórios, baseados em programação dinâmica dual estocástica, o modelo de geração de séries sintéticas deve ser aplicado diretamente à série temporal original e deve ser capaz de lidar com resíduos que apresentam um forte coeficiente de assimetria.

A solução adotada pelo Setor Elétrico Brasileiro foi ajustar uma distribuição lognormal com três parâmetros aos resíduos mensais a_t . A seguir são mostrados em detalhes os passos para modelagem da estrutura dos ruídos que serão gerados com base em simulações Monte Carlo – geração pseudoaleatória de observações de alguma distribuição de probabilidades e uso da amostra obtida para aproximar a função de interesse. Definindo as seguintes variáveis:

$$\xi_t \sim N(\mu_\xi, \sigma_\xi^2) \quad (4.11)$$

$$a_t = e^{\xi_t} + \Delta \quad (4.12a)$$

$$\xi_t = \ln(a_t - \Delta) \quad (4.12b)$$

Então, a variável a_t é lognormal, como segue:

$$a_t \sim \log normal(\mu_\xi, \sigma_\xi^2, \Delta) \quad (4.13)$$

Visto que os ruídos dos quatro subsistemas são gerados de forma conjunta, define-se o vetor aleatório multivariado $b_t \sim N_4(0,1)$. Logo, segue que:

$$\frac{\xi_t - \mu_\xi}{\sigma_\xi} = b_t \quad (4.14)$$

$$\xi_t = b_t \sigma_\xi + \mu_\xi \quad (4.15)$$

Logo, substituindo a (equação 4.15) em (4.12a) é possível reescrever a_t com a seguinte estrutura:

$$a_t = e^{b_t \sigma_\xi + \mu_\xi} + \Delta \quad (4.16)$$

O vetor b_t tem dimensão (4x1) e é gerado aleatoriamente com base em uma distribuição gaussiana padrão ($N(0,1)$). Os parâmetros Δ e σ_ξ são estimados de forma a preservar os momentos dos resíduos, conforme (Charbeneau, 1978).

$$\sigma_{\xi_t} = \sqrt{\ln(\theta)} \quad (4.17)$$

$$\theta = 1 + \frac{\sigma_{a(m)}^2}{\Delta^2} \quad (4.18)$$

$$\mu_{\xi_t} = 0,5 * \ln\left(\frac{\sigma_{a(m)}}{\theta(\theta - 1)}\right) \quad (4.19)$$

As definições anteriores são dadas no intuito de preservar os momentos dos resíduos. De acordo com (Charbeneau, 1978) a distribuição lognormal com três parâmetros só é aplicável a distribuições com coeficiente de assimetria positivo. Assim, a variável θ é a única raiz real (sempre ≥ 1) da equação a seguir, em que γ representa o coeficiente de assimetria dos resíduos gerados pelo modelo.

$$\theta^3 + 3\theta^2 - (4 + \gamma^2) = 0 \quad (4.20)$$

Para se gerar vazões mensais multivariadas, assume-se que os resíduos espacialmente não correlacionados, b_t , podem ser transformados em resíduos espacialmente correlacionados, W_t , da seguinte forma:

$$W_t = Db_t \quad (4.21)$$

D é uma matriz quadrada de dimensão igual ao número de subsistemas (4x4). A matriz D pode ser estimada por:

$$DD^T = U \quad (4.22)$$

Segundo PENNA (2009), no intuito de preservar dependências espaciais entre os subsistemas, faz-se U sendo a estimativa da matriz correlação espacial mensal entre as ENAs dos pares de subsistemas. A matriz D pode ser estimada pelo método da Decomposição de Choleski (Conte, 1965).

Para estimar D , adotam-se os seguintes procedimentos matemáticos: a equação matricial $DD^T = U$ é resolvida por meio da obtenção dos autovalores e autovetores da matriz U , assim:

$$PL^{1/2}L^{1/2}P^T = U \quad (4.23)$$

$$DD^T = U \quad (4.24)$$

Portanto, $D = PL^{1/2}$. Em que P e L são as matrizes ($k \times k$) de autovetores e autovalores de U , respectivamente. Assim, U é decomposta em P e L , obtendo-se posteriormente a matriz de carga, D . Tem-se, portanto, m matrizes D , $m = 12$ no caso de séries mensais.

A partir dos cenários gerados, o modelo NEWAVE calcula as estratégias ótimas de operação para diversos cenários hidrológicos. Na fase de otimização, o modelo estocástico “linear” PAR (p) será utilizado na estrutura de derivadas (Cortes de Benders) durante a fase de recursões do algoritmo de PDDE, responsável pela construção iterativa da função de custo futuro (Marcato, 2002).

Contudo, dada a necessidade de garantir vazões e/ou ENAs não-negativas e a partir de algumas manipulações matemáticas, é possível observar uma alteração na equação linear autorregressiva do modelo, assumido uma estrutura não-linear, gerando um problema de não convexidade na PDDE. Na próxima seção, observa-se, matematicamente, a origem desse problema.

4.1.5

A estrutura não-linear do PAR (p)

A demonstração da não linearidade do PAR (p) foi evidenciada por (Finardi, et al., 2009) e abordada por (Oliveira, 2010). Essa seção aborda as demonstrações feitas pelos autores num contexto de interconfigurações e avalia o impacto da equação final obtida na fase da construção da função de custo futuro.

Matematicamente, como na (equação 4.6), tem-se o “PAR (p) Interconfigurações” dado por:

$$\begin{aligned} \left(\frac{Y_{r,m,c_m} - \mu_{m,c_m}}{\sigma_{m,c_m}} \right) &= \varphi_1^{m,c_m} \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right) + \dots \\ &+ \varphi_{p_m}^{m,c_m} \left(\frac{Y_{r,m-p_m,c_{m-p_m}} - \mu_{m-p_m,c_{m-p_m}}}{\sigma_{m-p_m,c_{m-p_m}}} \right) + a_t \end{aligned} \quad (4.25)$$

Definindo as variáveis Φ_t^m e β_t , pode-se reescrever Y_t :

$$\Phi_t^{m,c_m} = \varphi_t^{m,c_m} \frac{\sigma_{m,c_m}}{\sigma_{m-p_m,c_{m-p_m}}} \quad (4.26)$$

$$\beta_t = \mu_{m,c_m} - \Phi_1^{m,c_m} \mu_{m-1,c_{m-1}} - \dots - \Phi_{p_m}^{m,c_m} \mu_{m-p_m,c_{m-p_m}} \quad (4.27)$$

$$Y_t = \Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} + \dots + \Phi_{p_m}^{m,c_m} Y_{r,m-p_m,c_{m-p_m}} + \beta_t + a_t \sigma_{m,c_m} \quad (4.28)$$

Assumindo que as vazões/ENAs não podem ser negativas, tem-se:

$$\Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} + \dots + \Phi_{p_m}^{m,c_m} Y_{r,m-p_m,c_{m-p_m}} + \beta_t + a_t \sigma_{m,c_m} > 0 \quad (4.29)$$

Isolando-se a_t :

$$a_t > \frac{(-\Phi_1^{m,c_m} Z_{r,m-1,c_{m-1}} - \dots - \Phi_{p_m}^{m,c_m} Z_{r,m-p_m,c_{m-p_m}} - \beta_t)}{\sigma_{m,c_m}} \quad (4.30)$$

Fazendo a_t seguir uma distribuição lognormal de três parâmetros e explicitando a variável Δ (equação 4.10):

$$a_t = e^{(b_t \sigma_{\xi_t} + \mu_{\xi_t})} + \Delta \quad (4.31)$$

$$\Delta = \frac{(-\Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} - \dots - \Phi_{p_m}^{m,c_m} Y_{r,m-p_m,c_{m-p_m}} - \beta_t)}{\sigma_{m,c_m}} \quad (4.32)$$

Voltando à (equação 4.28) e substituindo recursivamente, o modelo PAR (p), a princípio autorregressivo linear, assume a seguinte estrutura:

$$Y_{r,m,c_m} = \Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} + \dots + \Phi_{p_m}^{m,c_m} Y_{r,m-p_m,c_{m-p_m}} + \beta_t + (e^{(b_t \sigma_{\xi_t} + \mu_{\xi_t})} + \Delta) \sigma_{m,c_m} \quad (4.33)$$

$$Y_{r,m,c_m} = \Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} + \dots + \Phi_{p_m}^{m,c_m} Y_{r,m-p_m,c_{m-p_m}} + \beta_t + (e^{(b_t \sigma_{\xi_t} + \mu_{\xi_t})}) \sigma_{m,c_m} + \Delta \sigma_{m,c_m} \quad (4.34)$$

$$Y_{r,m,c_m} = \Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} + \dots + \Phi_{p_m}^{m,c_m} Y_{r,m-p_m,c_{m-p_m}} + \beta_t \quad (4.35)$$

$$+ (e^{(b_t \sigma_{\xi_t} + \mu_{\xi_t})}) \sigma_{m,c_m} + \frac{(-\Phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} - \dots - \Phi_{p_m}^m Y_{r,m-p_m,c_{m-p_m}} - \beta_t)}{\sigma_{m,c_m}} \sigma_{m,c_m}$$

$$Y_t = (e^{(b_t \sigma_{\xi_t} + \mu_{\xi_t})}) \sigma_{m,c_m} \quad (4.36)$$

Incorporando a correlação entre os subsistemas, tem-se:

$$Y_t = (e^{(W_t \sigma_{\xi_t} + \mu_{\xi_t})}) \sigma_{m,c_m} \quad (4.37)$$

Assim, dado a necessidade de garantir $Y_t > 0$, é possível notar a alteração na estrutura do modelo PAR (p), tornando-o não-linear. Na fase de otimização, a formulação matemática do problema de otimização é válida apenas em problemas lineares. Portanto, ao violar as hipóteses da formulação matemática do problema de otimização, perde-se a possibilidade de usar a técnica que está calcada em Programação Linear, gerando-se um problema de não convexidade na PDDE.

Sabe-se que durante as recursões do algoritmo de otimização, são calculadas derivadas e estas incluem as do modelo PAR(p). Para ilustrar a ideia exposta acima, a seguir são mostrados os cálculos da derivada em relação à primeira defasagem (análogo para as demais até a ordem do período m) (Finardi, et al., 2009).

$$\frac{\partial Y_{r,m,c_m}}{\partial Y_{r,m-1,c_{m-1}}} = \frac{\partial((e^{(W_t \sigma_{\xi_t} + \mu_{\xi_t})}) \sigma_{m,c_m})}{\partial Y_{r,m-1,c_{m-1}}} \quad (4.38)$$

$$\frac{\partial \Delta}{\partial Y_{r,m-1,c_{m-1}}} = \Delta' = -\varphi_1 \quad (4.39)$$

$$\frac{\partial \theta}{\partial Y_{r,m-1,c_{m-1}}} = \theta' = \sigma_a^2 \left(\frac{-2}{\Delta^3} \right) \Delta' = \frac{2\varphi_1 \sigma_a^2}{\Delta^3} \quad (4.40)$$

$$\frac{\partial \sigma_{\xi_t}}{\partial Y_{r,m-1,c_{m-1}}} = \sigma_{\xi_t}' = \frac{1}{2 \ln(\theta)} \frac{1}{\theta} \theta' = \frac{2\varphi_1 \sigma_a^2}{2\theta \Delta^3 \ln(\theta)} \quad (4.41)$$

$$\frac{\partial \mu_{\xi_t}}{\partial Y_{r,m-1,c_{m-1}}} = \mu_{\xi_t}' = \frac{(\theta^2 - \theta)^{0.5}}{\sigma_a} \frac{1}{2(\theta^2 - \theta)^{1.5}} (2\theta\theta' - \theta') = \frac{2\varphi_1\sigma_a(2\theta - 1)}{\Delta^3(\theta^2 - \theta)} \quad (4.42)$$

$$\frac{\partial Y_{r,m,c_m}}{\partial Y_{r,m-1,c_{m-1}}} = \sigma_m \frac{\partial (e^{(W_r\sigma_{\xi_t}) + \mu_{\xi_t}})}{\partial Y_{r,m-1,c_{m-1}}} = \sigma_m (e^{(W_r\sigma_{\xi_t}) + \mu_{\xi_t}}) (\xi_t \sigma_{\xi_t}' + \mu_{\xi_t}') \quad (4.43)$$

$$\frac{\partial Y_{r,m,c_m}}{\partial Y_{r,m-1,c_{m-1}}} = \sigma_m (e^{(W_r\sigma_{\xi_t}) + \mu_{\xi_t}}) \left(\xi_t \frac{2\varphi_1\sigma_a^2}{2\theta\Delta^3 \ln(\theta)} + \frac{2\varphi_1\sigma_a(2\theta - 1)}{\Delta^3(\theta^2 - \theta)} \right) \quad (4.44)$$

Nota-se, portanto, o impacto da não-linearidade assumida pelo PAR (p) no cálculo das derivadas mostradas anteriormente. Assim, uma vez que o modelo PAR(p), conforme está implementado no SEB, torna-se não linear ao ter que usar o artifício da distribuição lognormal para não gerar cenários negativos, essa tese aborda um modelo não linear, mas que não precisa de manipulações para não gerar afluências negativas.

Concluindo, essa seção mostrou algumas características inerentes à modelagem e geração dos cenários feitos pelo Setor Elétrico Brasileiro, as quais são utilizadas para cumprir certas restrições advindas do setor, mas que, de certo modo, não necessariamente estão de acordo com os pressupostos metodológicos dos modelos adotados. Ressaltando-as têm-se: o modelo PAR(p)-IC, a questão da afluência negativa, a definição da ordem máxima, o terceiro parâmetro da lognormal e a questão da não linearidade.

4.2

Estudo de caso: Modelagem e Geração de Energias Afluentes no contexto do SEB

Nesta seção apresentam-se os resultados da modelagem e da geração de cenários utilizando-se o modelo PAR(p) como implementado no contexto do SEB²⁷, isto é, modelo PAR(p) interconfigurações, identificação da ordem da direita para a esquerda (D-E), ordem máxima 6 com possíveis correções de

²⁷ O estudo apresentado faz uso do PMO de janeiro/2011 e da série histórica de ENA referente ao período de jan/1931 à dez/2010.

afluências negativas, geração dos cenários através da distribuição lognormal, via simulação de Monte Carlo e utilizando dados da tendência hidrológica.

4.2.1

Ajuste do Modelo “PAR(p) Interconfigurações”

Seguindo os critérios apresentados no último parágrafo da (seção 4.1.5), neste ponto trata-se das questões relacionadas ao ajuste do modelo PAR(p)-IC. Para iniciar a análise, aborda-se a questão da possível não estacionariedade do modelo gerada pelo aumento da vazão em alguns postos, mas principalmente, pela estimação do modelo PAR(p) intercorrelacionando as configurações, lembrando que a padronização dos dados só é feita para a geração dos cenários. Nas próximas seções são abordadas as ordens identificadas, os parâmetros estimados e os resíduos do modelo e a questão dos *outliers*.

4.2.1.1

A questão da estacionariedade

Conforme mencionado em (4.1.2), existe a não trivialidade de se testar a estacionariedade das séries históricas de ENAs, uma vez que a série se modifica a cada estágio. Dentro deste contexto de interconfigurações, a ideia foi testar a estacionariedade da série histórica a cada intervalo de 12 períodos, conforme pode ser observado na (Figura 4.9). Explicando de maneira mais clara, a cada 12 períodos, monta-se uma nova matriz de configurações (79x12, neste exemplo) e faz-se os testes de estacionariedade. Para um PMO de janeiro, estudo de caso apresentado nessa tese, observa-se que serão feitos 73 testes de estacionariedade para cada subsistema, neste caso, subsistemas Sudeste/Centro-Oeste, Sul, Nordeste e Norte.

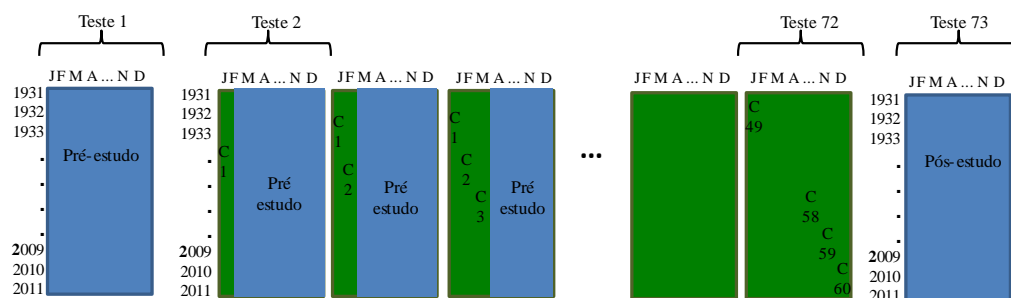


Figura 4.9 – Algoritmo para testar a estacionariedade “Intercorrelações”

Os dois testes escolhidos para verificar a estacionariedade foram o teste de Dickey Fuller Aumentado (ADF) (Dickey & Fuller, 1979) e Phillips Perron (PP) (Phillips & Perron, 1988). Ambos os testes tem como hipótese nula a presença de raiz unitária.

Como pode ser observado na (Tabela 4.2), para este estudo de caso, os 496 testes indicaram estacionariedade, ou seja, ambos rejeitaram a hipótese nula de presença de raiz unitária.

Apesar deste resultado, é importante destacar a importância de se testar a estacionariedade, uma vez que séries históricas não estacionárias podem levar a estimação de parâmetros viesados e não consistentes. Dessa forma, o SEB deveria, antes de estimar os parâmetros do modelo PAR(p) Interconfigurações, testar a estacionariedade da série histórica interconfigurações conforme apresentado nessa seção.

Tabela 4.2 – Testes de Estacionariedade

Testes/Subsistemas	Sudeste/Centro-Oeste	Sul	Nordeste	Norte
Phillips-Perron	-4,86 ¹ (0,001) ³	-9,20 ¹ (0,001) ³	-6,78 ¹ (0,001) ³	-6,34 ¹ (0,001) ³
Augmented Dickey Fuller	-4,86 ² (0,001) ³	-9,20 ² (0,001) ³	-6,78 ² (0,001) ³	-6,34 ² (0,001) ³

(1) Phillips-Perron test statistic (média dos 62 testes)

(2) Augmented Dickey-Fuller test statistic (média dos 62 testes)

(3) Estatística p-valor

Fonte: autor, utilizando o software Matlab.

4.2.1.2

A questão da identificação da ordem e os parâmetros

Conforme discutido no capítulo 3 dessa tese há diversas maneiras de identificar a ordem dos modelos da família Box & Jenkins ((Schwarz, 1978);

(Akaike, 1987) (Oliveira, 2010); (CEPEL, 2011a)). Ainda, verificou-se que o Setor Elétrico Brasileiro identifica os modelos PAR(p)-IC observando a FACP-IC da direita para esquerda, o qual foi denotada nesse trabalho como critério D-E. Tal critério significa que, observando-se a FACP-IC do último *lag* para o primeiro, ao encontrar um valor de $\phi_{k,k}^m$ fora do intervalo de confiança $1,96/\sqrt{N}$, assume-se tal ordem para o modelo PAR(p)-IC.

Outras duas peculiaridades do SEB, quanto à identificação da ordem, são, a fixação da ordem máxima e a restrição associada ao corte de Benders (Benders, 1962). Com relação à fixação da ordem máxima, definiu-se que a ordem do modelo PAR(p)-IC não pode ser maior que 6 (seis). Apesar de não haver na literatura nenhuma evidência matemática ou empírica que justifique essa decisão, pressupõe-se que a precipitação ocorrida há mais de seis meses não irá afetar a precipitação do mês corrente. Como não é objetivo desse trabalho validar essa decisão, nesse estudo de caso, será adotado ordem máxima 6, assim como definido pelo SEB.

Quanto à restrição associada ao corte de Benders, conforme observado na (seção 4.1.3) deste trabalho, visando conciliar os objetivos originais do modelo PAR(p)-IC com a sua incorporação no modelo de otimização NEWAVE de forma explícita, deve-se buscar uma ordem suficientemente grande, mas que não produza coeficientes positivos nos cortes de Benders associados à afluência prevista para o mês em curso (CEPEL, 2011a). A solução para o problema, quando ele ocorre, é analisar novamente a FACP-IC do mês em questão procurando o próximo coeficiente significativo, para assim, definir a ordem do modelo.

Dentro desse contexto, o modelo SEB assume as seguintes características: identificação da ordem da direita para a esquerda, *lag* máximo 6 e restrições impostas pela PDDE (modelo SEB). Nas figuras (Figura 4.10), (Figura 4.11), (Figura 4.12) e (Figura 4.13) observam-se os resultados referentes as ordens identificadas. Como se pode observar, de uma maneira geral as ordens não se alteram com as mudanças de configuração, mas há meses em que verificam-se mudanças, como, por exemplo, dezembro/Sudeste-Centro-Oeste, abril/Sul e julho, agosto, setembro, outubro e dezembro do subsistema Norte.

Quanto a identificação das ordens, alguns pontos precisam ser ressaltados e analisados pelo SEB. Como por exemplo, o critério de identificação “D-E”. Como já observado em (Oliveira & Souza, 2011), critérios mais parcimoniosos na identificação da ordem, podem gerar cenários tão bons quanto modelos com ordens elevadas. Ainda, a justificativa dada por (CEPEL, 2011a) de que ordens maiores são mais capazes de capturar longos períodos de chuva e/ou seca, parece não ser válida, uma vez que, como demonstrado em (Oliveira & Souza, 2011), modelos mais parcimoniosos não violam os testes de sequências positivas ou negativas.

Um exercício que deve ser feito pelo SEB é analisar os diferentes critérios de identificação (*e.g.* Bootstrap, Critério de Akaike, Critério de Schwarz, “esquerda-direita”, “direita-esquerda”), gerar as séries sintéticas e decidir qual seria o melhor critério para esse tipo de série temporal.

É importante lembrar também que, ordens menores podem evitar o problema das afluições negativas, isto é, podem evitar a redução *ad hoc*, da ordem dos modelos e ainda, a definição, também *ad hoc* da ordem máxima.

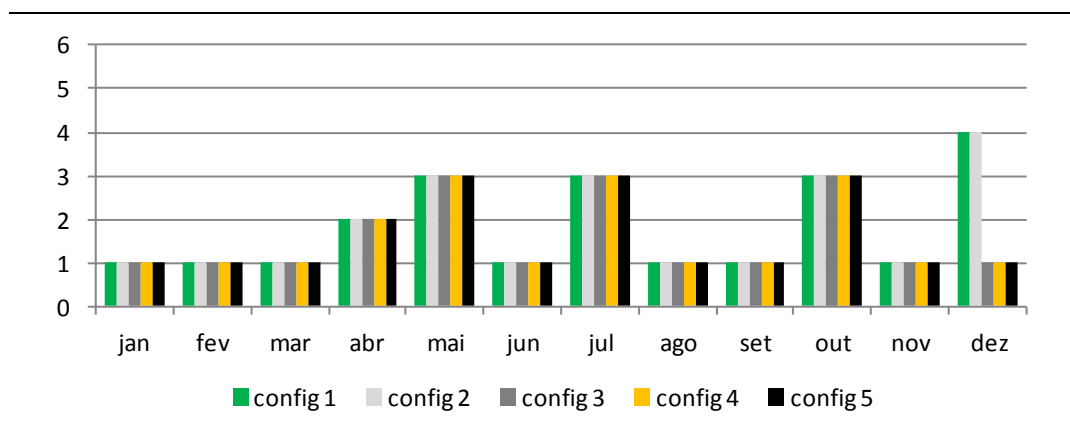


Figura 4.10 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Sudeste/Centro-Oeste

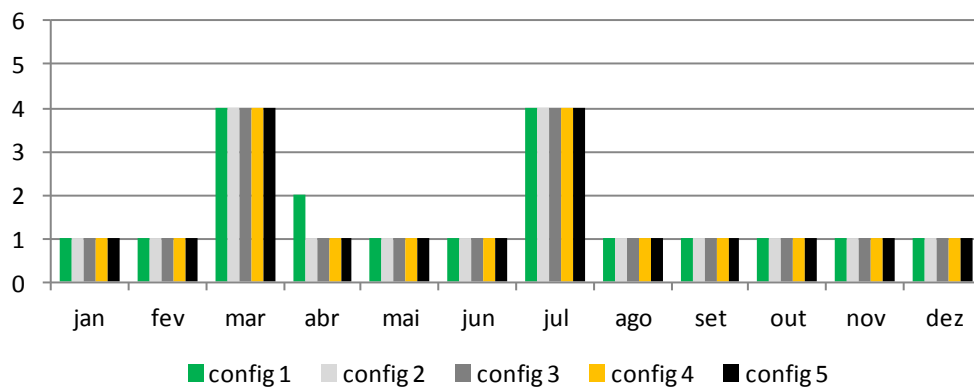


Figura 4.11 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Sul

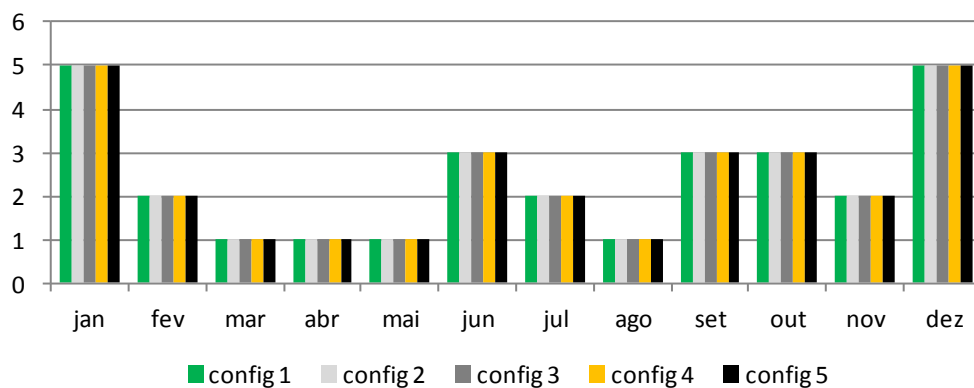


Figura 4.12 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Nordeste

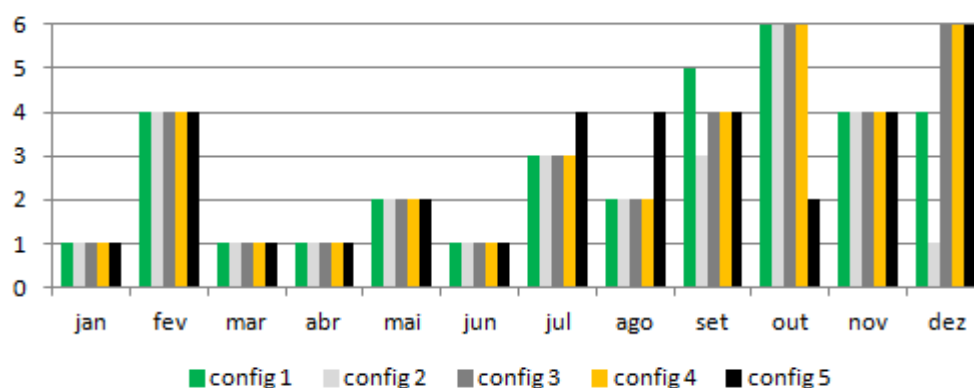


Figura 4.13 – Ordens identificadas de acordo com o mês e sua configuração correspondente – Subsistema Norte

Com relação aos parâmetros dos modelos, conforme apresentado, os mesmos foram estimados via Método dos Momentos, que, conforme ressaltado

por (Hipel & McLeod, 1994) são uma boa aproximação para modelos AR da estimação via Máxima Verossimilhança.

Um ponto que deve ser observado, é apenas com relação à estacionariedade dos parâmetros dos modelo, isto é, a escolha das ordens do modelo como é feita pelo SEB e uma possível não estacionariedade no histórico, pode fazer com que os parâmetros estimados resultem em modelos não estacionários, violando importantes restrições da teoria Box & Jenkins, o que pode acarretar em cenários que não refletem a realidade da série histórica.

Dessa forma, cabe ao Setor Elétrico Brasileiro, procurar alternativas que consigam evitar esses possíveis problemas, como por exemplo, um teste de estacionariedade dos parâmetros.

4.2.1.3

Análise dos Ruídos

De acordo com a teoria Box & Jenkins (Box, et al., 2008), na qual está inserida os modelos PAR(p)-IC, os resíduos do modelo devem assumir o seguinte comportamento:

$$E(a_t) = 0 \quad (4.45a)$$

$$E(a_t a_{t-k}) = 0 \dots \text{para todo } k \neq 0 \quad (4.45b)$$

$$E(a_t a_{t-k}) = \sigma^2 \quad \text{para } k=0 \quad (4.45c)$$

$$a_t \sim NID(0, \sigma^2) \quad (4.45d)$$

Neste contexto, o objetivo dessa seção é observar e testar os resíduos gerados pelos modelos PAR(p)-IC ajustados as séries históricas e confirmar ou não se os mesmos estão cumprindo os requisitos explicitados acima.

Vale lembrar que, como é um estudo de caso, testou-se apenas os resíduos para a primeira configuração do parque hidrológico, sendo que o ideal seria testá-los para todas as configurações, uma vez que, conforme observado em (4.2.1.2) pode haver alteração na ordem e, conseqüentemente nos parâmetros do modelo, logo, o comportamento dos resíduos não será o mesmo.

Ainda, conforme destacado na (seção 4.1.2), outra característica do ruído que precisa ser testada é a detecção ou não de *outliers*, pois tais observações podem comprometer os procedimentos usuais de modelagem linear de uma série histórica. Em outras palavras, podem induzir a uma identificação incorreta do modelo da família Box & Jenkins e a uma estimação viesada dos parâmetros deste modelo ((Miranda, 2001); (Ota, 1996)).

Com relação a (equação 4.45a), ao testar a média dos resíduos observou-se a não violação da hipótese nula da média ser igual a zero, dessa forma, validando essa restrição.

Quanto a (restrição 4.45b) observa-se, de acordo com o teste de Ljung-Box (Box, et al., 2008) e a FAC dos resíduos ((Figura 4.14); (Figura 4.15); (Figura 4.16) e (Figura 4.17)) que os modelos ajustados “capturaram” toda a relação linear existente no histórico, podendo-se concluir que os resíduos seguem um processo ruído branco.

FAC	FACP	AC	ACP	Q-Stat	Prob	
		1	0.008	0.008	0.0619	0.803
		2	-0.021	-0.021	0.4675	0.792
		3	-0.019	-0.019	0.8204	0.845
		4	-0.050	-0.050	3.1605	0.531
		5	0.011	0.011	3.2675	0.659
		6	0.021	0.018	3.6794	0.720
		7	-0.016	-0.017	3.9132	0.790
		8	0.038	0.037	5.2617	0.729
		9	-0.010	-0.010	5.3641	0.801
		10	0.008	0.011	5.4190	0.861
		11	0.058	0.057	8.6871	0.651
		12	-0.047	-0.045	10.836	0.543

Figura 4.14 – FAC dos resíduos – Subsistema Sudeste/Centro-Oeste

FAC	FACP	AC	ACP	Q-Stat	Prob	
		1	0.014	0.014	0.1812	0.670
		2	-0.051	-0.051	2.6269	0.269
		3	0.002	0.004	2.6325	0.452
		4	0.009	0.006	2.7042	0.608
		5	0.030	0.030	3.5575	0.615
		6	0.026	0.026	4.2045	0.649
		7	-0.039	-0.037	5.6531	0.581
		8	0.042	0.046	7.3821	0.496
		9	0.001	-0.005	7.3824	0.597
		10	-0.022	-0.019	7.8494	0.644
		11	0.024	0.023	8.3888	0.678
		12	0.062	0.060	12.043	0.442

Figura 4.15 – FAC dos resíduos – Subsistema Sul

FAC	FACP	AC	ACP	Q-Stat	Prob	
		1	0.010	0.010	0.0887	0.766
		2	-0.043	-0.043	1.8338	0.400
		3	-0.042	-0.041	3.4842	0.323
		4	0.015	0.014	3.6884	0.450
		5	0.025	0.021	4.2626	0.512
		6	0.027	0.026	4.9469	0.551
		7	-0.022	-0.019	5.3975	0.612
		8	0.060	0.064	8.8405	0.356
		9	0.035	0.034	10.001	0.350
		10	-0.011	-0.010	10.120	0.430
		11	0.020	0.028	10.521	0.484
		12	-0.017	-0.017	10.794	0.547

Figura 4.16 – FAC dos resíduos – Subsistema Nordeste

FAC	FACP	AC	ACP	Q-Stat	Prob	
		1	-0.021	-0.021	0.4044	0.525
		2	0.004	0.004	0.4228	0.809
		3	-0.016	-0.015	0.6546	0.884
		4	0.045	0.045	2.6192	0.623
		5	0.014	0.016	2.7976	0.731
		6	0.010	0.010	2.8860	0.823
		7	0.041	0.043	4.4936	0.721
		8	-0.004	-0.004	4.5063	0.809
		9	0.010	0.009	4.6087	0.867
		10	-0.020	-0.019	4.9864	0.892
		11	0.004	-0.001	5.0026	0.931
		12	-0.030	-0.030	5.8477	0.924

Figura 4.17 – FAC dos resíduos – Subsistema Norte

Para testar a (restrição 4.5c, variância constante) foram feitos dois testes nos resíduos, a FAC dos resíduos ao quadrado e o teste para heterocedasticidade ARCH (Engle & Ng, 1991); (Lumsdaine & Serena, 1993)). Para o subsistema Sudeste/Centro-Oeste, no que diz respeito à FAC dos resíduos ao quadrado é possível observar (Figura 4.18) algumas autocorrelações estatisticamente significativas, porém com valores muito pequenos. Para corroborar a hipótese de não existência de efeito GARCH, fez-se, o teste ARCH (hipótese nula de não existência de efeito ARCH nos erros) (Figura 4.18), evidenciando-se a não rejeição da hipótese nula.

FAC	FACP	AC	ACP	O-Stat	Prob	
		1	0.059	0.059	3.2894	0.070
		2	0.013	0.009	3.4472	0.178
		3	0.030	0.029	4.3060	0.230
		4	0.099	0.095	13.583	0.009
		5	0.013	0.002	13.727	0.017
		6	-0.033	-0.037	14.799	0.022
		7	-0.004	-0.005	14.812	0.038
		8	-0.014	-0.022	14.989	0.059
		9	0.029	0.032	15.788	0.071
		10	0.038	0.041	17.035	0.074
		11	0.082	0.081	23.548	0.015
		12	0.060	0.052	26.990	0.008

Teste para Heterocedasticidade: ARCH			
Estatística F	3.272	Prob. F(1,945)	0.071
Obs*R-quadrado	3.268	Prob. Chi-Square(1)	0.071

Figura 4.18 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Sudeste/Centro-Oeste

Quando a análise é feita para os subsistemas Sul (Figura 4.19), Nordeste (Figura 4.20) e Norte (Figura 4.21), os resultados vistos para o Sudeste não se

mantêm. Como se pode observar, para os três subsistemas têm-se uma FAC dos resíduos ao quadrado com correlações significantes, e ainda, corroborando tal análise, os testes ARCH feitos nos resíduos indicam a presença de um comportamento GARCH, uma vez que rejeitou-se a hipótese nula de não existência de efeito ARCH.

Dessa forma, ao observar os resíduos dos modelos dos subsistemas Sul, Nordeste e Norte verifica-se a presença de um componente não linear nos dados que não está sendo modelado pelo modelo PAR(p)-IC. Conforme destacado por (Engle & Ng, 1991), a violação da hipótese de variância constante tem como consequência parâmetros não tendenciosos, mas ineficientes e com variâncias estimadas tendenciosas; como consequência, as estatísticas de teste do modelo tornam-se não confiáveis.

FAC	FACP	AC	ACP	O-Stat	Prob
1	0.074	0.074	5.1748	0.023	
2	0.112	0.108	17.207	0.000	
3	-0.013	-0.029	17.378	0.001	
4	0.076	0.068	22.865	0.000	
5	0.011	0.006	22.985	0.000	
6	0.080	0.065	29.176	0.000	
7	0.000	-0.009	29.176	0.000	
8	0.090	0.073	36.969	0.000	
9	0.014	0.007	37.165	0.000	
10	-0.022	-0.051	37.642	0.000	
11	-0.009	-0.002	37.717	0.000	
12	0.019	0.013	38.059	0.000	

Teste para Heterocedasticidade: ARCH			
Estatística F	5.172	Prob. F(1,945)	0.023
Obs*R-quadrado	5.154	Prob. Chi-Square(1)	0.023

Figura 4.19 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Sul

FAC	FACP	AC	ACP	O-Stat	Prob
1	0.190	0.190	34.447	0.000	
2	0.136	0.103	51.937	0.000	
3	-0.028	-0.074	52.669	0.000	
4	-0.074	-0.074	57.867	0.000	
5	-0.069	-0.033	62.456	0.000	
6	-0.076	-0.044	67.991	0.000	
7	-0.013	0.014	68.158	0.000	
8	-0.043	-0.040	69.955	0.000	
9	0.074	0.081	75.259	0.000	
10	0.142	0.125	94.697	0.000	
11	0.079	0.008	100.74	0.000	
12	0.060	0.010	104.26	0.000	

Teste para Heterocedasticidade: ARCH			
Estatística F	35.678	Prob. F(1,945)	0.000
Obs*R-quadrado	34.453	Prob. Chi-Square(1)	0.000

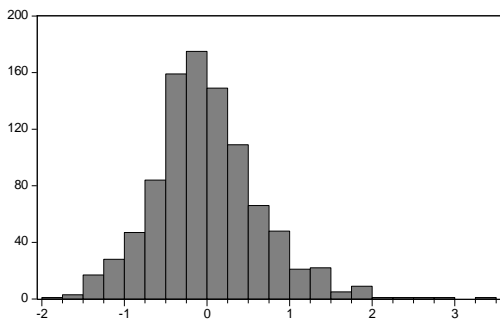
Figura 4.20 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Nordeste

FAC	FACP	AC	ACP	O-Stat	Prob
1	0.114	0.114	12.348	0.000	
2	0.041	0.028	13.931	0.001	
3	-0.004	-0.012	13.945	0.003	
4	-0.008	-0.008	14.012	0.007	
5	0.015	0.017	14.217	0.014	
6	-0.004	-0.007	14.233	0.027	
7	0.035	0.035	15.406	0.031	
8	-0.026	-0.034	16.051	0.042	
9	0.034	0.039	17.168	0.046	
10	0.011	0.005	17.289	0.068	
11	0.026	0.022	17.925	0.083	
12	0.075	0.069	23.391	0.025	

Teste para Heterocedasticidade: ARCH			
Estatística F	12.435	Prob. F(1,945)	0.000
Obs*R-quadrado	12.299	Prob. Chi-Square(1)	0.001

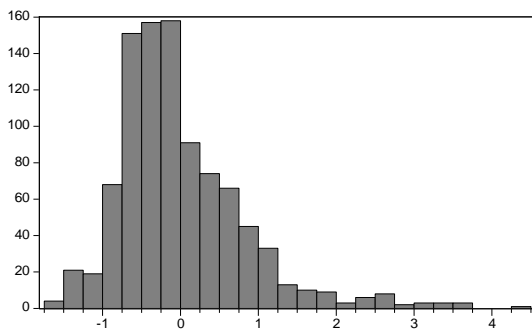
Figura 4.21 – FAC dos resíduos ao quadrado e Teste ARCH para os resíduos – Subsistema Norte

Quanto a hipótese (4.45d) de normalidade dos resíduos, conforme se observa nas figuras 4.22 a 4.25, tal restrição não foi atendida em qualquer subsistema. Em todos os casos verificou-se a existência de assimetria positiva e excesso de curtose. Para corroborar este resultado, ao observar a estatística de teste Jarque-Bera (Jarque & Bera, 1981), verifica-se que em todos os casos há rejeição da hipótese nula de normalidade.



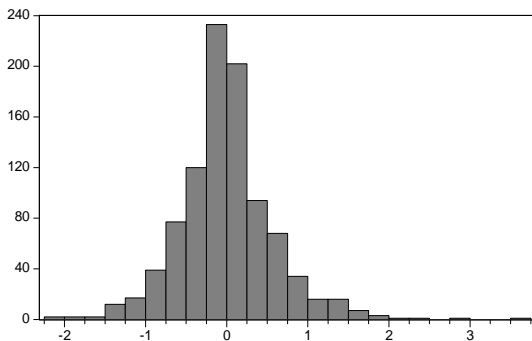
Estatística	Resultado
Média	0.00
Máximo	3.27
Mínimo	-1.98
Desv. Pad	0.64
Assimetria	0.60
Curtose	4.46
Jarque-Bera	143.32
Prob	0.00

Figura 4.22 – Histograma e estatísticas dos Resíduos – Subsistema Sudeste/Centro-Oeste



Estatística	Resultado
Média	0.00
Máximo	4.37
Mínimo	-1.71
Desv. Pad	0.83
Assimetria	1.43
Curtose	6.32
Jarque-Bera	762.30
Prob	0.00

Figura 4.23 – Histograma e estatísticas dos Resíduos – Subsistema Sul



Estatística	Resultado
Média	0.00
Máximo	3.52
Mínimo	-2.14
Desv. Pad	0.58
Assimetria	0.53
Curtose	6.06
Jarque-Bera	415.20
Prob	0.00

Figura 4.24 – Histograma e estatísticas dos Resíduos – Subsistema Nordeste

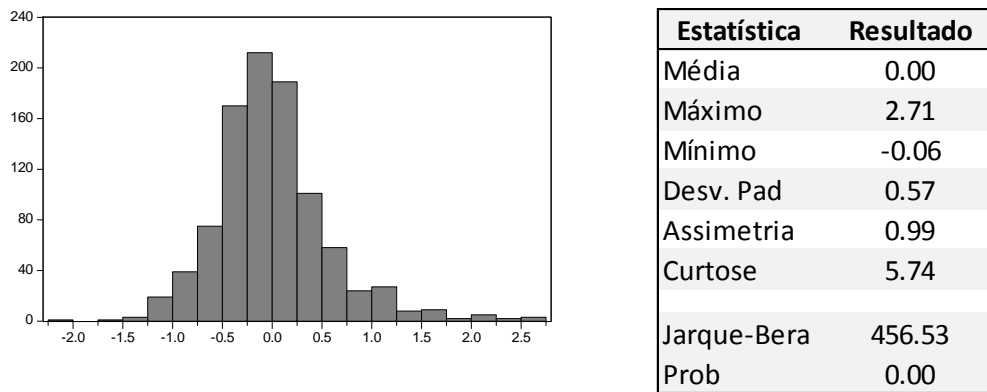


Figura 4.25 – Histograma e estatísticas dos Resíduos – Substema Norte

Por fim, lembrando o que foi dito nesse trabalho com relação a presença de *outliers* (seção 4.1.2) e testando a possível presença dos mesmos nas séries de energia, observa-se nas figuras 4.26 a 4.29, as quais mostram os resíduos do modelos PAR(p)-IC e dois “cortes” representando dois desvios-padrões, que para todos os subsistemas existem observações atípicas que podem viesar os parâmetros estimados pelo modelo. Para corroborar tal observação foram estimados modelos levando-se em conta a presença dessas observações atípicas e observou-se que as mesmas apresentavam significância estatística.

Ainda, como destacado no início dessa seção, todos os testes foram feitos utilizando-se os dados da primeira configuração e conforme destacado na (seção 4.2.1) deste trabalho, a questão do modelo interconfigurações pode potencializar esse tipo de problema, levando a parâmetros não confiáveis.

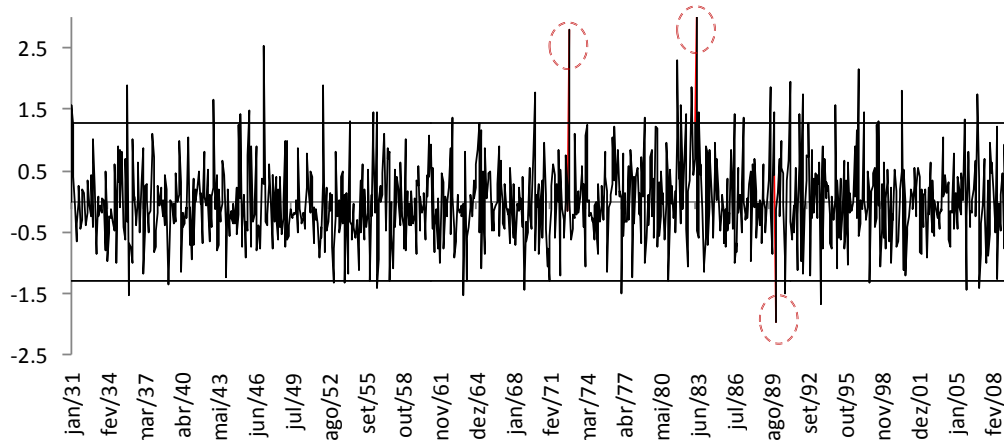


Figura 4.26 – Gráfico dos resíduos para detecção de *outliers* – Substema Sudeste/Centro-Oeste

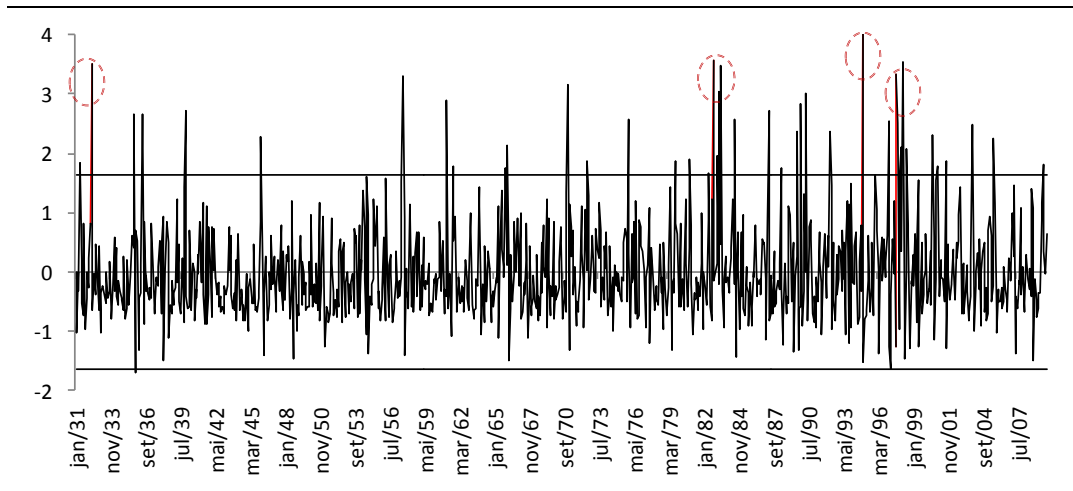


Figura 4.27 – Gráfico dos resíduos para detecção de outliers – Substema Sul

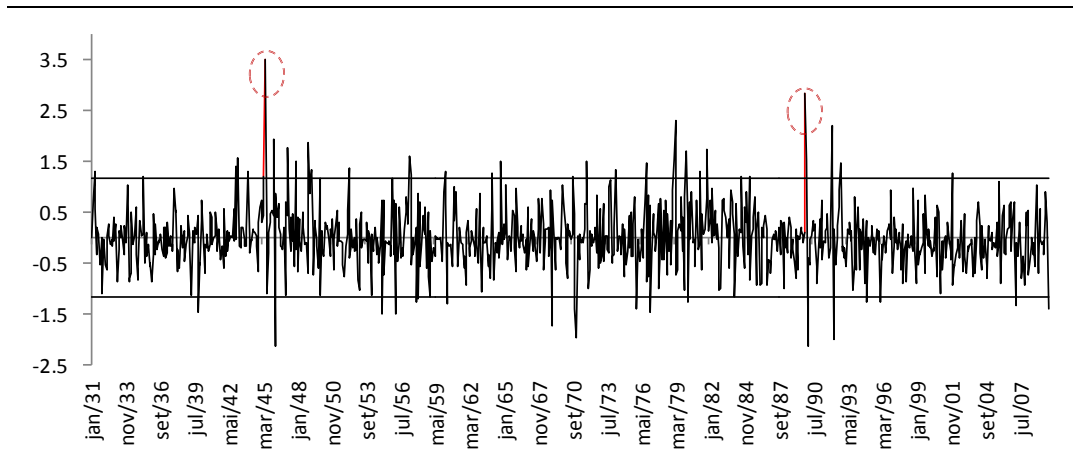


Figura 4.28 – Gráfico dos resíduos para detecção de outliers – Substema Nordeste

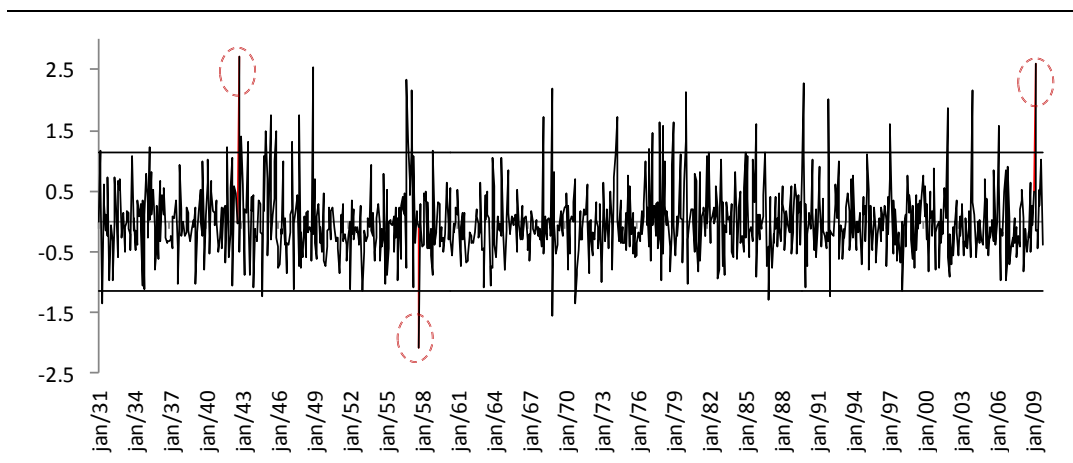


Figura 4.29 – Gráfico dos resíduos para detecção de outliers – Substema Norte

Dando sequência ao trabalho, na (seção 4.2.2) são apresentados os resultados de geração das séries sintéticas, bem como, algumas estatísticas de teste que servem para testar a adequabilidade dos cenários gerados.

4.2.2

Ajuste dos cenários

Nesta seção serão apresentados os resultados da geração de cenários sintéticos ENA dos quatro subsistemas brasileiros. Com intuito de avaliar o modelo PAR-IC inserido no contexto do SEB e para subsidiar possíveis análises e comparações com a proposta Gama.

4.2.2.1

Séries sintéticas e teste de média dos cenários

Nesta seção mostram-se as séries sintéticas geradas pelo modelo PAR(p)-IC, via simulação Monte Carlo, obtidas para cada um dos subsistemas e as médias das mesmas em comparação à média histórica. As figuras (Figura 4.30) e (Figura 4.31) retratam as séries sintéticas simuladas, os valores médios dos 200 cenários de energia natural afluente (linha pontilhada preta) e a média histórica (linha contínua vermelha) reproduzida para os dez anos de cenários sintéticos (120 meses). Como se pode observar, nos quatro subsistemas a média histórica e a média dos cenários têm um comportamento muito similar.

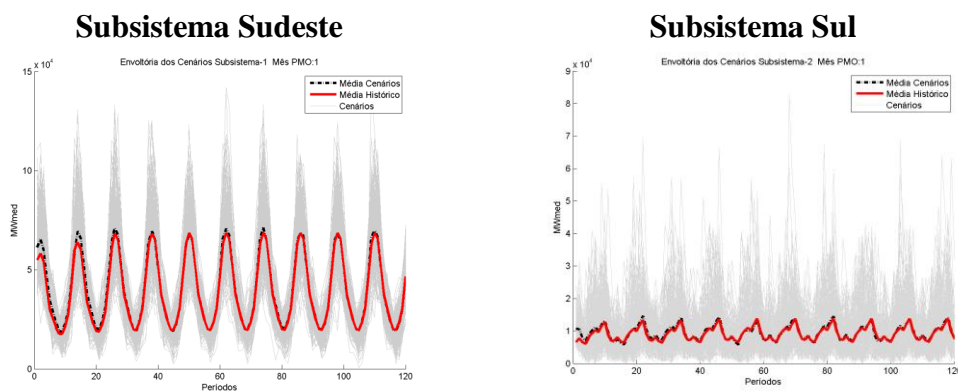


Figura 4.30 – Séries Sintéticas e médias dos cenários

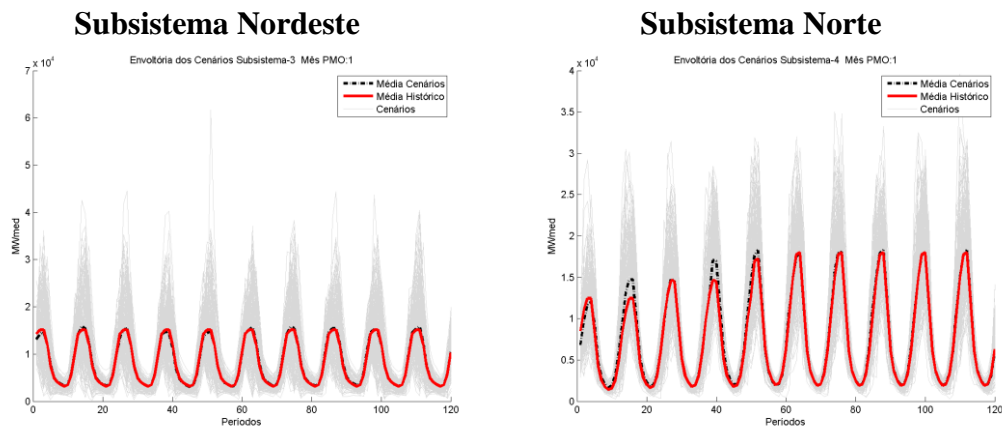


Figura 4.31 - Séries Sintéticas e médias dos cenários

Para corroborar o que foi observado nas figuras (Figura 4.30) e (Figura 4.31), se faz um teste de média (Casella & Berger, 2001) para verificar se a média dos cenários gerados é estatisticamente igual a média do histórico, apresentando a evolução dos p-valores ((Figura 4.32); (Figura 4.33)). Testa-se, por meio da estatística t , a hipótese da média dos 200 elementos de cada um dos 120 períodos da matriz de cenários ser igual à média das observações de cada um dos meses das configurações específicas da série histórica.

Conforme se observa nas figuras (Figura 4.32) e (Figura 4.33) há uma “aprovação” de mais de 90% dos cenários para os subsistemas Sudeste/Centro-Oeste, Sul e Nordeste. Para o subsistema Norte o resultado foi de 79% de aprovação. Pode-se concluir que os resultados são satisfatórios para todos os subsistemas, o que indica que as médias dos períodos gerados são estatisticamente coincidentes com o histórico de configurações.

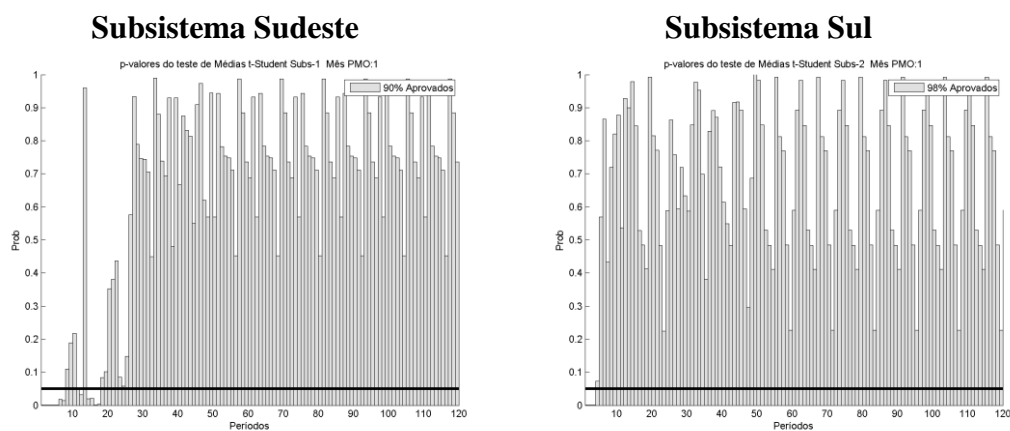


Figura 4.32 – P-valores do Teste t

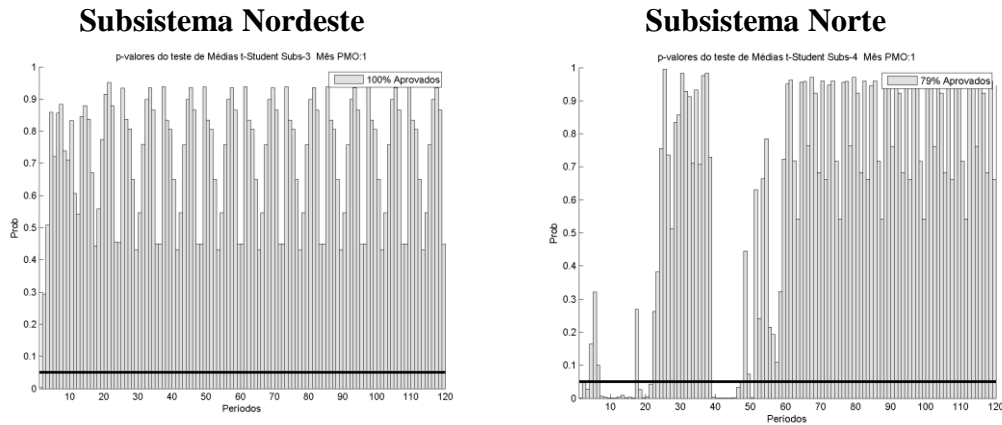


Figura 4.33 – P-valores do Teste t

4.2.2.2

Teste de Variância dos cenários

Analogamente ao realizado para os testes de médias, o Teste de Levene (Casella & Berger, 2001) verifica a igualdade de variância. Conforme pode ser observado a seguir, as variâncias históricas são estatisticamente iguais às variâncias dos cenários para a maior parte dos períodos de todas as configurações ((Figura 4.34) e Figura 4.35)).

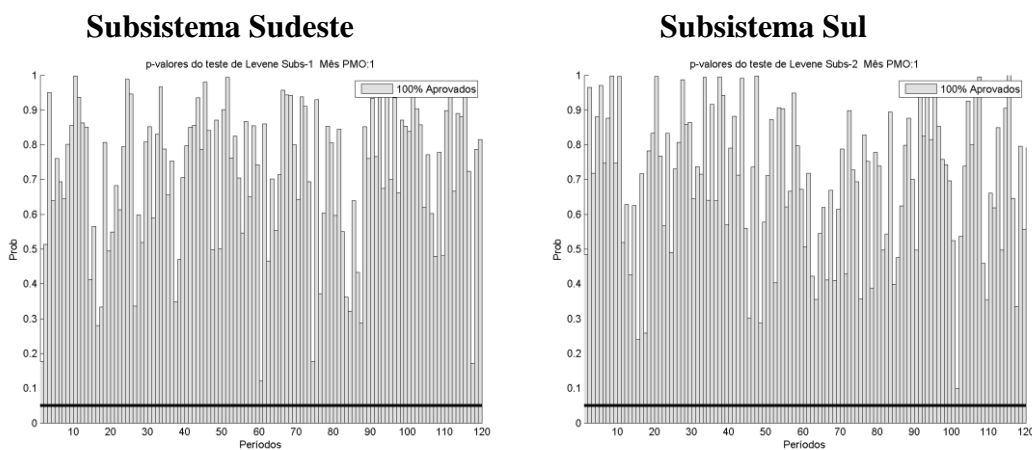


Figura 4.34 – P-valores do Teste de Levene

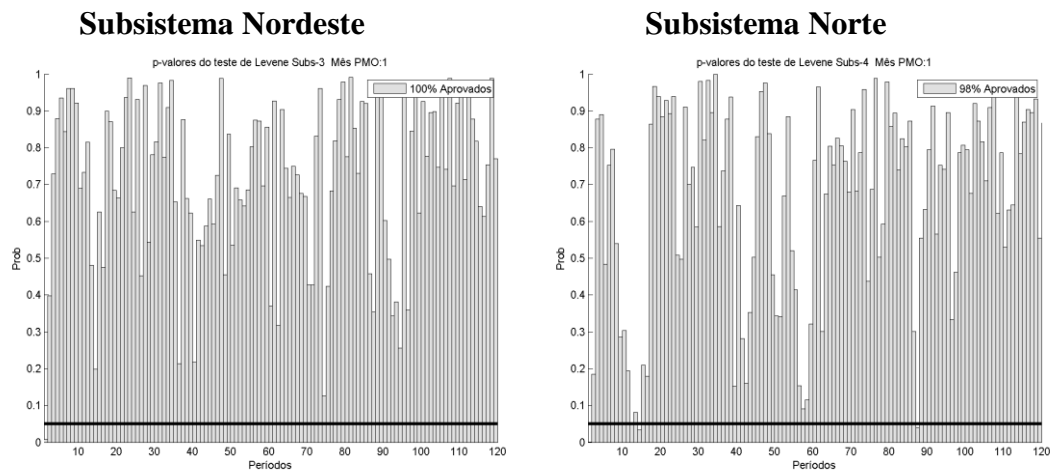


Figura 4.35 – P-valores do Teste de Levene

Assim, com relação aos testes mostrados nessa seção, observou-se que, segundo estes critérios, a estimação do modelo PAR(IC) e a geração de cenários via lognormal conforme implementada pelo SEB, gera “bons” cenários de afluência. Em outras palavras, os cenários gerados convergem para o histórico quando observa-se alguns momentos amostrais, tais como a média, a variância e a distribuição dos dados.

Para encerrar as discussões em torno do modelo PAR(p) interconfigurações, é importante observar que todos os testes e discussões feitos nessa seção são importantes requisitos da formulação Box & Jenkins e não se pode justificar imposições à modelagem para atender restrições ligadas à PDDE. Como observado ao longo dessa tese, o MGEA está intimamente ligado aos três pilares que sustentam o Setor Elétrico Brasileiro e “manipulações” sem justificativa teórica e matemática em seus modelos podem gerar importantes “perdas” ao bom planejamento da expansão e operação e na determinação do PLD.

Dessa forma, dado as dimensões do setor e a relação que a parte estocástica tem com os três pilares do SEB (capítulo 2), faz-se necessário, como feito nessa seção, observar os riscos que uma modelagem desse tipo gera e, mais ainda, permitir que outros modelos, que se utilizam de outras técnicas possam fornecer a PDDE novas séries de energia e novos modelos de afluência, minimizando-se dessa maneira, possíveis problemas como o ocorrido em junho de 2010.

Por fim, acredita-se que este capítulo tenha cumprido os três principais objetivos almejados, isto é, apresentar de maneira clara e didática o modelo

PAR(p) interconfigurações utilizado pelo SEB, levantar e discutir possíveis pontos que precisam ser melhores esclarecidos com relação a modelagem de séries temporais e a geração de séries sintéticas e fornecer informações empíricas acerca dos modelos e dos cenários para que os mesmos possam ser comparados com a modelagem Gama. Dando sequência a este trabalho, no próximo capítulo são apresentados os modelos Gama, bem como seus resultados e pontos de atenção.

5

Processo Periódico Autorregressivo Gama

Conforme discutido nessa tese, as séries sintéticas de vazão/ENA são comumente utilizadas em estudos de planejamento e operação do sistema hidroelétrico. Com esse propósito, um grande número de modelos de séries temporais utilizando um modelo Autorregressivo gaussiano foi sugerido, tendo como precursor (Thomas & Fiering, 1962).

Segundo (Fernandez & Salas, 1986), séries de vazão anuais podem ser consideradas aproximadamente normais, no entanto, para séries com intervalo de tempo mais curto, como mensal, semanal e diária as características da normalidade tornam-se não verificáveis, principalmente devido ao problema da assimetria. Tal dificuldade já foi abordada por diferentes métodos e um grande número de transformações foi feita nas séries para aproximar-se da normal. A maioria desses modelos são autorregressivos de ordem 1 constante para toda série ou com parâmetros periódicos e assumem normalidade dos resíduos. No entanto, o problema da assimetria permanece.

Neste estudo, a modelagem da série histórica de energia com distribuição marginal assimétrica é feita utilizando modelos de séries temporais com distribuição Gama e estrutura autorregressiva. O modelo linear e não linear adotado é uma extensão para acomodar o comportamento sazonal da distribuição marginal e da função de autocorrelação. O procedimento será estimar os parâmetros do modelo para cada mês e gerar os ruídos conforme estabelecido em cada caso.

A proposta de modelos Gama lineares e não lineares é restritiva, mas complementar no tratamento do coeficiente periódico de assimetria e o modelo combinado é proposto para resolver esse problema. O novo modelo é periódico, não Gaussiano, não linear e apresenta a estrutura autorregressiva. A vantagem dessa modelagem é que ela pode ser aplicada diretamente em séries históricas periódicas assimétricas, como é o caso da série histórica de energia. Para o caso do SEB a modelagem e geração de séries sintéticas, assumindo distribuição gama,

propicia a não necessidade de redução nos parâmetros para evitar problemas no corte de Benders e prescinde a questão dos cenários.

Nesse contexto, apresentam-se uma classe de modelos desenvolvidos por (Lawrance & Lewis, 1981) e (Fernandez & Salas, 1986), inseridos no contexto de “interconfigurações”, que incorporam as propriedades de assimetria e correlação sem a necessidade de transformação da série. Os modelos podem ser: Gama aditivo (PAGAR(1)), Gama multiplicativo (PMGAR(1)) ou modelos que incorporam as propriedades das duas estruturas (PGAR(1)).

5.1

O modelo teórico

5.1.1

Modelo Periódico Autorregressivo Aditivo Gama (PAGAR(1))

Considere o processo periódico Autorregressivo de ordem 1, baseado na estrutura usada pelo SEB, dado pela (equação 5.1):

$$\left(\frac{Y_{r,m,c_m} - \mu_{m,c_m}}{\sigma_{m,c_m}} \right) = \phi_1^{m,c_m} \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right) + \varepsilon_{r,m,c_m} \quad m=1,2\dots s \quad (5.1)$$

onde:

Y_{r,m,c_m} Série temporal em um período t qualquer e
 $Y_{r,m,c_m} \sim \text{Gama}(\alpha_{m,c_m}, \beta_{m,c_m}, \lambda_{m,c_m})$

ε_{r,m,c_m} Série de ruídos independente e identicamente distribuída

Assumindo-se estacionariedade cíclica com distribuição Gama, o modelo proposto em (5.1) representa um Modelo Periódico Autorregressivo com distribuição Gama aditiva e função densidade dada por:

$$f_x(x) = \frac{\alpha^{(m,c_m)} \beta^{(m,c_m)} (x - \lambda^{(m,c_m)})^{\beta^{(m,c_m)} - 1} \exp[-\alpha^{(m,c_m)} (x - \lambda^{(m,c_m)})]}{\Gamma(\beta^{(m,c_m)})} \quad (5.2)$$

onde:

$\alpha^{(m,c_m)}$ parâmetro de escala calculado com base no mês m da configuração c_m .

$\beta^{(m,c_m)}$ parâmetro de forma calculado com base no mês m da configuração c_m .

$\lambda^{(m,c_m)}$ parâmetro de localização calculado com base no mês m da configuração c_m .

tal que:

$$\alpha^{(m,c_m)} > 0; \beta^{(m,c_m)} > 0 \text{ e } x > \lambda^{(m,c_m)}$$

Seguindo a ideia proposta por (Fernandez & Salas, 1990), para a (equação 5.1) ser definida é preciso encontrar uma distribuição para o resíduo ε_t , dado que Y_t segue uma distribuição Gama.

Pode ser mostrado que a transformação de Laplace-Stieltjes do processo de erro, baseada na estrutura delineada pelo SEB, é dada por:

$$LS_{\varepsilon_t}(s) = e^{(\lambda^{(m,c_m)} - \phi_1^{(m,c_m)} \lambda^{(m-1,c_{m-1})})s} \left(\frac{1}{1 + \alpha^{(m,c_m)} s} \right)^{\beta^{(m,c_m)} - \beta^{(m-1,c_{m-1})}} \cdot \left[\phi_1^{(m,c_m)} \frac{\alpha^{(m-1,c_{m-1})}}{\alpha^{(m,c_m)}} + \left(1 - \frac{\alpha^{(m-1,c_{m-1})}}{\alpha^{(m,c_m)}} \right) \frac{1}{1 + \alpha^{(m,c_m)} s} \right]^{\beta^{(m-1,c_{m-1})}} \quad (5.3)$$

Lembrando que, para fins de simplificar a notação assumiu-se ε_{m,c_m} ao invés de ε_{r,m,c_m} . Para encontrar a inversa de (5.3) têm-se as seguintes restrições:

$$\beta^{(m,c_m)} \geq \beta^{(m-1,c_{m-1})} > 0 \quad (5.4a)$$

$$0 \leq \phi_1^{(m,c_m)} \frac{\alpha^{(m-1,c_{m-1})}}{\alpha^{(m,c_m)}} \leq 1 \quad (5.4b)$$

Utilizando as propriedades da transformação de Laplace-Stieltjes, ε_{m,c_m} pode ser dado como a soma de duas variáveis aleatórias:

$$\varepsilon_{m,c_m} = \varepsilon_{m,c_m}(0) + \varepsilon_{m,c_m}(1) \quad (5.5)$$

onde:

$$\varepsilon_{m,c_m}(0) \sim \text{Gama}(\alpha^{m,c_m}, \beta^{m,c_m} - \beta^{m-1,c_{m-1}}, \lambda^{m,c_m} - \phi_1^{(m,c_m)} \lambda^{m-1,c_{m-1}}) \quad (5.5a)$$

$$\varepsilon_{m,c_m}(1) = 0 \quad \forall N = 0 \quad (5.5b)$$

$$\varepsilon_{m,c_m}(1) = \sum_{j=1}^N X_j \left(\phi_1^{(m,c_{m_i})} \frac{\alpha^{(m-1,c_{m-1})}}{\alpha^{(m,c_m)}} \right)^{U_N} \quad \forall N > 0 \quad (5.5c)$$

Onde N é uma variável aleatória Poisson com valor esperado θ dado por:

$$\theta^{m,c_m} = -\beta^{m-1,c_{m-1}} \ln \left(\phi_1^{m,c_m} \frac{\alpha^{m-1,c_{m-1}}}{\alpha^{m,c_m}} \right) \quad (5.5d)$$

Em adição, a sequência U_1, U_2, \dots, U_N é formada por variáveis aleatórias independentes e identicamente distribuída com distribuição Uniforme (0,1) e a sequência X_1, X_2, \dots, X_N é formada por variáveis aleatórias independentes e identicamente distribuída com distribuição Exponencial (α^{m,c_m}).

Como acontece nos modelos PAR(p) tradicionais, a estrutura de autocorrelação dos modelos PAGAR(1) é obtida multiplicando ambos os lados de (5.5) por $Y_{r,m-k,c_{m-k}}$ e tomando o valor esperado. Assim, tem-se:

$$\rho_k^{m,c_m} = \frac{\sigma^{m-k,c_{m-k}}}{\sigma^{m,c_m}} \prod_{j=0}^{k-1} \phi^{m-j,c_{m-j}} \quad k = 1, 2, \dots \quad (5.6a)$$

onde:

ρ_k^{m,c_m} representa a correlação entre o mês m e o mês $m-k$ da configuração c_m .

É importante observar que para $k = 1$, a estrutura de autocorrelação será idêntica à obtida para o modelo “PAR(p) Interconfigurações” que pressupõe distribuição Normal, como pode ser observado em (5.6b).

$$\rho_1^{m,c_m} = \frac{\sigma^{m-1,c_{m-1}}}{\sigma^{m,c_m}} \phi^{m,c_m} \quad (5.6b)$$

Ao usar a expressão (5.6b) e expressar a variância e o coeficiente de assimetria em termos dos parâmetros do modelo, as restrições (5.4a) e (5.4b) podem ser reduzidas para:

$$\left(\rho_1^{m,c_m} \right)^2 \leq \left(\beta^{m-1,c_{m-1}} / \beta^{m,c_m} \right) \leq 1 \quad (5.7a)$$

$$\left(\rho_1^{m,c_m} \right) \leq \left(\gamma^{m,c_m} / \gamma^{m-1,c_{m-1}} \right) \leq 1 \quad (5.7b)$$

onde:

γ^{m,c_m} representa o coeficiente de assimetria do mês m da configuração c_m .

As equações (5.7a) e (5.7b) representam a região de admissão dos modelos em função do coeficiente de autocorrelação periódico de *lag* 1. Após descrever o Modelo Periódico Autorregressivo Gama Aditivo de *lag* 1, torna-se relevante tecer alguns comentários sobre o modelo e suas restrições. Conforme pode ser verificado na (equação 5.4a), o parâmetro de forma ou é constante ou aumenta de um mês para o outro. Se β é constante, pode ser mostrado que $\varepsilon_{m,c_m}(0)$ de (5.5) torna-se igual a $\lambda^{m,c_m} - \phi^{m,c_m} \lambda^{m-1,c_{m-1}}$. Nesse caso, as restrições (5.4b) e (5.6b) conduzem para a restrição $0 \leq \rho_1^{m,c_m} \leq 1$. Portanto, o modelo PAGAR(1) terá média, variância e autocorrelação periódica, mas assimetria constante.

Ainda, para um modelo mais simples que assume $\beta^{m,c_m} = 1$, para todo mês m , a distribuição marginal será uma $Exp(\alpha^{m,c_m}, \lambda^{m,c_m})$. Por isso, o modelo PAGAR(1) sozinho com distribuição Exponencial ou Gama, assume que todos os parâmetros são periódicos, exceto o parâmetro de forma. Segundo (Fernandez & Salas, 1990), esse modelo pode ser útil para modelar séries hidrológicas, nas quais a assimetria não varia de mês para mês e são as mesmas durante o ano.

Por outro lado, considerando que β varia nos meses, é possível observar que a restrição (5.7a) implica que $\beta_m > \beta_{m-1} > \dots > \beta_2 > \beta_1$, mas a condição $\beta_1 > \beta_m$ não é possível, o que implica que a correlação entre $Y_{r,1,c_1}$ e Y_{r,s,c_s} não se mantém.

Portanto, o modelo (5.1) sob a restrição (5.4a) apresenta uma sequência Gama com média, variância e correlação periódica e assimetria decrescente. Mas, a correlação entre o último mês do ano e o primeiro é perdida, o que implica que as sequências anuais derivadas de cada um dos modelos são descorrelatadas.

Dessa forma, o modelo PAGAR(1) somente é aplicável para um processo Gama com o parâmetro de forma constante através dos anos, o que implica assimetria constante, enquanto os demais momentos amostrais podem ser constantes ou periódicos. Vale lembrar que a solução para essa limitação será tratada no decorrer deste capítulo.

5.1.2

Modelo Periódico Autorregressivo Multiplicativo Gama (PMGAR(1))

O processo autorregressivo multiplicativo é um processo não linear. Este tipo de modelo é usado, de maneira implícita em hidrologia, quando uma transformação logarítmica é feita na série temporal. Por exemplo, se for aplicada uma transformação logarítmica em um modelo linear gaussiano, as variáveis assumem comportamento lognormal, então, implicitamente, um processo autorregressivo multiplicativo, pode ser pensado como um processo linear no domínio log. Contudo, conforme destacado por (Fernandez, 1984) esse processo multiplicativo implícito, tão comumente aplicado em hidrologia, é limitado, uma vez que não há representação matemática formal do processo de dependência, especialmente em relação à distribuição marginal requerida em cada estrutura de dependência. Neste contexto, o objetivo dessa seção é formalizar o processo autorregressivo periódico com distribuição marginal Gama.

O Modelo Periódico Autorregressivo Multiplicativo Gama – PMGAR(1), pode ser escrito em sua forma intercorrelacionada como:

$$\left(\frac{Y_{r,m,c_m} - \mu_{m,c_m}}{\sigma_{m,c_m}} \right) = \left(\frac{Y_{r,m-1,c_{m-1}} - \mu_{m-1,c_{m-1}}}{\sigma_{m-1,c_{m-1}}} \right)^{\delta_1^{m,c_m}} \cdot V_{r,m,c_m} \quad m=1,2\dots s \quad (5.8)$$

onde:

δ_1^{m,c_m} parâmetro periódico do processo autorregressivo

V_{r,m,c_m} Série de ruídos independente e identicamente distribuída

Objetivando encontrar a distribuição do erro apresentado em (5.8), faz-se uso da transformação de Mellin. (Fernandez, 1984) demonstrou que a transformação do resíduo para este tipo de modelo segue a formulação apresentada em (5.9), a qual é apresentada, sem perda de generalidade sob os moldes da configuração intercorrelacionada.

$$M_{V_{m,c_m}}(s) = E(V_{m,c_m}^{s-1}) = \left(\frac{\alpha^{m,c_m}}{(\alpha^{m,c_m})^{\delta^{m,c_m}}} \right)^{s-1} \cdot \left(\frac{\Gamma(\beta^{m-1,c_{m-1}})\Gamma(\beta^{m,c_m} + s - 1)}{\Gamma(\beta^{m,c_m})\Gamma(\beta^{m-1,c_{m-1}} + \delta^{m,c_m}s - \delta^{m,c_m})} \right) \quad (5.9)$$

Assumindo-se que foi utilizado V_{m,c_m} ao invés de V_{r,m,c_m} por questões de notação, a inversa da (equação 5.9) pode ser feita, dada as seguintes restrições:

$$0 < \beta^{m,c_m} \leq \beta^{m-1,c_{m-1}} \quad (5.10a)$$

$$0 \leq \delta^{m,c_m} \leq 1 \quad (5.10b)$$

Segue que o ruído V_{m,c_m} pode ser escrito como:

$$V_{r,m,c_m} = \frac{\alpha^{m,c_m}}{(\alpha^{m-1,c_{m-1}})^{\delta^{m,c_m}}} (Z^{m,c_m})^{\delta^{m,c_m}} W^{m,c_m} \quad (5.11)$$

onde:

$$Z^{m,c_m} \quad Z^{m,c_m} \sim \text{Beta}(\beta^{m,c_m}, \beta^{m-1,c_{m-1}} - \beta^{m,c_m})$$

W^{m,c_m} é o ruído de um modelo estacionário multiplicativo Gama, definido por (Mckenzie, 1982).

O ruído W^{m,c_m} pode ser simulado usando a propriedade de decomposição do logaritmo de uma variável Gama. (Fernandez, 1984) mostra que W^{m,c_m} pode ser simulado como:

$$W^{m,c_m} = \exp\left(V'_{m,c_m} + \lim_{\kappa \rightarrow \infty} \sum_{\kappa=1}^{\kappa} V_{\kappa,m,c_m}\right) \quad (5.12)$$

Seguindo as seguintes regras de probabilidade:

$$V'_{m,c_m} = -\gamma(1 - \delta^{m,c_m}) \quad \text{com probabilidade } \delta^{m,c_m} \quad (5.13a)$$

$$V'_{m,c_m} = -\gamma(1 - \delta^{m,c_m}) - \text{Ex}(\beta^{m,c_m}) \quad \text{com probabilidade } (1 - \delta^{m,c_m})$$

$$V_{m,c_m} = (1 - \delta^{m,c_m}) / \kappa \quad \text{com probabilidade } \delta^{m,c_m} \quad (5.13b)$$

$$V_{m,c_m} = \frac{(1 - \delta^{m,c_m})}{\kappa - Ex(\beta^{m,c_m} + \kappa)} \quad \text{com probabilidade } (1 - \delta^{m,c_m})$$

onde:

$Ex(a)$ representa uma variável aleatória exponencial. $Ex(a) = 1/a$

γ $\gamma = 0,577216 \dots$ (Constante de Euler)

Por fim, é importante ressaltar que as equações (5.13a) e (5.13b) convergem quase certamente, e, na prática, não é necessário um valor de κ muito grande para obter uma boa representação de Z^{m,c_m} (Fernandez, 1984).

A função de autocorrelação do modelo PMGAR(1), definida demonstrada por (Fernandez & Salas, 1986), pode ser descrita, em sua forma intercorrelacionada, como se segue:

$$\rho_k^{m,c_m} = \left(\frac{\beta^{m,c_m}}{\beta^{m-k,c_{m-k}}} \right)^{1/2} \prod_{j=0}^{k-1} \delta^{m-j,c_{m-j}} \quad k = 1, 2, 3, \dots \quad (5.14a)$$

Para o caso de $k = 1$, tem-se:

$$\rho_1^{m,c_m} = \left(\frac{\beta^{m,c_m}}{\beta^{m-1,c_{m-1}}} \right)^{1/2} \delta^{m,c_m} \quad (5.14b)$$

Segundo (McKenzie, 1982), a (equação 5.14a) é válida somente quando a série histórica tem um comportamento Gama. Essa consideração reafirma o fato de que quando a série histórica é lognormalmente distribuída e um processo AR(1) é aplicado para o log dessa série, a estrutura de autocorrelação da série temporal original não é a mesma que para uma série com a transformação log. É importante observar também que, se um processo $\{X\}$ é assumido Gama, o processo $Y = \log(X)$ não é loggama distribuído.

Dando continuidade, assumindo que os parâmetros do modelo são escritos em termos do coeficiente de assimetria periódico e da autocorrelação de lag 1, apresentado em (5.14b), as restrições apresentadas em (5.10a) e (5.10b) podem ser reescritas como:

$$1 \leq \frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \leq \frac{1}{(\rho_1^{m,c_m})^2} \quad (5.15a)$$

ou:

$$1 \leq \frac{\gamma^{m,c_m}}{\gamma^{m-1,c_{m-1}}} \leq \frac{1}{\rho_1^{m,c_m}} \quad (5.15b)$$

Como no caso do modelo PAGAR(1), as equações (5.15a) e (5.15b) representam a região de admissão dos modelos em função do coeficiente de autocorrelação periódico de *lag* 1.

É importante analisar quais são as implicações das restrições apresentadas acima, isto é, quais são as limitações que as mesmas trazem para os parâmetros do modelo.

De acordo com (Fernandez & Salas, 1986) há basicamente duas importantes implicações: a restrição (5.10a) relacionada ao parâmetro β pode variar ou manter-se constante ao longo dos anos. Para o caso do parâmetro β constante, similar ao que acontece para o modelo PAGAR(1), a transformação (5.9) e a restrição (5.10a) conduzem para a restrição $0 \leq \rho_1^{m,c_m} \leq 1$, fazendo com que o PMGAR(1) tenha média, variância e autocorrelação periódica, mas assimetria constante.

Considerando β não constante, é possível observar que a restrição (5.10a) implica que $\beta_m < \beta_{m-1} < \dots < \beta_2 < \beta_1$, mas a condição $\beta_1 < \beta_m$ não é possível, o que implica, como no caso do modelo PAGAR(1), que a correlação entre $Y_{r,1,c_1}$ e Y_{r,s,c_s} não se mantém. Portanto, o modelo (5.8) sob a restrição (5.10a) apresenta uma sequência Gama com média, variância e correlação periódica e assimetria decrescente. Mas a correlação entre o último mês do ano e o primeiro mês do ano seguinte é perdida, o que implica que as sequências anuais derivadas de cada um dos modelos são descorrelatadas.

Assim, considerando que é desejável a variação de β e uma vez que as restrições (5.4a – Modelo PAGAR(1)) e (5.10a – Modelo PMGAR(1)) são opostas, (Fernandez & Salas, 1986) propuseram um modelo misto que pudesse capturar as vantagens de ambas as metodologias, o modelo PGAR(1) – Periódico Gama Autorregressivo. Tal modelo, apresentado a seguir, é mais adequado para representar fenômenos hidrológicos, uma vez que reproduzirá um processo Gama, com média, variância, autocorrelação e assimetria periódica.

5.1.3

Modelo Periódico Autorregressivo Gama (PGAR(1))

As duas seções anteriores mostraram que os modelos autorregressivos Gama podem ser ajustados considerando parâmetros periódicos, isto é, eles podem acomodar a estrutura periódica tanto na distribuição marginal quanto na autocorrelação respeitando certas restrições.

Na verdade, o processo periódico Gama aditivo pode ser propriamente definido somente se o parâmetro de forma da distribuição Gama for constante ou aumentar, dentro de um certo limite, ao longo do tempo. Por outro lado, o processo periódico Gama multiplicativo foi desenvolvido para um parâmetro de forma constante ou quando o mesmo decresce ao longo do tempo. Dessa forma, segundo (Fernandez, 1984), tais modelos são complementares mais que competitivos. Então, nesse contexto, a ideia foi desenvolver um modelo híbrido que poderia acomodar simultaneamente as condições impostas por ambos os modelos.

5.1.3.1

Definição

Seja Q_{r,m,c_m} um processo periódico autorregressivo assimétrico, com distribuição marginal Gama com três parâmetros, $Gama(\alpha^{m,c_m}, \beta^{m,c_m}, \lambda^{m,c_m})$, como apresentada em (5.2), tem-se:

$$Y_{r,m,c_m} = Q_{r,m,c_m} - \lambda^{m,c_m} \quad (5.16)$$

Assim, assume-se uma Gama com dois parâmetros, $Gama(\alpha^{m,c_m}, \beta^{m,c_m})$. Para modelar o processo Y_{r,m,c_m} , propõem-se o seguinte modelo:

$$Q_{r,m,c_m} = \lambda^{m,c_m} + \phi_1^{m,c_m} Y_{r,m-1,c_{m-1}} + (Y_{r,m-1,c_{m-1}})^{\delta^{m,c_m}} \xi_{r,m,c_m} \quad (5.17)$$

onde:

ξ_{r,m,c_m} é o ruído do processo

Conforme se pode observar o modelo (5.17) é não linear e não gaussiano, representando uma mistura entre os modelos PAGAR(1) e PMGAR(1). Tal estrutura foi denominada por (Fernandez & Salas, 1986) como modelo PGAR(1).

Para utilizar ambas as estruturas os autores definiram um método em que os parâmetros ϕ_1^{m,c_m} e δ^{m,c_m} são definidos em função dos parâmetros α^{m,c_m} e β^{m,c_m} da distribuição marginal Gama e da autocorrelação de lag 1, ρ_1^{m,c_m} , como:

$$\begin{aligned} \phi^{m,c_m} = 0 & & \beta^{m,c_m} < \beta^{m-1,c_{m-1}} \\ \phi^{m,c_m} = \rho_1^{m,c_m} \frac{\alpha^{m,c_m}}{\alpha^{m-1,c_{m-1}}} \left(\frac{\beta^{m,c_m}}{\beta^{m-1,c_{m-1}}} \right)^{1/2} & & \beta^{m,c_m} \geq \beta^{m-1,c_{m-1}} \end{aligned} \quad (5.18)$$

$$\begin{aligned} \delta^{m,c_m} = \rho_1^{m,c_m} \left(\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \right)^{1/2} & & \beta^{m,c_m} < \beta^{m-1,c_{m-1}} \\ \delta^{m,c_m} = 0 & & \beta^{m,c_m} \geq \beta^{m-1,c_{m-1}} \end{aligned} \quad (5.19)$$

Seguindo a mesma lógica, o ruído, ξ_{r,m,c_m} , corresponde ao modelo aditivo ou multiplicativo, a depender do valor relativo do parâmetro de forma periódico, β^{m,c_m} (equação 5.20).

$$\begin{aligned} \xi_{r,m,c_m} = V_{r,m,c_m} & \quad \text{dado por (5.11)} & \beta^{m,c_m} < \beta^{m-1,c_{m-1}} \\ \xi_{r,m,c_m} = \varepsilon_{r,m,c_m} & \quad \text{dado por (5.5)} & \beta^{m,c_m} \geq \beta^{m-1,c_{m-1}} \end{aligned} \quad (5.20)$$

Quanto às restrições dos parâmetros do modelo, obtidas em (5.7) e (5.15), tem-se:

$$\begin{aligned} 0 < \delta^{m,c_m} < 1 & & \beta^{m,c_m} < \beta^{m-1,c_{m-1}} \\ 0 < \phi^{(m,c_m)} \frac{\alpha^{(m-1,c_{m-1})}}{\alpha^{(m,c_m)}} < 1 & & \beta^{m,c_m} \geq \beta^{m-1,c_{m-1}} \end{aligned} \quad (5.21)$$

Ao expressar os valores dos parâmetros $\phi^{(m,c_m)}$ e δ^{m,c_m} de acordo com as equações (5.20) e (5.21), as restrições podem ser simplificadas para as seguintes equações:

$$\left(\rho_1^{m,c_m} \right)^2 \leq \left(\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \right) \leq \left(\rho_1^{m,c_m} \right)^{-2} \quad (5.22a)$$

ou:

$$\rho_1^{m,c_m} \leq \frac{\gamma^{m,c_m}}{\gamma^{m-1,c_{m-1}}} \leq \frac{1}{\rho_1^{m,c_m}} \quad (5.22b)$$

As expressões (5.22a) e (5.22b) definem o espaço paramétrico do modelo PGAR(1) como função da autocorrelação ρ_1^{m,c_m} , a qual assumirá valores positivos. Uma vez que as imposições acima são válidas, não há outras restrições aos parâmetros de escala e forma impostos pela distribuição marginal Gama.

Portanto, além do modelo PGAR(1) manter a estrutura autorregressiva, ele permite que média, a variância, a autocorrelação e a assimetria sejam periódicas. Conforme ressaltado por (Fernandez & Salas, 1986), essa deve ser a melhor forma de modelar séries geofísicas observadas de forma periódica.

5.1.3.2

Estimação dos parâmetros dos modelos

O conjunto de parâmetros descritos no modelo PGAR (1) são $\{\lambda^{m,c_m}, \alpha^{m,c_m}, \beta^{m,c_m}, \phi_1^{m,c_m}, \delta^{m,c_m}\}$, os quais envolvem parâmetros relacionados com a distribuição marginal e os parâmetros autorregressivos aditivos e multiplicativos. Como os parâmetros ϕ_1^{m,c_m} e δ^{m,c_m} são definidos por (5.18) e (5.19), respectivamente, o conjunto de parâmetros a serem estimados reduz-se para $\{\rho_1^{m,c_m}, \lambda^{m,c_m}, \alpha^{m,c_m}, \beta^{m,c_m}\}$.

Os valores de $\lambda^{m,c_m}, \alpha^{m,c_m}$ e β^{m,c_m} são os parâmetros da distribuição marginal de Q_{r,m,c_m} , a qual assume uma Gama com três parâmetros ($Gama(\alpha^{m,c_m}, \beta^{m,c_m}, \lambda^{m,c_m})$). Dado que os valores da série $Q_{r,m,c_m} = \{Q_{1,m,c_m}, Q_{2,m,c_m}, \dots, Q_{N,m,c_m}\}$, onde N é o número de anos, estão separados por um ano, é razoável imaginar que são aproximadamente não correlacionados, sendo possível estimar os parâmetros pelo método de Máxima Verossimilhança (MV). Contudo, embora algumas práticas e procedimentos eficientes estejam disponíveis para obter os parâmetros da distribuição Gama através do método de MV, tal estimador não tem solução se o parâmetro β^{m,c_m} for menor que 1, o que

implica que o coeficiente de assimetria não pode ultrapassar 2, conforme observado por (Kite, 1977) e (Salas & Smith, 1981).

Entretanto, não é incomum séries de vazão/ENA com assimetria maior que 2, tornando-se limitado o uso do estimador de MV para os modelos descritos nesse trabalho. Embora o estimador por Métodos dos Momentos, em geral, não tenha propriedades tão boas quanto o estimador de MV, neste caso ele é bastante apropriado, uma vez que pode ser aplicado para todo intervalo de valores da densidade Gama. Os estimadores por Método dos Momentos são:

$$\hat{\lambda}^{m,c_m} = \bar{Q}^{m,c_m} - 2S^{m,c_m} / g^{m,c_m} \quad (5.23)$$

$$\hat{\alpha}^{m,c_m} = S^{m,c_m} g^{m,c_m} / 2 \quad (5.24)$$

$$\hat{\beta}^{m,c_m} = 4 / g^{m,c_m} \quad (5.25)$$

$$\hat{\rho}_1^{m,c_m} = r_1^{m,c_m} \quad (5.26)$$

5.2

Análise Empírica

5.2.1

A série de ENA e a distribuição Gama

Conforme comentado ao longo desse trabalho e destacado por (Fernandez & Salas, 1986), séries temporais de vazão/energia apresentam um comportamento próximo à distribuição Gama. Mesmo com a afirmação dos autores e de outros trabalhos que corroboraram tal afirmativa, é preciso testar empiricamente como as séries temporais de energia, base de dados foco deste trabalho, se comportam em termos de distribuição, justificando assim, a adoção ou não de modelos que tem como pressupostos a distribuição Gama da série histórica.

Para atingir o objetivo explicitado acima, foram feitos dois procedimentos de ajuste à série histórica de energia. O primeiro procedimento, mais exploratório, foi plotar em um mesmo gráfico o histograma das séries históricas de energia e as

funções densidade de probabilidade (FDP) da distribuição Gama e Normal ajustadas aos dados e comparar a adequação do ajuste das distribuições aos dados.

O segundo procedimento, com validade estatística, constitui em verificar a adequação do ajuste da distribuição Gama aos dados observados. Para tal utilizou-se o Teste Qui-Quadrado (χ^2), como proposto por (Triola, 2005), no qual, a hipótese nula é que os dados seguem uma Distribuição Gama. Para atender a dependência serial, ajustou-se uma distribuição Gama para cada mês do ano.

Com relação ao subsistema Sudeste/Centro-Oeste, conforme observa-se na figura 5.1, os dados de energia ajustam-se melhor a FDP da Gama. Claramente observa-se que a questão da assimetria das séries de vazão/ENA influencia fortemente nesse ajuste.

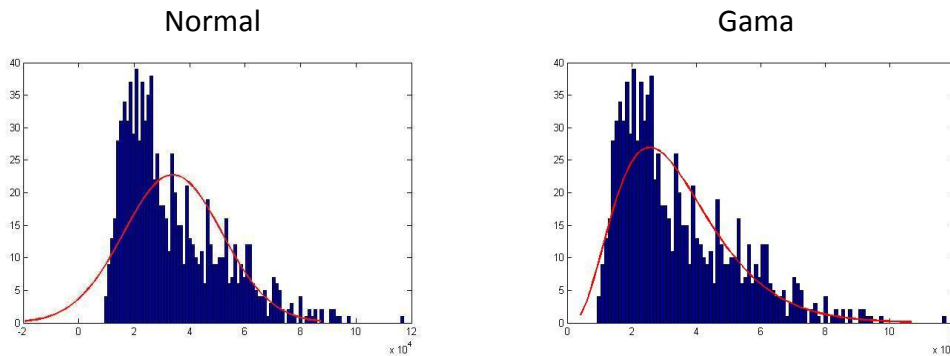


Figura 5.1 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Sudeste/Centro-Oeste

Com relação ao Teste Qui-Quadrado, verifica-se, com 95% de confiança, que apenas no mês de junho a hipótese nula de que os dados seguem uma distribuição Gama foi rejeitada (Figura 5.2).

Ambos os procedimentos reforçam o argumento de que as séries de energia seguem uma distribuição Gama e que, portanto, podem ser modeladas por modelos que pressupõem tal distribuição.

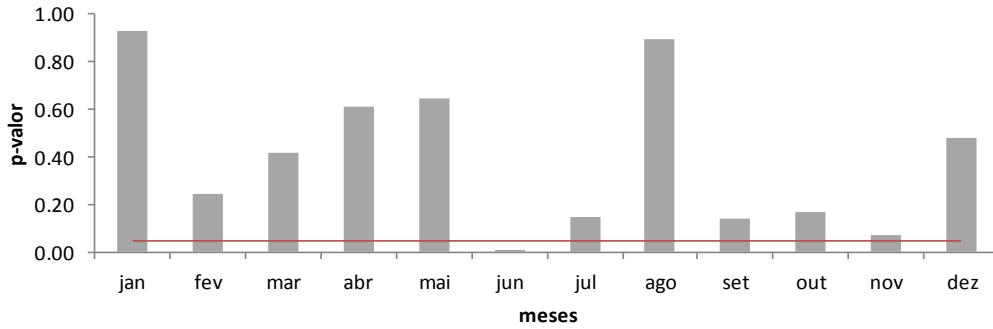


Figura 5.2 – Teste do Qui-Quadrado (χ^2) para série de ENA mensal do subsistema Sudeste/Centro-Oeste

Quanto ao subsistema Sul, a análise é muito parecida com a feita para o subsistema Sudeste/Centro-Oeste, isto é, a série histórica se ajusta perfeitamente à função densidade de probabilidade da distribuição Gama (Figura 5.3) e, conforme se observa na figura 5.4, a hipótese nula de que os dados seguem uma distribuição Gama não foi rejeitada em nenhum dos meses.

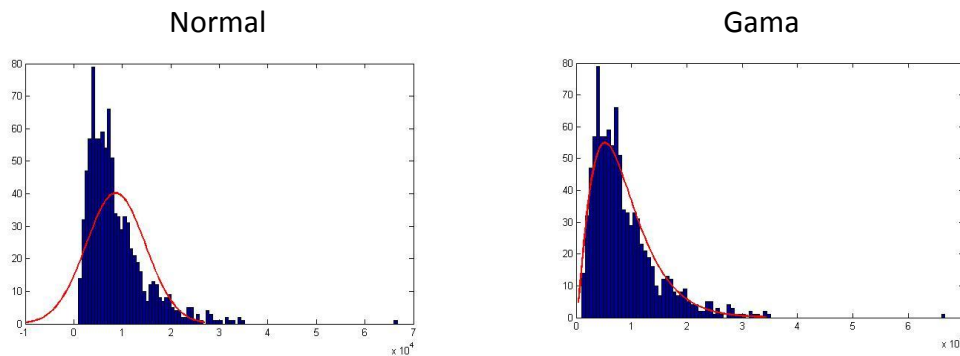


Figura 5.3 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Sul

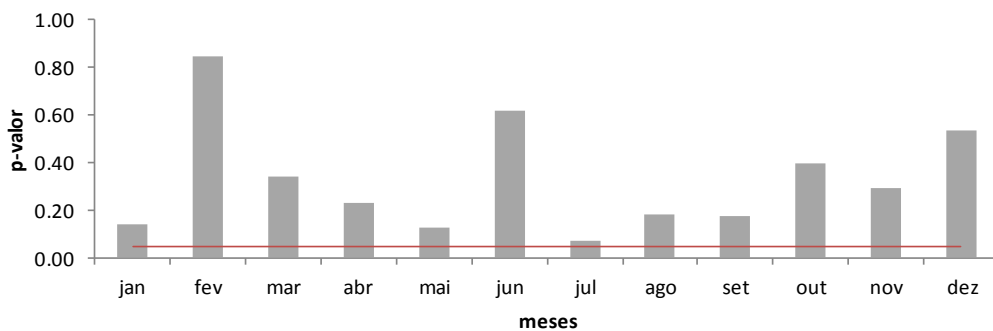


Figura 5.4 – Teste do Qui-Quadrado (χ^2) para série de ENA mensal do subsistema Sul

Para o subsistema Nordeste, quando se analisa o histograma da série histórica com sobreposição da FDP da distribuição Gama e Normal, se observa que os dados de energia se adequam mais à Gama que a Normal (Figura 5.5), como aconteceu nos subsistemas Sudeste/Centro-Oeste e Sul. No entanto, ao testar a série de dados à distribuição Gama (Figura 5.6), verifica-se a rejeição da hipótese nula em seis meses dos doze testados. Tal resultado não invalida a utilização dos modelos Gama a essa séries de dados, além disso, se o nível de confiança for flexibilizado para 90%, todos os testes não rejeitam a hipótese nula.

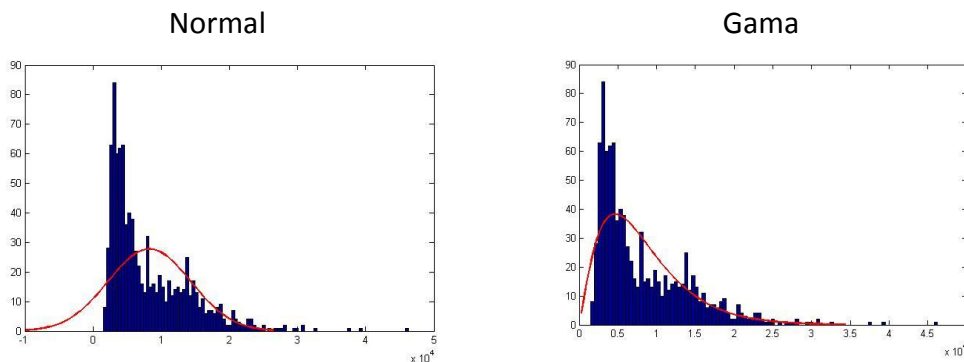


Figura 5.5 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Nordeste

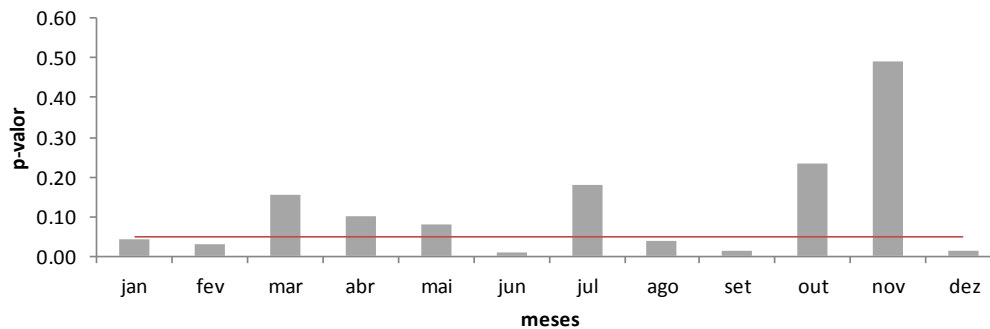


Figura 5.6 – Teste do Qui-Quadrado (χ^2) para série de ENA mensal do subsistema Nordeste

Por fim, ao observar o subsistema Norte, verifica-se novamente o bom ajuste dos dados à distribuição Gama (Figura 5.7) e, mais ainda, o Teste Qui-Quadrado mostra que a hipótese nula de Distribuição Gama não é rejeitada para todos os meses deste subsistema.

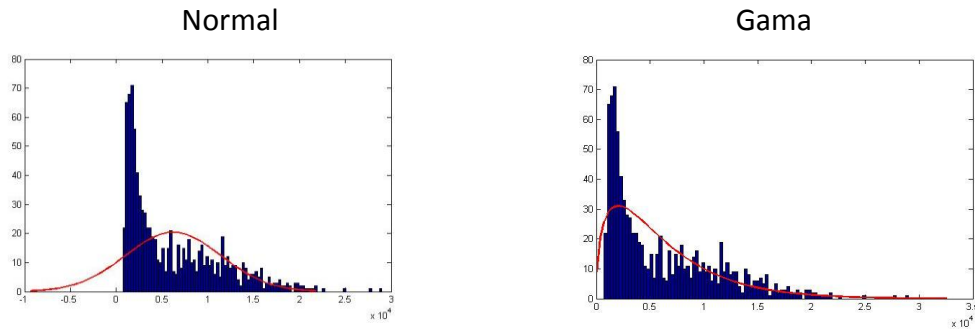


Figura 5.7 – Histograma da série histórica com sobreposição da FDP da distribuição Normal e Gama – Subsistema Norte

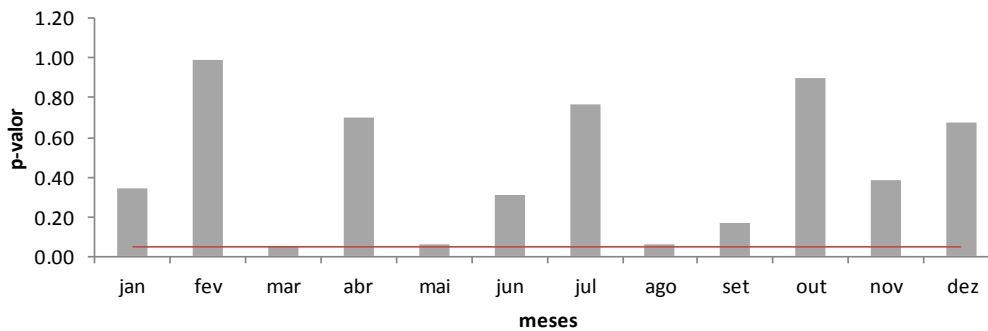


Figura 5.8 – Teste do Qui-Quadrado (χ^2) para série de ENA mensal do subsistema Norte

5.2.2

A modelagem Gama periódica

Conforme observado, a seção (5.2.1) corroborou os argumentos descritos nessa tese e salientados por (Fernandez & Salas, 1986) de que séries de ENAs aproximam-se muito mais da distribuição Gama que da distribuição Normal, justificando o uso de modelos e a estimação de parâmetros que não dependam da normalidade dos dados.

Conforme também observado nesse trabalho, (Fernandez & Salas, 1986) propuseram o modelo PGAR(1) que prescinde da normalidade e que, segundo os autores deve ser aplicado a séries históricas que se aproximam da distribuição Gama, como é o caso das séries de ENAs. Outro ponto defendido pelos autores é a questão da formulação matemática, a qual foi bem definida para os modelos Gama, não sendo o caso, por exemplo, de modelos que utilizam a distribuição lognormal para simular as séries sintéticas.

Tal arcabouço matemático desenvolvido pelos autores, mostrou que o modelo PGAR(1) só é válido quando a restrição (5.22a), a qual condiciona uma área de admissão aos parâmetros β^{m,c_m} , é atendida. Dessa forma, antes de iniciar a modelagem da série histórica de ENA, testou-se tal restrição imposta pelo modelo. Como resultado, verificou-se que as séries históricas de energia, que compreendem o período de 1931 à 2009, não se mostraram adequadas ao ajuste do modelo periódico autorregressivo Gama proposto por (Fernandez & Salas, 1986).

Para testar tal restrição adotou-se o seguinte procedimento: para cada mês calculou-se $\beta^{m-1,c_{m-1}} / \beta^{m,c_m}$ e o valor de ρ_1^{m,c_m} , em seguida, plotou-se os mesmos em um gráfico de barras e linhas, onde as barras representam os valores $\beta^{m-1,c_{m-1}} / \beta^{m,c_m}$, a linha verde $(\rho_1^{m,c_m})^2$ e a linha azul $(\rho_1^{m,c_m})^{-2}$. Dessa forma, a restrição (5.22a) é atendida quando a barra vermelha está entre as linhas verde e azul.

Analisando-se as figuras (Figura 5.9, Figura 5.10, Figura 5.11 e Figura 5.12), observa-se, para o subsistema Sudeste/Centro-Oeste que nos meses de maio, junho, julho, agosto, setembro e dezembro a restrição não foi atendida. Já ao analisar o subsistema Sul (Figura 5.10), verifica-se que a restrição não foi atendida, marginalmente, apenas no mês de julho. No subsistema Nordeste (Figura 5.11) a restrição (5.22a) foi respeitada apenas nos meses de janeiro, fevereiro, novembro e dezembro. E no subsistema Norte, o parâmetro não está na região de admissão apenas no mês de agosto. Dessa forma, a julgar pela restrição (5.22a) pode-se dizer que nenhum dos subsistemas se adequa à mesma, no entanto, observou-se também que nos subsistemas Sul e Norte tal violação é marginal.

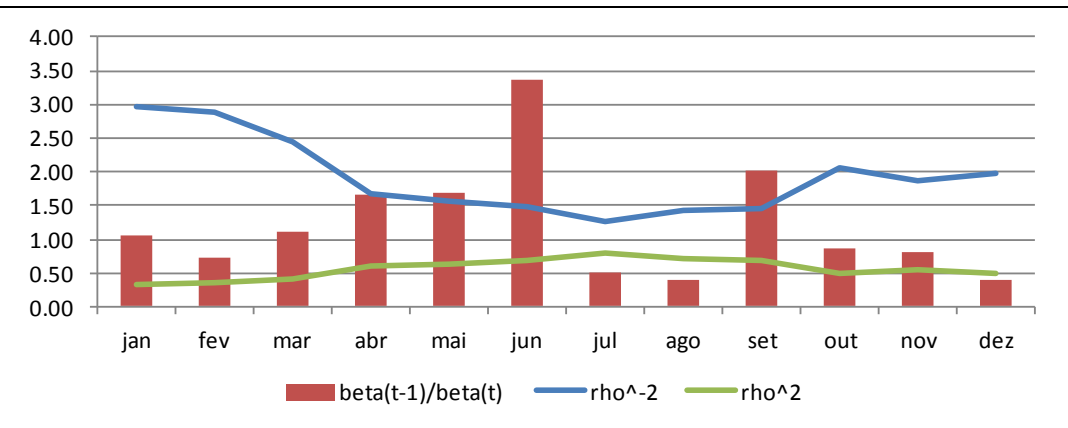


Figura 5.9 – Estudo da restrição (5.22a) para a série histórica do Subsistema Sudeste/Centro-Oeste

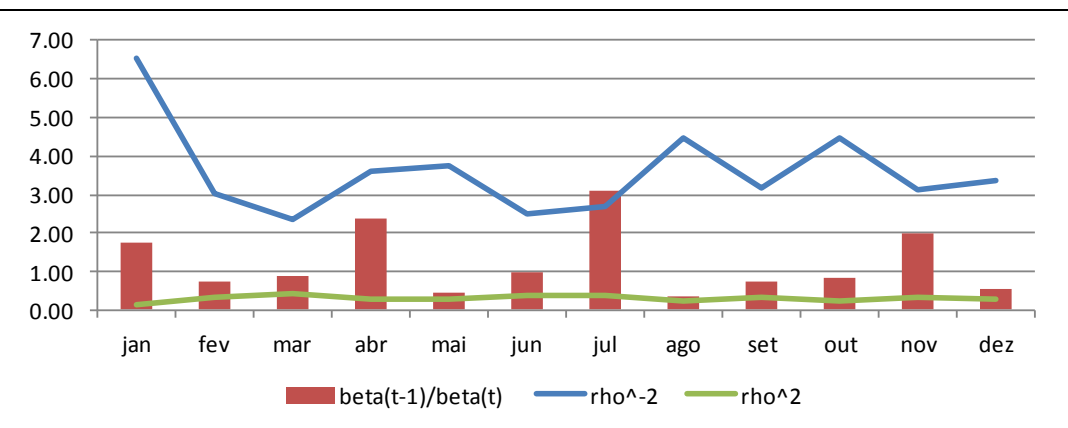


Figura 5.10 – Estudo da restrição (5.22a) para a série histórica do Subsistema Sul

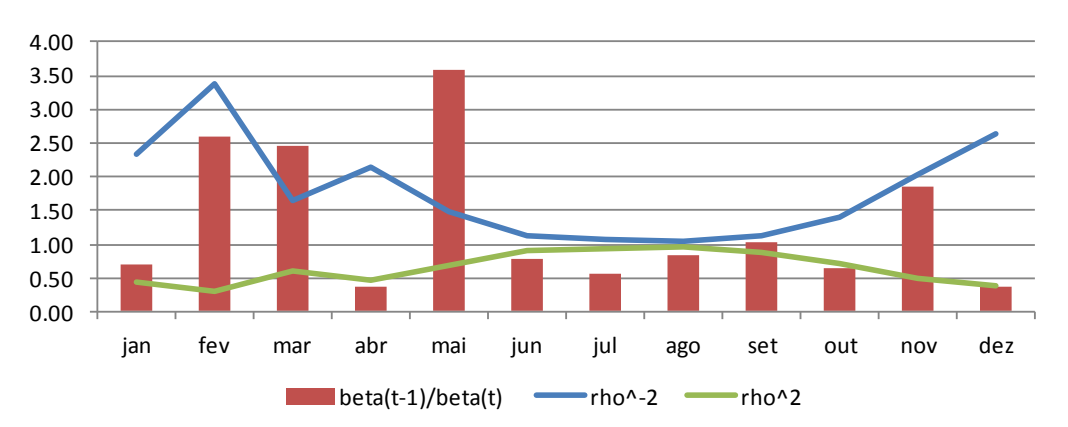


Figura 5.11 – Estudo da restrição (5.22a) para a série histórica do Subsistema Nordeste

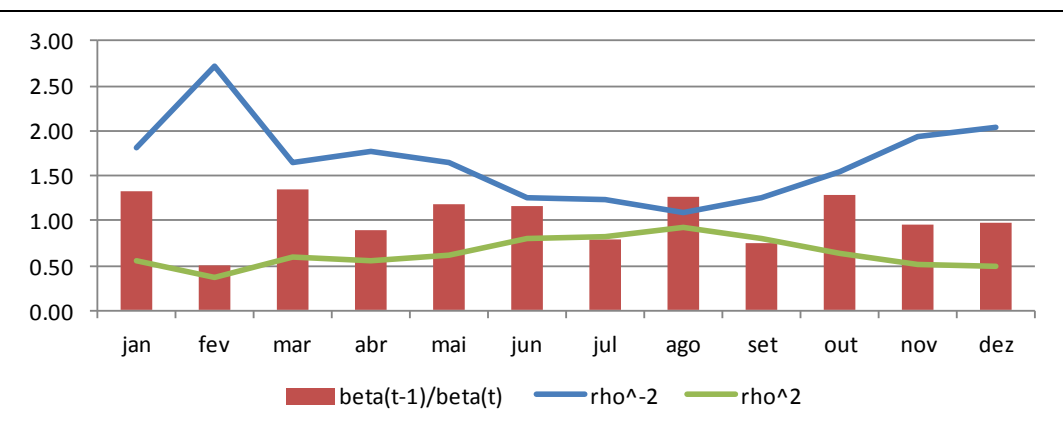


Figura 5.12 – Estudo da restrição (5.22a) para a série histórica do Subsistema Norte

Ao confirmar a violação da restrição (5.22a) e entendendo a conveniência de se ajustar um modelo Gama as séries históricas de energia, comprovada na seção (5.2.1) deste trabalho, pensou-se em maneiras de “corrigir” tal limitação. Em outras palavras, na tentativa de suplantar a restrição (5.22a) foram propostos e testados os seguintes ajustes à modelagem: (i) Estimar os parâmetros da distribuição Gama via Máxima Verossimilhança; (ii) Estimar o modelo Gama com a base de dados padronizada; (iii) Testar o valor de “k” (equação 5.12); (iv) Estimar o modelo Gama utilizando correções na estimação dos parâmetros proposta por (Fernandez & Salas, 1990); (v) Propor uma “correção” à restrição (5.22a); e, por fim, (vi) fazer “cortes” na série histórica de ENA.

Quanto a tentativa (i), apesar dos pontos discutidos em (5.1.3.2) de que a estimação via Método dos Momentos não traria nenhum “prejuízo” aos parâmetros aferidos, estimaram-se os parâmetros da distribuição Gama (parâmetros de escala e localização) via Máxima Verossimilhança (Hahn & Shapiro, 1994). Contudo, tal “correção” não proporcionou “ganhos” à geração dos cenários, mostrando-se, conforme já ressaltado por (Fernandez & Salas, 1986), não necessária. Além disso, os “novos” parâmetros estimados continuaram violando a restrição (5.22a).

Outras duas tentativas para corrigir a restrição (5.22a), implementadas neste trabalho foram, a estimação do modelo PGAR(1) com a base de dados padronizada e o teste de diversos valores para o “k” apresentado na equação (5.12) deste trabalho.

Quanto à estimação do modelo Gama com os dados padronizados, a mesma mostrou-se não factível uma vez que a padronização possibilitou o surgimento de

valores negativos, tornando-se inviável a estimação e a geração de ruídos Gama, uma vez que os mesmos são definidos positivos.

Com relação ao valor de “ k ” (equação 5.12), (Fernandez & Salas, 1986) e (Fernandez, 1984) apesar de ressaltarem que o valor desse parâmetro não é grande, não definem o seu valor real. Dessa forma, objetivando “varrer” todo o espaço paramétrico de “ k ”, testaram-se diversos valores para este parâmetro, concluindo-se que os melhores cenários são gerados quando “ k ” é igual a dois.

Dando sequência à tentativa de “melhorar” os cenários gerados pelo modelo Gama e, principalmente, atender a restrição (5.22a), optou-se por testar as correções nos parâmetros estimados propostas por (Fernandez & Salas, 1990). Neste artigo os autores propõem o uso do modelo Gama Autorregressivo (GAR(1)) e definem “correções” aos parâmetros de autocorrelação, variância e assimetria. Seguindo a mesma metodologia proposta por (Fernandez & Salas, 1990), testaram-se as correções ao modelo PGAR(1), contudo, as mesmas não resolveram os problemas na restrição (5.22a) e também não trouxeram nenhuma melhora significativa aos cenários.

Nesse contexto de violação da restrição, manipulou-se matematicamente a restrição (5.22a), de maneira que os parâmetros estimados se adequassem à mesma. Foram testadas duas correções, a primeira via β^{m,c_m} e a segunda via ρ_1^{m,c_m} . As correções aos parâmetros β^{m,c_m} podem ser observadas nas equações (5.27a) e (5.27b).

$$\beta^{m,c_m} = \rho_1^2 \beta^{m-1,c_{m-1}} \quad \left(\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \right) \geq (\rho_1^{m,c_m})^{-2} \quad (5.27a)$$

$$\beta^{m,c_m} = \rho_1^{-2} \beta^{m-1,c_{m-1}} \quad \left(\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \right) \geq (\rho_1^{m,c_m})^2 \quad (5.27b)$$

Outra alternativa foi “corrigir” o parâmetro ρ_1^{m,c_m} , conforme apresentado nas equações (5.28a) e (5.28b).

$$\rho_1^{m,c_m} = \sqrt{\frac{\beta^{m,c_m}}{\beta^{m-1,c_{m-1}}}} \quad \left(\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \right) \geq (\rho_1^{m,c_m})^{-2} \quad (5.28a)$$

$$\rho_1^{m,c_m} = \sqrt{\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}}} \quad \left(\frac{\beta^{m-1,c_{m-1}}}{\beta^{m,c_m}} \right) \geq (\rho_1^{m,c_m})^2 \quad (5.28b)$$

Após a formulação matemática das equações (5.27a), (5.27b), (5.28a) e (5.28b), testaram-se as séries sintéticas simuladas e o atendimento à restrição (5.22a). Conforme esperado, a identidade $\left(\frac{\beta^{m-1, c_{m-1}}}{\beta^{m, c_m}} \right)$ ficou dentro da região de admissibilidade do modelo, no entanto, ao contrário do esperado, os cenários continuaram, em alguns casos, voláteis e não respeitando os momentos amostrais do histórico.

Ainda neste contexto de violação da restrição e tentativa de melhorar os cenários gerados, optou-se por fazer “cortes” na série de ENA na tentativa de encontrar períodos históricos em que a mesma se encaixasse à restrição do modelo Gama proposto por (Fernandez & Salas, 1986). Conforme se pode observar na (

Figura 5.13), para o período de 1931 à 1980, o subsistema Sul atende à restrição imposta pelo modelo Gama. Atendida tal restrição, esperava-se séries sintéticas que “respeitassem” os momentos amostrais do histórico e menos voláteis, contudo, os cenários gerados para o subsistema Sul, estimando-se os parâmetros no período de 1931 à 1980, não se mostraram adequados.

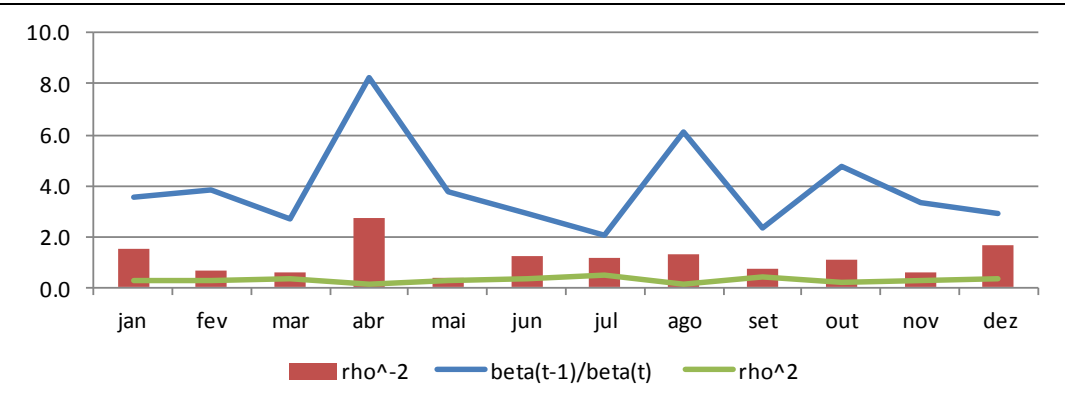


Figura 5.13 – Restrição (5.22a) corrigida para a série histórica do Subsistema Sul

Por fim, após diversas tentativas para atender a restrição (5.22a) imposta pelo modelo e, mais ainda, de tentar gerar “bons” cenários utilizando-se o modelo proposto por (Fernandez & Salas, 1986), apresenta-se nas figuras (Figura 5.14, Figura 5.15, Figura 5.16, Figura 5.17), os cenários gerados menos voláteis e mais próximos aos momentos amostrais do histórico. Os mesmos foram gerados seguindo os pressupostos e o algoritmo proposto por (Fernandez & Salas, 1986),

apenas corrigindo a restrição (5.22a) através da “correção” do parâmetro β^{m,c_m} proposta em (5.28a) e (5.28b).

As figuras (Figura 5.14, Figura 5.15, Figura 5.16, Figura 5.17) apresentam os seguintes resultados: (i) os cenários obtidos para cada um dos subsistemas e as médias destas séries em comparação à média histórica; (ii) gráficos comparativos entre os desvios-padrão dos cenários e da série histórica; (iii) teste de média dos cenários, isto é, por meio do *T-Test*, testou-se a hipótese da média dos 200 elementos de cada um dos 60 períodos da matriz de cenários ser igual à média das 79 observações de cada um dos meses da série histórica; e (iv) a evolução temporal dos *p-values* do teste de Levene para comparação de igualdade de variâncias.

Conforme se pode observar, de maneira geral, principalmente se comparado aos resultados do modelo PAR(p)-IC, os cenários gerados pelo modelo Gama são mais voláteis. Quanto aos testes e análises dos momentos amostrais, observa-se que os cenários gerados para os subsistemas Sudeste/Centro-Oeste e Norte apresentam médias diferentes àquelas verificadas pelo histórico, conforme mostrado pelo Teste *t* (o nível de significância adotado é de 5%, o valor crítico do *p-value* é representado por uma linha contínua), enquanto nos subsistemas Sul e Nordeste os cenários parecem convergir para a média histórica. Com relação ao desvio-padrão, observa-se que o mesmo está “bem representado” nos subsistemas Sul e Norte, mas no subsistema Sudeste/Centro-Oeste apresenta-se diferente do histórico (Figura 5.14, Figura 5.15, Figura 5.16, Figura 5.17).

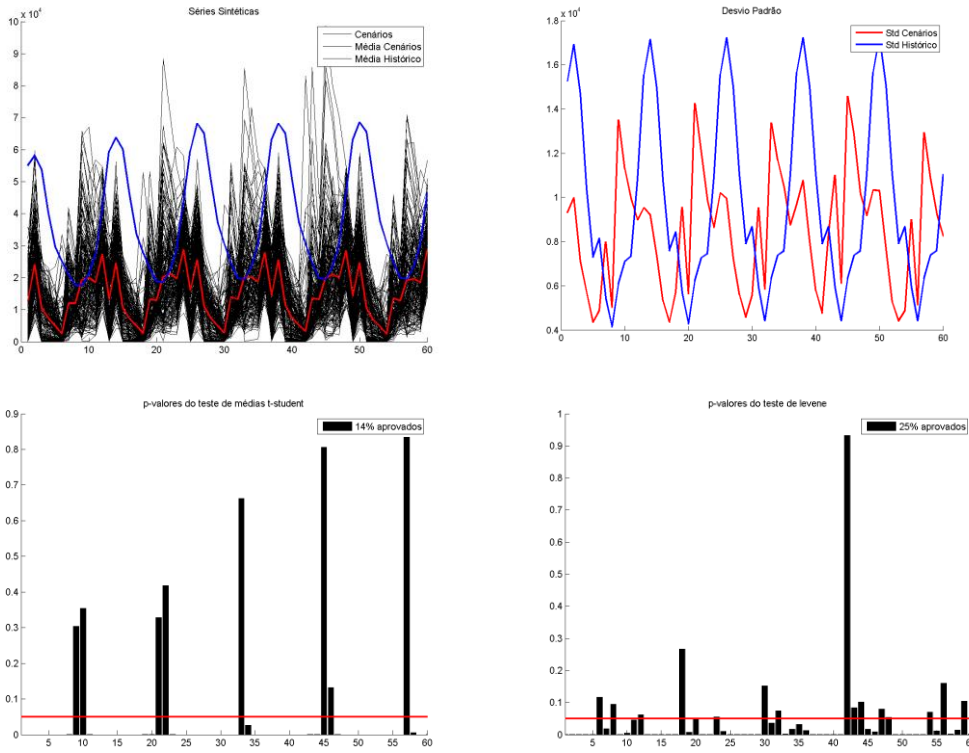


Figura 5.14 – Cenários e testes - Subsistema Sudeste/Centro-Oeste

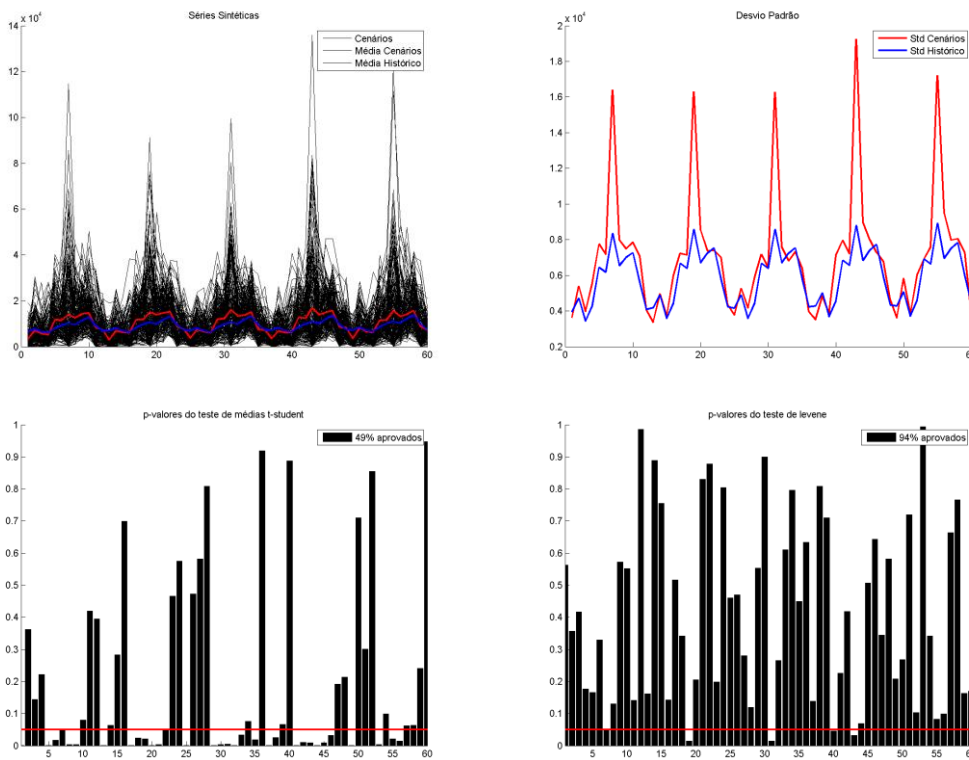


Figura 5.15 – Cenários e testes - Subsistema Nordeste

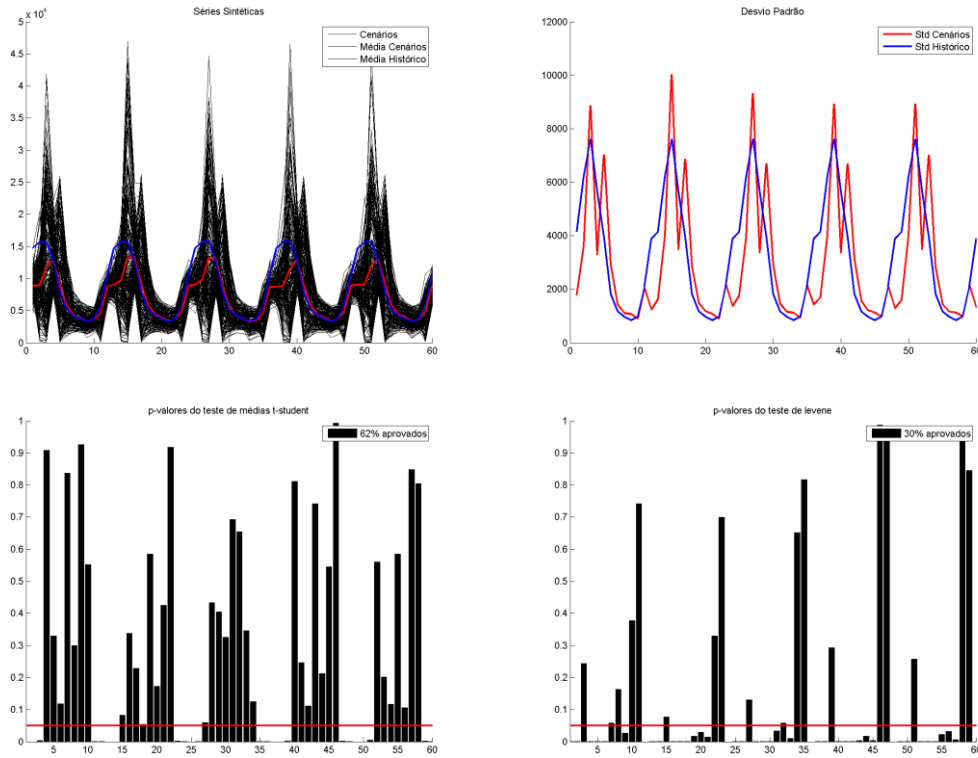


Figura 5.16 – Cenários e testes - Subsistema Nordeste

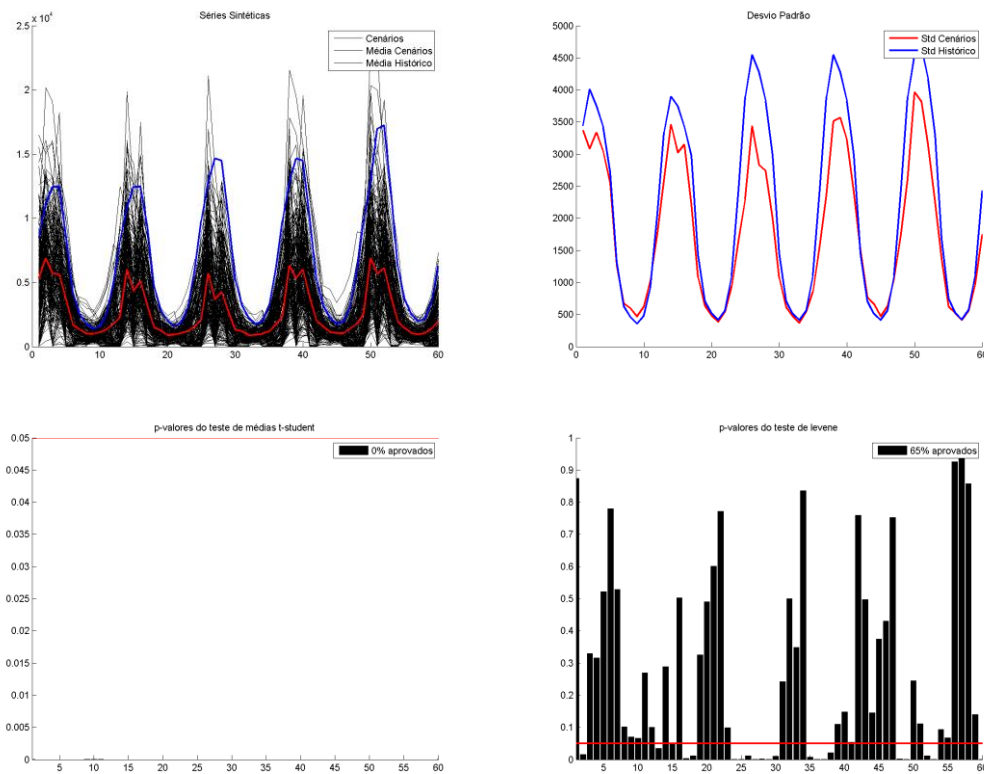


Figura 5.17 – Cenários e testes - Subsistema Norte

Após a discussão sobre a modelagem Gama, a análise às restrições impostas e a apresentação dos cenários gerados, torna-se importante levantar alguns pontos com relação ao exposto. Conforme observado anteriormente, os modelos Gama não geraram cenários tão “bons” quanto aos apresentados pelo modelo PAR(p)-IC (seção 4.2.2), contudo, também se observou a necessidade de uma modelagem Gama, comprovada pelos testes Qui-quadrado e pelo argumento de (Fernandez & Salas, 1986), os quais defendem a “robustez” matemática desse modelo.

Neste contexto, conclui-se que a utilização dos modelos Gama, da forma que foram propostos por (Fernandez & Salas, 1986) não podem ser usados pelo SEB, pois os resultados da geração dos cenários mostraram-se não factíveis.

Considerações finais e trabalhos futuros

A ideia dessa tese, basicamente, foi discutir três importantes tópicos relacionados ao Setor Elétrico Brasileiro, a saber, a estocasticidade intrínseca ao setor, as peculiaridades metodológicas que norteiam o módulo de geração de energias afluentes, peça chave às atividades de planejamento da expansão, operação e precificação de curto prazo, e a “atenção” para a necessidade de um modelo Gama, mais adequado às características da série histórica de energia a ser modelada.

Remetendo-se aos objetivos desse trabalho, observa-se que os mesmos foram essencialmente atendidos, isto é, foi apresentada ao leitor uma visão diferente do Setor Elétrico Brasileiro, evidenciando-se a intrínseca relação entre a estocasticidade e as atividades executadas pelo SEB, um importante setor de infraestrutura do país. Ressalta-se que as séries sintéticas de energia/vazão são cruciais para determinar qual é a melhor maneira de operar o setor, subsidiam decisões sobre quando se deve ou não expandi-lo, evitando-se assim custo e/ou perdas desnecessárias. E, ainda, são um fator preponderante na determinação do preço de curto prazo de energia elétrica, dado que a quantidade simulada/prevista de água nos reservatórios no futuro será um dos determinantes do preço da energia no curto prazo.

Seguindo a lógica proposta, apresentou-se o Módulo de Geração de energias Afluentes, destacando-se a sua contribuição ao Setor Elétrico Brasileiro, como o mesmo está ligado à PDDE e quais as “ressalvas” que esta ligação estabelece. Ainda neste tópico, levantaram-se algumas peculiaridades do módulo, como por exemplo, a questão das configurações intercorrelacionadas, algo inédito até então, e pontos que levantam discussão com relação à teoria de Séries Temporais, como por exemplo, a definição da ordem do modelo, a questão da estacionariedade e a necessidade de tratar possíveis *outliers*.

Quanto à metodologia adotada pelo Setor Elétrico Brasileiro para a geração de séries sintéticas discutiram-se pontos metodológicos importantes, com ênfase para a geração de séries sintéticas utilizando a distribuição lognormal, o

deslocamento adotado para evitar a geração de cenários negativos e os problemas que essa escolha estabelece, que são a não linearidade do PAR(p)-IC e a falta de uma robustez matemática do modelo, por exemplo, com relação aos parâmetros estimados, ressaltada por (Fernandez & Salas, 1986).

Ao unir as especificidades abordadas no parágrafo anterior com a demonstrada característica da série histórica de energia relacionada à distribuição Gama, abordou-se, seguindo as peculiaridades do SEB, uma nova metodologia que prescinde da normalidade e evita a possível geração de cenários negativos. Observando os problemas levantados, concluiu-se que a utilização dos modelos Gama, da forma que foram propostos por (Fernandez & Salas, 1986) não podem ser usados pelo SEB, pois os resultados da geração dos cenários mostraram-se não factíveis.

Este trabalho abre um leque de possibilidades e discussões sobre como está sendo feito o planejamento da operação, o planejamento da expansão e a determinação do preço spot da energia elétrica. Nesse sentido a ideia desse trabalho é ter como resultado intangível a criação de uma agenda de pesquisa sobre a modelagem estocástica que envolve o SEB.

Em outras palavras, é preciso discutir pontos de melhorias ao modelo adotado pelo setor, saber se o modelo hoje adotado é o mais adequado e discutir pontos relacionados aos efeitos das mudanças climáticas que vem ocorrendo ao longo dos anos no planeta, pois, conforme se evidenciou, mudanças no regime de chuvas podem provocar grandes danos ao SEB e conseqüentemente, prejuízos a economia brasileira.

Assim, a ideia dessa tese, muito mais que apresentar a relação estocasticidade vs SEB, apresentar o modelo PAR(p)-IC e propor a modelagem Gama, é levantar um problema real, onde os *policymakers* devem estar atentos não apenas aos problemas de modelagem que hoje existem, mas também, devem pensar quais serão as atitudes do País, caso algum efeito climático afete de uma maneira “não previsível” as ENAs de um determinado subsistema.

Nesse sentido, apesar do amplo leque de discussões abordado neste trabalho, uma intensa agenda de pesquisa surge e deve ser colocada em prática. Citam-se, por exemplo, a questão da não linearidade gerada pela adoção da lognormal deslocada para evitar a geração de séries sintéticas negativas. Conforme foi observado, o modelo Gama proposto transcende a questão da ENA

negativa, no entanto não resolve o problema da não convexidade ao ser acoplado à PDDE. Algumas soluções já foram pensadas, como uma expansão de Taylor ao modelo Gama, no entanto precisam ser colocadas em prática.

Outro trabalho de importante execução é estimar as funções de custo futuro geradas pelo modelo NEWAVE utilizando as séries sintéticas geradas pelo modelo Gama. Nessa mesma linha, também vale a pena testá-las em um modelo de Programação Dinâmica Estocástica (PDE), o qual prescinde a necessidade de linearidade do modelo.

Uma terceira e importante agenda de trabalho é pesquisar outros modelos que utilizam outras distribuições, como a Weibull, na geração de séries sintéticas. Mesmo ressaltando que após uma vasta pesquisa bibliográfica não foi encontrado nenhum trabalho nessa direção.

Outra extensão deste trabalho é o ponto de atenção que se colocou à necessidade de se adotar modelos que consigam capturar efeitos ocasionados pelas mudanças climáticas. Conforme se observou, o Setor Elétrico Brasileiro é amplamente dependente das vazões e todo seu planejamento está calcado em modelos que as utilizam previstas e/ou simuladas no longo prazo. Dessa forma, entender como os efeitos climáticos podem afetar os três pilares do setor elétrico é fundamental.

Referências

ABINEE. (2012). *Desempenho Setorial*. Acesso em 15 de February de 2012, disponível em Associação Brasileira da Indústria Elétrica e Eletrônica: <http://www.abinee.org.br>

Akaike, H. (1987). Factor Analysis and AIC. *Psychometrika* 52(3), pp. 317–332.

Bacanli, U. G. (2012). Stochastic Modeling of Monthly Streamflow Data. *Balwois, Ohrid, Republic of Macedonia*.

Ballini, R., França, E., Kadowaki, M., Soares, S., & Andrade, M. G. (1997). Modelos de Redes Neurais e Box&Jenkins para Previsão de Vazões Médias Mensais. *Proceedings of the XII Simpósio Brasileiro de Recursos Hídricos. Vitória, Brasil*.

Barros, M., Mello, M. F., & Souza, R. C. (2009). Aquisição de energia no mercado cativo brasileiro: simulações dos efeitos da regulação sobre o risco das distribuidoras. *Pesquisa Operacional* 29, pp. 303-322.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society*, 27, pp. 27-41.

Beard, L. R. (1965). Use of interrelated records to simulate streamflow. *Journal of Hydrology. ASCE*, pp. 13-22.

Beard, L. R. (1967). Hydrologic Simulation in water yield analysis. *Hydrology Engineering Center. US Army Corps of Engineers*.

Benders, J. F. (1962). Partitioning Procedures for Solving Mixed Variables Program-ming Problems. *Numerische Mathematik*, 4, pp. 238-252.

BNDES. (2012). *Arquivo 2012*. Acesso em 15 de February de 2012, disponível em Banco Nacional de Desenvolvimento Econômico e Social: <http://www.bndes.gov.br/>

Bobé, B., & Robitaille, R. (1975). Correction of bias in the estimation of the coefficient of skewness. *Water Resources Research*, pp. 851-854.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B (Methodological)*, 26, pp. 211-252.

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series analysis: forecasting and control*. Willey.

Campos, L. C. (2010). Modelo Estocástico Periódico baseado em Redes Neurais. *Tese de Doutorado, DEE, PUC-Rio, Brasil*.

- Canuto, O. (2000). A Crise Asiática e seus desdobramentos. *Revista Econômica* 4.
- Casella, G., & Berger, R. L. (2001). *Statistical Inference*. Duxbury, 2nd ed.
- Castro, N. J., & Leite, A. L. (2010). Preço spot de eletricidade: teoria e evidências do caso brasileiro. *IV Encontro de Economia Catarinense, Criciúma. IV Encontro de Economia Catarinense*.
- CCEE. (2012). *Arquivo 2012*. Acesso em 15 de March de 2012, disponível em Câmara de Comercialização de Energia Elétrica: <http://www.ccee.org.br>
- CEPEL. (2011a). *Manual de referência do modelo NEWAVE*. Rio de Janeiro, Brazil: Centro de Pesquisas em Energia Elétrica.
- CEPEL. (2011b). *Centro de Pesquisas de Energia Elétrica*. Acesso em 12 de July de 2011, disponível em <http://www.cepel.br/>
- Charbeneau, R. J. (1978). Comparison of the Two and Three-Parameter Log Normal Distributions Used in Streamflow Synthesis. *Water Resources Research*, 14(1), pp. 149-150.
- Conte, S. D. (1965). *Elementary Numerical Analysis*. MacGraw-Hill.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74, pp. 427-431.
- ELETROBRÁS. (2011). *Centrais Elétricas brasileiras*. Acesso em 21 de March de 2011, disponível em <http://www.eletronbras.gov.br/>
- Engle, R., & Ng, V. (1991). Measuring and testing the impact of news on volatility. *Journal of Finance*.
- EPE. (2009). *Atualização do valor para patamar único de custo de déficit*. Rio de Janeiro, Brazil: Empresa de Pesquisa Energética.
- EPE. (2012). *2012 Report*. Acesso em 15 de February de 2012, disponível em Empresa de Pesquisa Energética: <http://www.epe.gov.br/>
- Fernandes, E., Fonseca, M., & Alonso, P. (2005). Natural gas in Brazil's energy matrix: demand for 1995–2010 and usage factors. *Energy Policy* 33, pp. 365-386.
- Fernandez, B. (1984). Stochastic modeling of periodic streamflow series with gamma distributions. *Ph.D. dissertation, Colorado State University., Fort Collins*.
- Fernandez, B., & Salas, J. (1990). Gamma-Autoregressive Models for Stream-Flow Simulation. *Journal of Hydraulic Engineering* 116.
- Fernandez, B., & Salas, J. D. (1986). Periodic Gamma Autoregressive Processes for Operational Hydrology. *Water Resources Research*, 22, pp. 1385-1396.

Ferreira, P. G., Oliveira, F. L., & Souza, R. C. (2012). Autoregressive Periodic Model applied to the generation of synthetic scenarios, incorporating the impact of Climate Phenomena through SOI index. *10th International Conference on Operations Research, Havana*.

Finardi, E. C., Gonçalves, R. E., Garcia, A. G., Matos, V. L., Takigawa, F. Y., & Zucarato, A. N. (2009). *Investigações de propostas metodológicas nos modelos de precificação para minimizar a volatilidade do preço de liquidação de diferenças: Possíveis problemas nos cortes de Benders*. Florianópolis, Brazil: Tractebel Energia, Labplan (UFSC).

Gaver, D. P., & Lewis, P. A. (1980). First-order autoregressive gamma sequences and point processes. *Adv. Appl. Probab.*, pp. 727-745.

Goldemberg, J., & Prado, L. T. (2003). Reforma e crise do setor elétrico no período FHC. *Tempo Social 15 (2)*. USP. São Paulo.

Gujarati, D. N. (2006). *Econometria Básica, 4 ed.* São Paulo: Makron Books.

Hahn, G. J., & Shapiro, S. S. (1994). *Statistical Models in Engineering*. Hoboken, NJ: John Wiley & Sons, Inc.

Hamilton, J. D. (1994). *Times Series Analysis*. Princeton University Press, 1^a edition.

Hannan, E. J. (1955). A test for singularities in Sydney rainfall. *Aust. J. Phys.*, pp. 289-297.

Hipel, K. W., & McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*. Amsterdam, The Netherlands: Elsevier.

Hirsch, R. M. (1979). Synthetic hydrology and water supply reliability. *Water Resouces Research*, pp. 1603-1615.

Jarque, C., & Bera, A. K. (1981). Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters*.

Jones, R. H., & Brelford, W. M. (1967). Time Series with periodic structure. *Biometrika 54*, pp. 403-408.

Justel, A., Peña, D., & Tsay, R. S. (2001). Detection of outliers patches in autoregressive time series. *Statistics Sinica 11*, pp. 651-673.

Kelman, J. (2001). *Relatório da Comissão de Análise do Sistema Hidrotérmico de Energia Elétrica*. Brasília, Brazil.

Kelman, J., & Pereira, M. V. (1977). Critérios de avaliação para modelos de séries hidrológicas. *IV SNPTEE*.

Kite, G. W. (1977). Frequency and Risk analysis in hydrology. *Water Resources Publications*.

Lawrance, A. J. (1982). The innocation distribution of a Gamma distributed autoregressive process. *Scand. J. Stat.* 9, pp. 234-236.

Lawrance, A. J., & Lewis, P. A. (1981). A new autoregressive time series model in exponential variables (NEAR(1)). *Adv. Appl. Probab.*, pp. 826-845.

Lepecki, J., & Kelman, J. (1985). Brazilian Hydroelectric System. *Water International*, 10(4).

Lima, C. H., & Lall, U. (2010). Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model. *Journ al of Hydrology* 380, pp. 438–449.

Lumsdaine, R. L., & Serena. (1993). Testing for ARCH in the presence of a possibly misspecied conditional mean. *Journal of Econometrics*.

Lund, R., Shao, Q., & Basawa, I. (2006). Parsimonious Periodic Time Series Modeling. *Australian & New Zealand Journal of Statistics*, 48, pp. 33-47.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. New York: Springer-Verlag.

Maceira, M. E., & Damázio, J. M. (2004). The use of PAR(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the Brazilian hydropower system. *8th International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, Iowa*.

Maceira, M. E., Damázio, J. M., Ghirardi, A. O., & Dantas, H. M. (1999). Periodic ARMA Models Applied to weekly streamflow ferecasts. *IEEE Power Tech 99 Conference. Budapest, Hungary*.

Maceira, M. E., Penna, D. D., & Damazio, J. M. (2005, November 20-24). Geração de Cenários Sintéticos de Energia e Vazão para o Planejamento da Operação Energética. *XVI Simpósio Brasileiro de Recursos Hídricos*.

Maceira, M. E., Terry, L. A., Costa, F. S., Damázio, J. M., & Melo, A. C. (2002). Chain of Optimization models forSetting the Energy Dispatch and Spot price in the Brazilian System. *14th PSCC*.

Marcato, A. L. (2002). Representação híbrida de sistemas equivalentes e individualizados para o planejamento da operação a médio prazo de sistemas de potência de grande porte. Rio de Janeiro: Tese de Doutorado, DEE, PUC-Rio.

Marien, J. L. (2001). *The Use of the PAR(p) model in the Stochastic Dual Dynamic Programming Optimization scheme used in the Operation of the Brazilian Hydropower System*. Belgium: Report of consulting assignment.

Marreco, J. M. (2007). Planejamento de Longo Prazo da Expansão da Oferta de Energia Elétrica no Brasil sob uma Perspectiva da Teoria das Opções Reais. *Tese de Doutorado. Universidade Federal do Rio de Janeiro*.

Matalas, N. C. (1967). Mathematical assessment of synthetic hydrology. *Water Resource Research*, pp. 937-947.

McKenzie, E. (1982). Product Autoregression: A Time-Series characterization of the Gama distribution. *J. Appl. Prob.* 19, pp. 463-468.

Mckenzie, E. (1982). Product Autoregression: A Time-Series Characterization of the Gamma Distribution. *J. Appl. Prob.*

Mcmahon, T. A., & Miller, A. (1971). Application of the Thomas and Fiering model to skewed hydrologic data. *Water Resources Research*, pp. 1338-1340.

Ministério de Minas e Energia. (2013). Acesso em 2013, disponível em <http://www.mme.gov.br/mme>

Miranda, C. F. (2001). Modelação linear de séries temporais na presença de outliers. *Tese de Mestrado - Departamento de Matemática Aplicada, Universidade do Porto.*

Neto, A. C., & Souza, R. C. (1996). A Bootstrap Simulation Study in ARMA (p,q) Structures. *Journal of Forecasting*, pp. 343-353.

OECD. (2012). *2011 Report*. Retrieved February 15, 2012, from Organisation for Economic Co-operation and Development: <http://www.oecd.org/home/>

Oliveira, F. L. (2010). Nova abordagem para geração de cenários de afluências no planejamento da operação energética de médio prazo. Rio de Janeiro: Dissertação de Mestrado, DEE, PUC-Rio.

Oliveira, F. L., & Souza, R. C. (2011). A new approach to identify the structural order of par (p) models. *Pesquisa Operacional*, 31(3), pp. 487-498.

ONS. (2012). www.ons.org.br. Acesso em Agosto de 2012

Ota, R. (1996). Valores Aberrantes em Séries Temporais: Teste de detecção e Efeito na previsão de valores agregados. *Mestrado em Estatística, UNICAMP.*

PAC. (2012). *2011 Report*. Retrieved February 15, 2012, from Programa de Aceleração do Crescimento: <http://www.brasil.gov.br/pac>

Pagano, M. (1978). On periodic and multiple autorregressions. *The Annals of Statistics*, 6(6), pp. 1310-1317.

Pêgo, B., & Campos Neto, C. A. (2008). O PAC e o Setor Elétrico: Desafios para o abastecimento do mercado brasileiro (2007-2010). *Texto para Discussão nº 1329. IPEA.*

Penna, D. D. (2009). Definição da árvore de cenários de afluências para o planejamento da operação energética de médio prazo. Rio de Janeiro: Tese de Doutorado, DEE, PUC-Rio.

Pereira, G. A. (2011). Modelos de Memória longa para geração de cenários hidrológicos sintéticos. *Dissertação de Mestrado, DEE, PUC-Rio, Brasil.*

Pereira, M. V. (1989). Optimal Stochastic Operations Scheduling of Large Hydroelectric Systems. *International Journal of Electric Power and Energy Systems, 11(3)*, pp. 161-169.

Pereira, M. V., & Pinto, L. M. (1982). Decomposition Approach to the Economic Dispatch of Hydrothermal Systems. *WM.*

Pereira, M. V., & Pinto, L. M. (1983). Application of decomposition techniques to the mid and short term scheduling of hydrothermal systems. *IEEE Transactions on Power Apparatus and Systems(11)*.

Pereira, M. V., & Pinto, L. M. (1985). Stochastic optimization of multireservoir hydroelectric system: A decomposition approach. *Water Resources Research, 21(6)*.

Pereira, M. V., & Pinto, L. M. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming.*

Pereira, M. V., Campodónico, N., & Kelman, R. (1998). Long-term Hydro Scheduling based on Stochastic Models. *EPSOM'98.*

Phillips, P. C., & Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika 75*, pp. 335-346.

Pires, J. C., Giambiagi, F., & Sales, A. F. (2002). As perspectivas do setor elétrico após o Racionamento. *Texto para Discussão nº 97. BNDES.*

PREVIVAZ. (2000). Manual de previsão de vazões semanais.

Rasmussen, R. F., Salas, J., Fagherazzi, L., Rassam, J. C., & Bobee, R. (1996). Estimation and validation of contemporaneous PARMA models for streamflow simulation. *Water Resources Research, 32(10)*, pp. 3151-3160.

Rodrigues, R. D. (2007). Gerenciamento de Risco no Setor Elétrico Brasileiro através do uso de derivativos. *Master Thesis. Federal University of Rio de Janeiro.*

Salas, J. D. (2012). *Books and Chapters of Books.* Fonte: <http://www.engr.colostate.edu/~jsalas/publications.html>

Salas, J. D., & Obeysekera, J. T. (1982). ARMA Model Identification of Hydrologic Time Series. *Water Resources Research, 18*, pp. 1011-1021.

Salas, J. D., & Smith, R. A. (1981). Computer Programs of Probability Distribution in Hydrology. *Department of Civil Engineering, Colorado State University, Fort Collins.*

Salas, J. D., Boes, D. C., & Smith, R. A. (1982). Estimation of ARMA models with seasonal parameters. *Water Resources Research, 18(4)*, pp. 1006-1010.

- Salas, J. D., Sveinsson, O. G., Lane, W. L., & Frevert, D. K. (2007). Stochastic Analysis, Modeling, and Simulation (SAMS). *Colorado State University*.
- Sarnaglia, A. J., Oliveira, R. L., Coser, M., & Reisen, V. A. (2008). Estudo da Vazão Média Mensal do Rio de Castelo-ES: Uma Aplicação do Modelo PAR.
- Schaeffer, R., Cohen, C., Almeida, M., Achão, C., & Cima, F. (2003). Energia e pobreza: problemas de desenvolvimento energético e grupos sociais marginais em áreas rurais e urbanas do Brasil. *Chile 2003*.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, Hayward*, v.6, pp. 461-464.
- Shao, Q., & Lund, R. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *Journal of Time Series Analysis*, pp. 359-372.
- Shapiro, A., Tekaya, W., Costa, J. P., & Soares, M. P. (2012). Risk neutral and risk averse Stochastic Dual Dynamic Programming method. *European Journal of Operational Research*.
- Souza, R. C., Marcato, A. L., Dias, B. H., & Oliveira, F. L. (2012). Optimal Operation of Hydrothermal Systems with Hydrological Scenario Generation through Bootstrap and Periodic Autorregressive Models. *European Journal of Operational Research*, 222(3), pp. 606-615.
- Srikanthan, R., & McMahon, T. A. (1978). A review of lag-one Markov models for generation of annual flows. *Journal of Hydrology*.
- Stedinger, J. R. (1981). Estimating correlations in multivariate streamflow models. *Water Resource Research*, pp. 200-208.
- Stedinger, J. R., & Taylor, M. R. (1982). Synthetic Streamflow Generation: Model Verification and Validation. *Water Resource Research* 18, pp. 909-918.
- Stedinger, J. R., Lettenmaier, D. P., & Vogel, R. M. (1985). Multisite ARMA(1,1) and Disaggregation Models for Annual Streamflow Generation. *Water Resources Research* 21, pp. 497-509.
- Terry, L. A., Pereira, M. V., Neto, T. A., Silva, L. F., & Sales, P. R. (1986). Coordinating the Energy Generation of the Brazilian National Hydrothermal Electrical Generating System. *Interfaces*, 16.
- Tesfaye, Y. G. (2005). Seasonal Time Series Models and their application to the modelling of river flows. *Dissertation of Doctor of Philosophy in Hydrology. University of Nevada*.
- Thomas, H. A., & Fiering, M. B. (1962). Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. *Design of Water Resource Systems*, pp. 459-463.

Tolmasquim, M. T. (2011). *Novo Modelo do Setor Elétrico Brasileiro*. Brasília, Brazil: Editora Synergia.

Triola, M. F. (2005). *Introdução à Estatística*. LTC.

Vecchia, A. V. (1985a). Maximum likelihood estimation for periodic autorregressive moving average models. *Technometrics*, 27, pp. 375-384.

Vecchia, A. V. (1985b). Periodic autoregressive moving average (PARMA) modeling with application to water resources. *Water Resources Bulletin*, 21, pp. 721-730.

Vieira, A. M., Pereira, B. B., & Sales, P. R. (1998). Estimação conjunta dos parâmetros de um modelo estocástico multivariado contemporâneo periódico autorregressivo - PAR(p). *RBRH - Revista Brasileira de Recursos Hídricos* 3.